Characterizing Merkel Cell Polyomavirus Life Cycle Regulators and Viral Encoded Circular RNAs

By

Hongzhao Zhou

Bachelor of Science, China Agricultural University, 2015

Master of Science, Georgetown University, 2016

Submitted to the Graduate Faculty of the

School of Medicine in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2023

UNIVERSITY OF PITTSBURGH

SCHOOL OF MEDICINE

This dissertation was presented

by

Hongzhao Zhou

It was defended on

July 31, 2023

and approved by

Haitao Guo, Professor, Department of Microbiology and Molecular Genetics

Saleem A. Khan, Professor, Department of Microbiology and Molecular Genetics

Kathy H.Y. Shair, Assistant Professor, Department of Microbiology and Molecular Genetics

Masahiro Shuda, Assistant Professor, Department of Microbiology and Molecular Genetics

Dissertation Director: Patrick S. Moore, Professor, Department of Microbiology and Molecular Genetics

Characterizing Merkel Cell Polyomavirus Life Cycle Regulators and Viral Encoded Circular RNAs

Hongzhao Zhou, PhD

University of Pittsburgh, 2023

This dissertation, initiated in 2018, delves into the characterization of viral circular RNAs (circRNAs) and investigates the regulators of the Merkel Cell polyomavirus (MCV) life cycle. In Chapter 2, we examined Kaposi's sarcoma-associated herpesvirus (KSHV) circRNAs, which are expressed in infection-associated diseases. The expression of these circRNAs fluctuates based on the viral life cycle, and they are incorporated into viral particles for pre-emptive delivery. This hints at their potential function in the early phases of virus infection. Chapter 3 discusses our exploration of circMCV-T, which reveals the dynamic interplay between MCV linear RNA, circRNA, and miRNA. This insight deepens our understanding of the intricate relationships between these different RNA types in the context of MCV infection. Finally, in Chapter 4, we explore several of MCV's self-encoded viral restriction mechanisms, designed to favor viral latency over lytic replication. Our findings shed light on these intricate, latency-promoting mechanisms, which can now be studied using advanced recombinase technology. Overall, the work comprising this dissertation significantly enhances our understanding of viral circRNAs and elucidates the intricate interactions of various RNA types within the MCV life cycle. Furthermore, it has facilitated the development of effective tools for MCV research. As such, this work not only deepens our knowledge in this field but also opens up promising avenues for future exploration.

Table of Contents

List of Tables

List of Figures

Acknowledgments

The past half-decade of my life has unfolded as an enthralling odyssey, abundant with a myriad of experiences, engaging interactions, and priceless lessons. This journey, despite being laden with formidable challenges, has acted as a catalyst, nurturing my growth and resilience, and honing me into a more dauntless individual. Throughout this period, I've had the honor of acquiring and mastering a wide array of skills, extending my intellect and broadening my horizons. The accomplishment of various objectives, both large and small, has fortified my confidence, strengthening my faith in my capabilities. This voyage required adaptability and constant personal development - attributes I have learned to value and refine.

This transformative stage in my life has also been distinguished by the creation of enduring relationships and friendships. I've been privileged to meet and gain knowledge from a diverse group of people - mentors, colleagues, and friends. Each one has contributed distinctively to my odyssey, offering guidance and support as I traversed the intricate maze of graduate school. Their priceless advice, steadfast support, and belief in me have been critical in guiding me to my current position. My gratitude for their contribution is beyond measure. Though it's impossible to individually acknowledge each individual who has influenced my life during this time, I want them to know that their impact is greatly valued. The inability to name everyone doesn't diminish the deep gratitude I feel. Each one of you holds a special place in my journey and in my heart, and for that, I offer my sincerest appreciation. Your generosity, support, and fellowship have left an indelible mark on the course of my life, and for that, I am forever grateful. Thank you for being a part of this extraordinary chapter of my life.

In the forefront, I would like to express my heartfelt gratitude to my mentors, Drs. Patrick Moore and Yuan Chang, for their unyielding support throughout my research and writing endeavors. Their expertise, patience, and guidance have proven invaluable, and their encouragement has inspired me to fulfill my potential. I also seize this opportunity to express my profound appreciation to all members of the CM Lab, past and present. Their collective wisdom, cooperation, and unwavering support have been invaluable in the completion of this dissertation. I want to particularly express my gratitude to Dr. Bizunesh Abere, Dr. Lindsey Robinson-McCarthy, Dr. James Li Wan, Dr. Jennifer Alvarez-Orellana, Dr. Tuna Toptan, and Dr. Justin Wendzicki for their continual assistance, advice, and instruction over the years. My thanks also extend to Felicia Steele and Emily Pellegrini for their exceptional administrative work.

Serving at the University of Pittsburgh, particularly at the Hillman Cancer Center, has been a truly warm experience. I must thank other faculty members for fostering such a collaborative research atmosphere, especially our collaborators Dr. Robin Lee and Dr. Donna Stolz, as well as my committee members, Dr. Haitao Guo, Dr. Saleem Khan, Dr. Masahiro Shuda, and Dr. Kathy Shair. Without your guidance, reaching this point would have been unattainable.

I also extend my gratitude to the current and former directors of the MGDB program, Dr. Arjumand Ghazi, Dr. Kara Bernstein, and Dr. Michael Tsang. I must acknowledge the work of our program coordinator, Kristin DiGiacomo, who has ensured my life as a graduate student progressed as smoothly as possible.

Finally, I want to express my deepest appreciation to my family and friends, who have remained a constant pillar of support throughout the past five years. I am grateful to my parents for their encouragement during moments of adversity. I also owe a great deal of gratitude to my girlfriend, Jiali Lyu, for her emotional backing. Moreover, I'd like to extend my thanks to my

friends. Listed in the order we met, Mingde Liu, Yue Xin, Xingzhe Song, Shiying Liu, Lulu Shao, Housayin Li, Shenyu Sun, Veronica Jinghui Li, Simon Cao, Stella Siying Guo, Wenyu Peng, and Victor So, each one of you has lent your kindness and support in unique ways. Your presence has illuminated my life's journey. With all of you, life indeed becomes brighter.

I sincerely appreciate each and every one of you for all you've done for me. Your support and kindness truly mean the world to me.

Sincerely,

Hongzhao Zhou

周泓兆

1.0 Introduction

1.1 Viruses and Tumors

1.1.1 The Foundation of Tumor Virology

Cancer is the second leading cause of death in the United States. Although it is well accepted now that viruses cause a subset of these cancers, this association was not appreciated until the early twentieth century (**Figure 1**). In 1908, Vilhelm Ellermann and Olaf Bang identified the avian leukosis virus (ALV) as the first virus capable of transformation [1]. ALV received little attention in the field of virology initially. In 1910, Peyton Rous discovered a filterable agent that causes transmissible chicken sarcoma, which was later named as Rous sarcoma virus (RSV) [2]. Following the identification of this second example of avian tumor viruses, scientists began to consider the possibility of viruses causing tumors in mammals. In 1933, Richard Shope isolated the cottontail rabbit papillomavirus, which caused papillomas in rabbits [3]. Then, in the early 1950s, Ludwik Gross discovered a retrovirus and, at the same time with Sarah Stewart, first isolated a polyomavirus that cause murine tumors [4, 5]. The discovery of simian vacuolating virus 40 (SV40) in 1960 by Ben Sweet and Maurice Hilleman, a contaminant in vaccines produced in rhesus monkey kidney cells, re-invigorated the study of tumor viruses since studies of this virus showed potent transformation potential in rodent cells [6]. Although SV40 is now known to be incapable of replicating in human cells, it has fueled research in tumor virology as well as cell biology (see Section 1.2.1) [7].

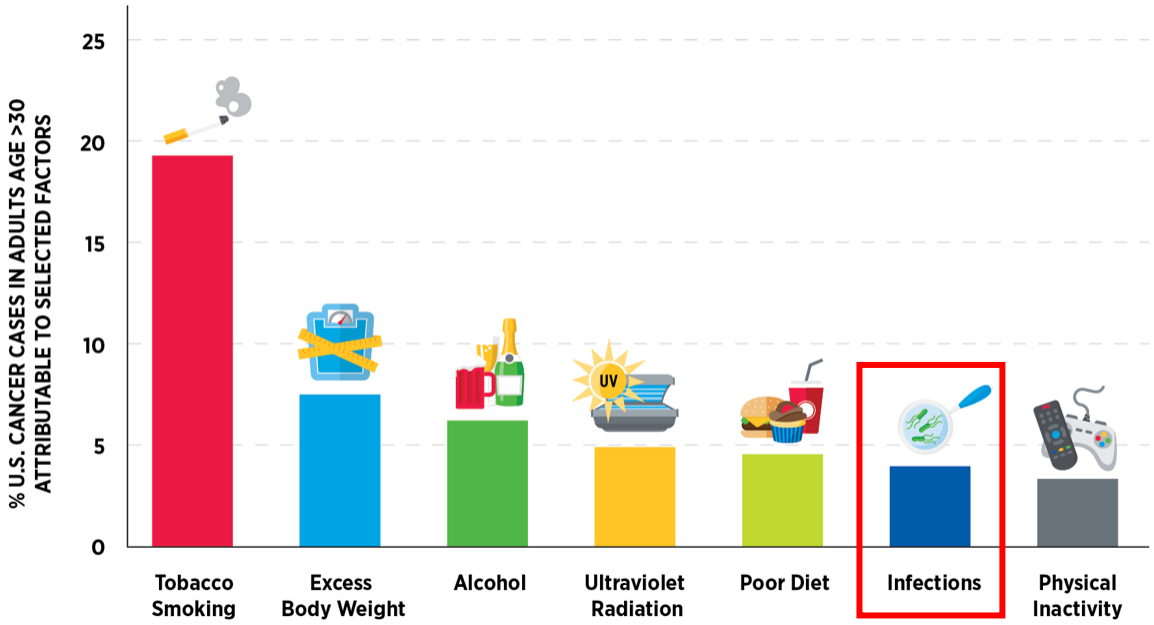


Figure 1 U.S. Cancer Causes in 2022.

Smoking, obesity, alcohol, ultraviolet radiation, poor diet, infections, and physical inactivity are major risk factors causing cancer in U.S. (Adopted from American Association for Cancer Research (AACR) Cancer Disparities Progress Report 2022)

1.1.2 Discovery of Human Tumor Viruses

The discovery of animal tumor viruses raised concern about the possibility of viruses causing human tumors. Since 1964, seven human tumor viruses have been identified (**Table 1**). In that year, the first human tumor virus, Epstein-Barr virus (EBV), was identified by Michael Anthony Epstein, Bert Achong, and Yvonne Barr from Burkitt's lymphoma through electron microscopic visualization [8]. EBV is now classified as a human γ-herpesvirus and has been found to infect more than 90% of adults worldwide. Despite EBV's significant ability to transform B lymphocytes *in vitro*, its widespread, lifelong, asymptomatic infection only causes cancer in a low percentage of infected individual [9, 10]. This suggests other risk factors in addition to EBV to

initiate or maintain transformation *in vivo*. Subsequent decades of studies have demonstrated that EBV encodes biomolecules including proteins and non-coding RNAs that play important roles in the pathogenesis of cancer and immune evasion [11-17].

In 1965, the second human tumor virus, hepatitis B virus (HBV), was identified by Baruch Blumberg through serologic screening [18]. By examining thousands of blood samples, an antigen from blood of an Aboriginal Australian (Au antigen), which reacted with an antibody in the serum of an American hemophilia patient, was found to be a surface protein of HBV. The sequelae of acute infection by HBV in the majority of cases is asymptomatic and the virus is cleared by the host immune system. In ~5% of patients chronic HBV infection develops due to a failure to eliminate the virus. However, scientists did not establish a relationship between chronic HBV infection and hepatocellular carcinoma (HCC) until 1975 [19]. This finding was later verified by a study in Taiwan, which showed that individuals with chronic HBV infection have a significantly increased risk of developing HCC by ~200 fold [20]. The HBV vaccine, the first human tumor virus vaccine, was developed by Maurice Hilleman and colleagues at the Merck pharmaceutical company the following year [21]. After the discovery of HBV, another virus causing hepatitis, hepatitis A virus (HAV) was identified as a cause of post-transfusion hepatitis, but not associated with liver cancer. However, a subset of individuals with non-A, non-B transfusion related hepatitis lead to the discovery of another human cancer virus [22]. In 1989, with the aid of advanced molecular biological techniques, Qui-Lim Choo, Michael Houghton, and their colleagues generated a cDNA library from a patient diagnosed with non-A, non-B hepatitis, and an infected chimpanzee model was used to increase the detectable viral sequences. From the cDNA library , a novel antigen was expressed by lambda phage, which was found to be the antigen encoded by the hepatitis C virus (HCV), a causative agent for HCC, much like HBV [23].

In 1980, Robert C. Gallo and colleagues identified human T-cell lymphotropic virus-I (HTLV-I) in cultured human T-cell lymphoma cells from a patient with cutaneous adult T-cell lymphoma (ATL), which was the first human retrovirus discovered [24]. One year later, Yorio Hinuma, Isao Miyoshi, and colleagues suggested that HTLV-I as the cause of ATL since only ATL patients have antibodies that target HTLV-I-infected T cells [25]. Similar to infection with EBV, only 5% of HTLV-I carriers develop ATL, again suggesting that risk factors in addition to viral infection are required to cause cancer [26]. Three years after the discovery of HTLV-I, Harald zur Hausen's group identified strain specific association of human papillomaviruses (HPV) DNA in cervical cancer [27, 28]. These two HPV strains, HPV16 and HPV18, which cause ~70% of cervical cancers worldwide, are now classified as high-risk HPVs along with HPV 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, and 68 [29]. In addition to cervical cancer, HPV is also linked to several other human cancers including over 72% of head-and-neck cancer, 70% of vaginal and anal cancers, 40% of vulvar cancers, and 47% of penile cancers [30]. Shortly after the discovery of HPV and its association with human cancer, scientists began developing vaccine against HPV infection. In 1991, Ian H. Frazer and Jian Zhou successfully produced HPV virus-like particles (VLP), which are comprised of only the capsid of HPV but not the genome. Nearly two decades later, HVP vaccines are commercially produced based on HPV VLP and clinically proven to prevent specific HPV infections [29]. This has changed the landscape of cancer epidemiology world-wide.

In 1994, Yuan Chang and Patrick S. Moore identified a novel herpesvirus using representational difference analysis (RDA), a technique that had only been developed the previous year [31]. By subtracting DNA libraries from Kaposi sarcoma (KS) tissue and healthy tissue, the team found two non-human DNA sequences that were specific to KS [32]. These sequences were

subsequently used to walk the entire ~185 kb genome of a new human herpesvirus designated as the Kaposi sarcoma-associated herpesvirus (KSHV) or human herpesvirus 8 (HHV8). KSHV DNA is found in all cases of AIDS as well as non-AIDS associated KS cases [33-37]. In addition to KS, KSHV also causes primary effusion lymphoma (PEL), multicentric Castleman disease (MCD), and KSHV inflammatory cytokine syndrome (KICS). KSHV possesses numerous open reading frames that exhibit remarkable homology to cellular genes. These viral gene products perform a diverse array of functions in KSHV-associated pathogenesis by disrupting essential cellular signal transduction pathways, including those involved in antiviral responses, cellular growth regulation, and apoptosis. The discovery of KSHV underscored the association of immunosuppression and the development of virus-associated cancers. In 2008, the Chang-Moore lab discovered another new human tumor virus. Huichen Feng, using digital transcriptome subtraction (DTS), identified a non-human RNA sequence in Merkel cell carcinoma (MCC) that was shown to encode the tumor antigen (T-Ag) of Merkel cell polyomavirus (MCV) [38]. MCV was found to be integrated clonally into approximately 80% of MCC tumors [38].

In the present era, the pursuit of identifying novel human tumor viruses remains a challenging area of scientific inquiry. This has been pursued using cutting-edge technologies, including deep sequencing and mass spectrometry, which have substantially improved the precision and efficacy of virus identification. Each newly identified tumor virus presents an opportunity for scientists to gain a deeper understanding of the intricate mechanisms underlying virus-induced oncogenesis, ultimately enabling the development of more effective diagnostic and therapeutic strategies. As such, ongoing research in new tumor virus discovery represents a critical aspect of both cancer biology and virology, with the potential to uncover important insights into the complex interplay between viruses and human cells.

Table 1 Identified Human Tumor Viruses

Virus	Discovery Year	Discovery Method	Associated Cancers	Taxonomy	Reference
Epstein-Barr virus (EBV)	1964	Electron microscopy	nasopharyngeal cancer, Burkitt lymphoma, gastric carcinoma, Hodgkin's lymphoma	Herpesvirus (dsDNA)	[8]
Hepatitis B virus (HBV)	1965	Serologic screening	hepatocellular carcinoma	Hepadnavirus (ssDNA /dsDNA)	[18]
Human T-cell lymphotropic virus-I (HTLV-I)	1980	Tissue culture	adult T-cell leukemia/lymphoma (ATL)	Retrovirus (ssRNA, Positive strand)	[24]
High-risk human papillomaviruses (HPV) 16 and 18	1983-84	DNA cloning	Anal, cervical, Penile, vaginal, vulvar, and oropharyngeal cancer	Papillomavirus (dsDNA)	[27], [28]
Hepatitis C virus (HCV)	1989	cDNA cloning	hepatocellular carcinoma, cholangiocarcinoma, non-Hodgkin's lymphoma	Flavivirus (ssRNA, Positive strand)	[23]
Kaposi sarcoma-associated herpesvirus (KSHV)	1994	Representational difference analysis	Kaposi sarcoma, multicentric Castleman disease, primary effusion lymphoma	Herpesvirus (dsDNA)	[32]
Merkel cell polyomavirus (MCV)	2008	Digital transcriptome subtraction	Merkel cell carcinoma	Polyomavirus (dsDNA)	[38]

1.2 Polyomaviruses

1.2.1 Discovery

Polyomaviruses are small, non-enveloped viruses that possess a 4.6-5.4 kilobases double-stranded, covalently closed, supercoiled, circular DNA genome. The name "polyomavirus" reflects the ability of members of the family to cause many ("poly") tumors ("oma"), as was demonstrated by the discovery of the first polyomavirus in 1953 by Ludwik Gross and Sarah Stewart [4]. They isolated a filterable agent that could cause salivary gland carcinoma in mice, which was later identified as murine polyomavirus (MuPyV). The second polyomavirus discovered was SV40, which was identified as a contaminant in the poliovirus vaccine and shown to induce tumors in hamsters [39]. Extensive research has been conducted on the simian vacuolating virus 40 (SV40) and its interactions with host cells, particularly with respect to the viral proteins expressed during its replication cycle and its mechanisms of replication, packaging, and transmission. Additionally, the ability of SV40 to induce tumorigenesis has been the subject of considerable investigation. This research has provided valuable insights into the molecular mechanisms underlying viral pathogenesis and cancer development. Despite over 50 years of research, there is little evidence that SV40 infects humans or causes cancer in humans [7]. One of the first human polyomaviruses, BK virus (BKV), was identified in the urine sample from a patient with renal transplantation in 1971 through electron microscopy [40]. The same year, via the same technology, JC virus (JCV) was discovered in human fetal glial cells from a brain with progressive multifocal leucoencephalopathy [41]. The two viruses remained the only polyomaviruses known to infect humans until, after nearly four decades, newly developed technologies such as high-

throughput sequencing facilitated the field of virus discovery. To date, fourteen human polyomaviruses have been identified (**Table 2**).

Table 2 Identified Human Polyomaviruses

Virus	Discovery Year	Discovery Method	Tissue Source	Related Diseases	Seroprevalence in adults	Reference
BK polyomavirus (BKV)	1971	Electron microscopy	Urine	Hemorrhagic cystitis, polyomavirus-associated nephropathy	80-100%	[40, 42-47]
JC polyomavirus (JCV)	1971	Electron microscopy	Urine, brain	Multifocal leukoencephalopathy	50-70%	[41-47]
Karolinska Institute polyomavirus (KIPyV)	2007	High-throughput sequencing	Nasopharyngeal tissue	Not identified	80-90%	[47-49]
Washington University polyomavirus (WUPyV)	2007	High-throughput sequencing	Nasopharyngeal tissue	Not identified	90-100%	[47, 49, 50]
Merkel cell polyomavirus (MCV)	2008	Digital transcriptome subtraction	Merkel cell carcinoma	Merkel cell carcinoma	70-80%	[38, 47, 49, 51]
Human polyomavirus 6 (HPyV6)	2010	Rolling circle amplification	Skin	Not identified	70-90%	[47, 49, 51-53]
Human polyomavirus 7 (HPyV7)	2010	Rolling circle amplification	Skin	Not identified	60-80%	[47, 49, 51-53]
Trichodysplasia spinulosa-associated polyomavirus (TSPyV)	2010	Rolling circle amplification	Trichodysplasia spinulosa spicules	Trichodysplasia spinulosa	60-90%	[47, 49, 51, 53, 54]

Human polyomavirus 9 (HPyV9)	2011	Generic polyomavirus PCR	Skin, blood, urine	Not identified	20-40%	[47, 49, 53, 55]
Malawi polyomavirus (MWPyV)	2012	Rolling circle amplification	Stool, wart	Not identified	80-100%	[47, 49, 53, 56]
St Louis polyomavirus (STLPyV)	2013	Rolling circle amplification	Stool	Not identified	60-70%	[49, 57, 58]
Human polyomavirus 12 (HPyV12)	2013	Generic polyomavirus PCR	Liver	Not identified	10-50%	[49, 59, 60]
New Jersey polyomavirus (NJPyV)	2014	High-throughput sequencing	Muscle	Not identified	10-70	[49, 60-62]
Lyon IARC polyomavirus (LIPyV)	2017	Generic polyomavirus PCR	Skin	Not identified	2-5%	[49, 63, 64]

1.2.2 Genome Arrangement

Polyomaviruses possess a chromatinized, double-stranded, covalently closed, circular DNA genome of approximately five kilobases, which is partitioned into three distinct regions: the early region (ER), the late region (LR), and the non-coding control region (NCCR) (**Figure 2**). The NCCR serves as the origin of replication and functions as a bidirectional enhancer and promoter for ER and LR transcription [65]. Shown here is the SV40 genome; some genes are variably present in other polyomaviruses.

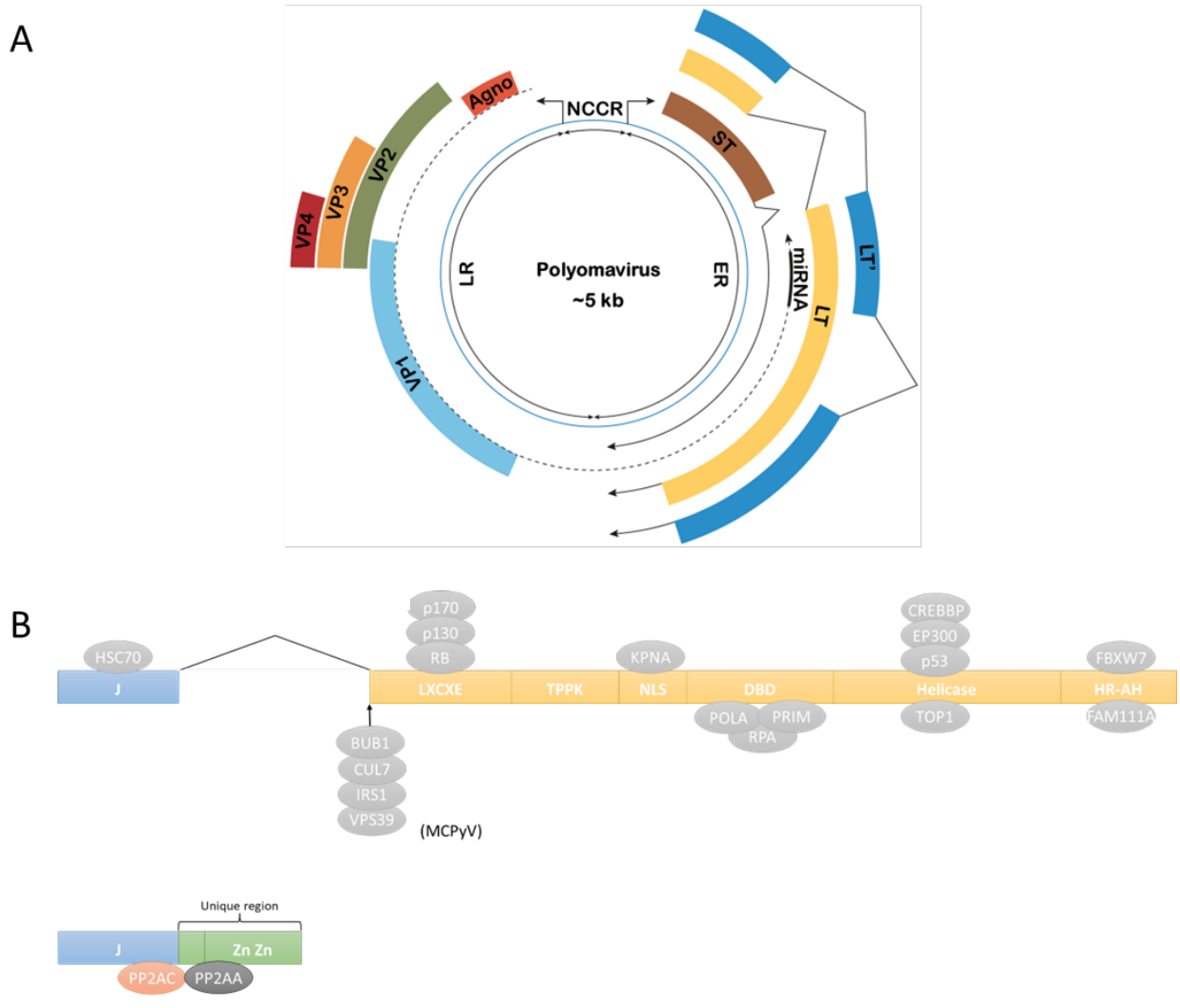


Figure 2 Schematics of SV40 Genome and Large/Small T antigen functional domains.

A. The typical polyomavirus genome is divided into three main regions: an early region (ER) encoding large T antigen (LT), small T antigen (sT), and variable numbers of alternatively spliced LT (LT') depending on the specific strain; a late region encoding the Agnoprotein (Agno) and viral capsid proteins VP1, VP2, and variably VP3, VP4; and a non-coding control region located between the early and late region, which contains the viral replication origin as well as early/late region promoters. B. LT contains a DnaJ (J) domain, a retinoblastoma-associated protein (RB)-binding LXCXE motif, a threonine-proline-proline-lysine (TPPK) motif, a nuclear-localization sequence (NLS), a DNA-binding domain (DBD), a helicase domain and a host range and adenovirus helper function (HR-AH) domain. sT also contains a J domain, followed by two zinc-finger domains. The cellular proteins which interact with LT and sT through specific domains are indicated. (Figure modified from DeCaprio and Garcea, 2013.)

The ER encodes multiple T antigens (T-Ags) via alternative splicing and expression begins upon viral entry (**Figure 2A**). Large T-antigen (LT) and small T-antigen (sT) are the two T-antigens found in all polyomaviruses. Other virus-type-specific T-Ags include the 57kT antigen [66] and middle T antigen/alternative large T antigen open reading frame (MT/ALTO) [67] for MCV, 17kT antigen for SV40 [68], and truncated T-Ag for BKV [69]. T-antigens are crucial for viral replication, viral RNA transcription, and modulation of host cell signaling pathways. In addition to T-antigens, many polyomaviruses also encode non-coding RNAs such as microRNAs (miRNAs) and circular RNAs (circRNAs) to regulate the viral life cycle [70-76].

LT includes a DnaJ domain that activates Hsc70 molecular chaperones (**Figure 2B**) [77]. Following the DnaJ domain, an LXCXE motif is conserved in human polyomaviruses, except for HPyV12 [59], enabling retinoblastoma-associated protein (RB) and RB-related proteins (p107, p130) to bind to LT (**Figure 2B**) [77]. Hsc70 binding to the DnaJ domain is necessary for LT to disrupt the interaction between the Rb protein and E2F complex, activating E2F to promote cell cycle progression [78, 79]. LT of SV40 and MCV contains a threonine-proline-proline-lysine (TPPK) motif downstream of the LXCXE motif, which is a classic cdc2/CDK1 consensus

sequence (**Figure 2B**). Phosphorylation of the first threonine is essential for the replication capacity of LT, indicating that cdc2/CDK1 phosphorylates the threonine to activate LT [80]. LT contains several conserved domains among polyomaviruses following the TPPK motif, including a nuclear localization sequence (NLS), an origin-binding domain (OBD) that recognizes pentanucleotide sequences (PS) in the replication origin, and a helicase domain (**Figure 2B**). Through these domains, LT assembles double head-to-head hexamers on the replication origin to unwind DNA and recruit cellular DNA polymerase-primase to initiate viral genome replication [81-86].

MCV sT, a product that reads through the first splicing site of LT, contains the N-terminal DnaJ domain and a cellular protein phosphatase 2A (PP2A) binding domain that is not present in LT (**Figure 2B**). PP2A is a serine/threonine phosphatase composed of A, B, and C subunits. The A subunit provides structural support, the B subunit determines substrate specificity, and the C subunit catalyzes dephosphorylation. By interacting with the A and C subunits of PP2A, sT alters its function by removing the B subunit [87-89].

In addition to LT and sT, some polyomaviruses express virus type-specific middle T-Ags (MT). For instance, MT has been extensively studied in MuPyV and is necessary for both MuPyV replication and cell transformation through its various domains including an N-Terminal PP2A binding sequence, various SH2 and SH3 interaction motifs, and C-terminal transmembrane domain [90, 91]. In contrast to the MT of MuPyV, MCV MT does not contain an N-Terminal PP2A binding sequence. This function is partitioned to its sT protein. Notably, MCV MT is not expressed in virus infected MCC and therefore is not likely to contribute to MCV mediated transformation of MCC. To date, the contributions of MCV MT to virus maintenance and reactivation has not been defined.

The LR expresses various viral capsid proteins (VPs) that assemble the viral capsid by alternative splicing and alternative translational initiation sites [92]. Polyomavirus capsids are composed of 72 self-assembled pentamers of the major capsid protein VP1, associated with a single molecule of VP2 or VP3 per VP1 pentamer [93-95]. In addition to VP proteins, certain polyomaviruses, including SV40, JCV, and BKV, produce a small peptide called agnoprotein from the LR. Unlike other viral homologous proteins, agnoproteins have highly diverse sequences, which accounts for their varied functions across different polyomaviruses. Nonetheless, agnoprotein plays a significant role in several stages of the polyomavirus lifecycle, such as viral gene expression, genome replication, maturation, and release [96, 97].

1.2.3 Life Cycle

The life cycle of polyomaviruses consists of several stages, including viral attachment, entry, early transcription, replication, late transcription, virion assembly, and release (**Figure 3**). Gangliosides are the primary receptors for most polyomaviruses [98-103]. Polyomaviruses, except for JCV, utilize caveolin-mediated endocytosis for host cell entry [104, 105]. Once inside the host cell, the virus is transported into the smooth endoplasmic reticulum by microtubules [106-112], and the capsid is isomerized by redox proteins such as PDI, ERdj5, and ERp57 (SV40, JCV), ERp29 (MuPyV, JCV), or ERp72 (JCV) [113-116]. After decapsidation, the virus is exported into the cytosol by endoplasmic reticulum-associated degradation (ERAD), and the viral genome is translocated into the nucleus using the host cell nuclear transport machinery and the nuclear localization sequence (NLS) in residual VP proteins [115, 117, 118]. In the nucleus, the early region of the genome is expressed driven by enhancer and promoter activities of the NCCR. T-Ags disrupt Rb interaction with E2Fs through it DnaJ domain and LXCXE motif, driving the host

cell into the S phase to replicate the viral genome using the cellular DNA replication machinery. The LT protein also acts as a helicase to initiate viral DNA replication at the replication origin. Transcription of the late region of the genome occurs after replication and leads to the production of viral capsid proteins, which encapsulate the chromatinized viral genome [119, 120]. While host cell lysis is generally considered the primary mode of progeny virion release, some evidence suggests that endocytic trafficking to the plasma membrane may also be involved as an alternative mechanism [121-125].

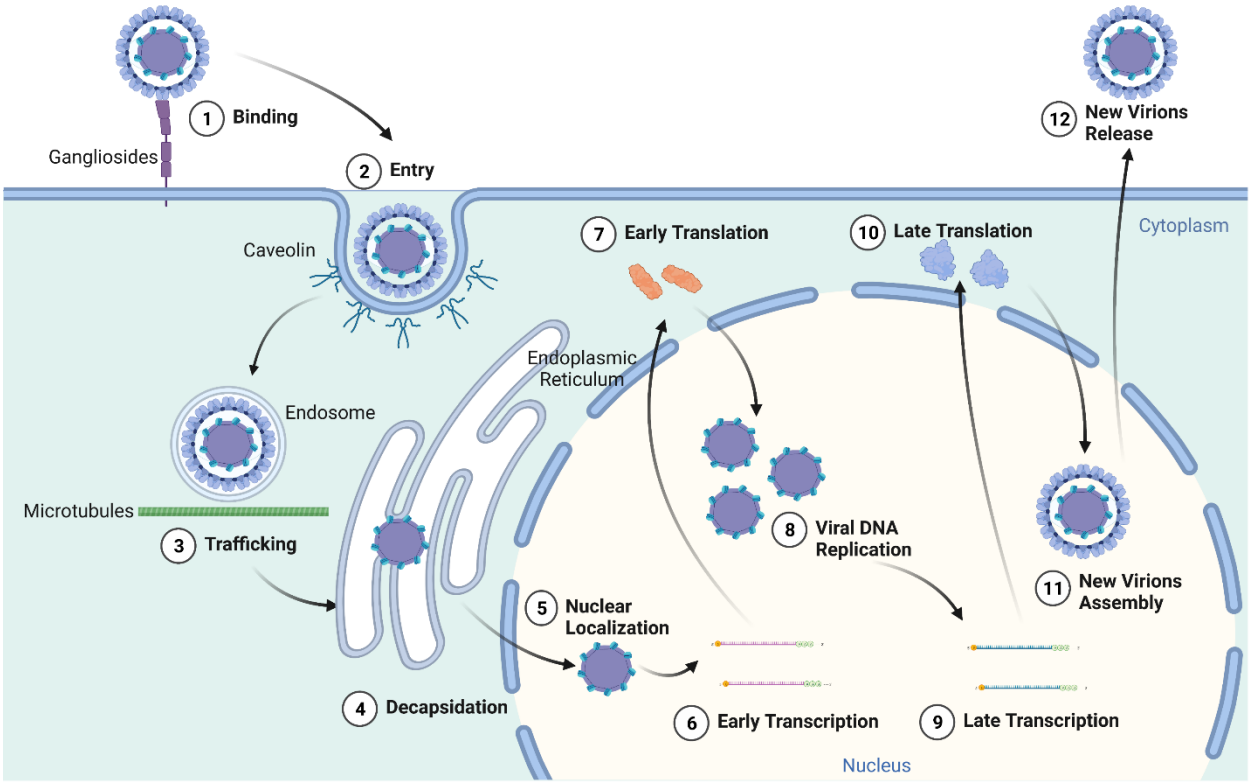


Figure 3 Schematic of Polyomavirus Life Cycle.

Most polyomaviruses interact with host cell receptors, such as gangliosides, and gain entry into host cells through caveolin-mediated endocytosis. Following entry, the viral particles are transported to the endoplasmic reticulum (ER) via endosomes and microtubules. Within the ER, the viral capsids disassemble, allowing the viral DNA genomes to translocate into the nucleus and initiate early transcription. Early

transcripts encode for viral T antigens, which activate viral genome replication and stimulate late transcription to produce capsid proteins. The replicated viral genome is then packaged by capsid proteins and subsequently released from the host cells. (Figure created with BioRender.com)

1.3 Merkel Cell Polyomavirus

1.3.1 Discovery

Merkel cell carcinoma (MCC) is a rare and aggressive cutaneous malignancy associated with a high mortality rate, with a 46% 5-year disease-associated mortality rate [126]. The incidence of MCC has rapidly increased from 0.15 cases per 100,000 people in 1986 to 0.7 cases per 100,000 in 2013 [127, 128]. MCC typically presents clinically as a fast-growing, painless, and firm nodule on sun-exposed skin [129]. The disease predominantly affects older adults with light skin and immunocompromised individuals [130-132]. For instance, individuals with AIDS or solid organ transplant recipients, particularly for kidney, liver, and heart transplantation, are at a higher risk of developing MCC [133]. Autoimmune diseases such as Crohn's disease, inflammatory bowel disease, and ankylosing spondylitis, but not rheumatoid arthritis, have been associated with an increased risk of developing MCC [133, 134].

The association between MCC and immune suppression motivated researchers at the University of Pittsburgh, led by Yuan Chang and Patrick Moore, to develop a bioinformatics-based technology known as Digital Transcriptome Subtraction (DTS), which differentiates non-human RNA sequences by computational subtraction of human sequences from cDNA deep sequencing library of MCC tumor specimens [38]. Through this high-throughput method, they isolated one fusion transcript with high homology to the African green monkey lymphotropic polyomavirus

(LPyV) and BKV T-Ag sequences and another transcript with no homology to previously identified polyomaviruses, which was later shown to be contained in the viral genome [38]. Since the new polyomavirus was identified from MCC and its genome was integrated into ~80% of MCC, it was named Merkel cell polyomavirus (MCV). MCV infection is near-ubiquitous, asymptomatic, and lifelong. Serological studies indicate that the prevalence of MCV infection rises from 37% to 42% among 1- to 6-year-olds, reaching 92% in adults within the general population, which points to a potential infection during early childhood [135]. To date, approximately 50 MCV variants have been identified from MCCs or skin swabs, displaying a relatively low genetic diversity [38, 52, 66, 136-145]. To study MCV life cycle, our lab designed a consensus genome based on isolates MCV350 (EU375803), MCV339 (EU375804), MCV344 (JF812999), MCV349 (JF813000), MCV352 (JF813001), MCVMKL-1 (FJ173815), and MCV85 (JF813002) [138].

1.3.2 Genome Arrangements

As with other polyomaviruses, the MCV genome is divided into the early region (ER) and the late region (LR) by the non-coding control region (NCCR) (**Figure 4**). The NCCR contains regulatory elements for both early and late gene expression, as well as the viral origin of replication [38, 146, 147]. The ER produces several alternatively spliced mRNAs that encode four T antigens: large T antigen (LT), small T antigen (sT), 57kT antigen (57kT), and middle T antigen or alternative LT open reading frame (MT/ALTO) [38, 66, 67]. In addition, the ER also encodes a miRNA, miR-M1, which targets all MCV early transcripts [72, 76]. The LR encodes two viral structural proteins, VP1 and VP2, that are responsible for forming the capsids during lytic replication [38]. Unlike other polyomaviruses, MCV does not express VP3, VP4, or agnoprotein [38, 102, 148, 149]. Furthermore, MCV has been found to encode multiple circular RNAs

(circRNAs), some of which are capable of regulating protein expression or encoding protein products [150, 151].

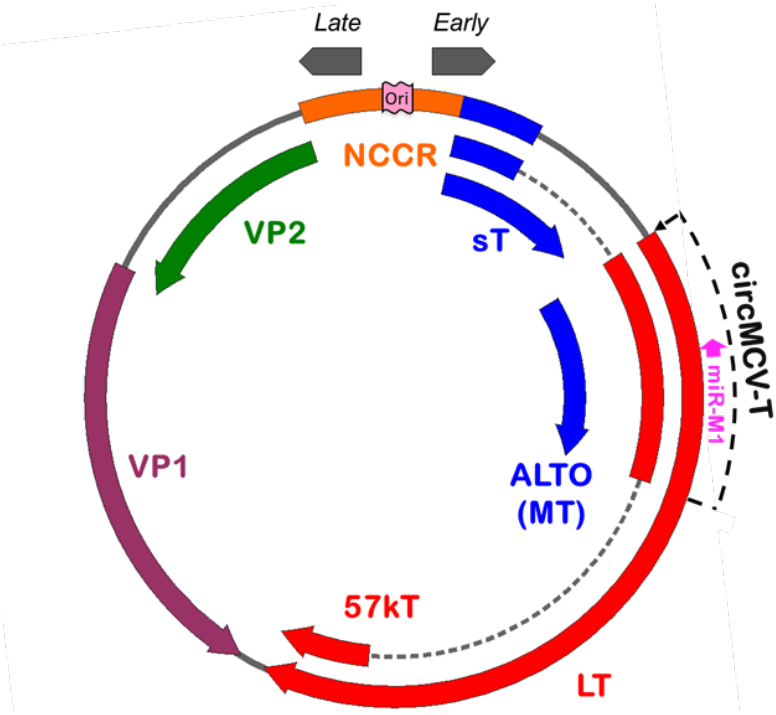


Figure 4 A Schematic of MCV Genome.

MCV genome can be divided into three distinct regions: early (ER), late (LR), and non-coding control (NCCR). The NCCR contains the MCV replication origin (Ori) as well as promoters for both early and late regions. The ER region encodes several proteins, including large T antigen (LT), small T antigen (sT), 57kT antigen (57kT), and middle T antigen/alternative LT open reading frame (MT/ALTO). Additionally, the ER region gives rise to a microRNA (miRNA), miR-M1, and four circular RNAs (circRNAs), with circMCV-T exhibiting the highest expression level (see Chapter 3). The LR region is responsible for encoding the capsid proteins VP1 and VP2.

1.3.3 Association With MCC

As mentioned, while MCV was discovered in MCCs, the causal relationship between MCV and MCC had yet to be established. Several pieces of evidence, however, suggested that MCV is

indeed involved in the development of MCC. First, the MCV genome was found to be clonally integrated into the genome of all virus infected MCC, indicating that viral integration occurs prior to tumorigenesis [38, 152-154]. Second, MCV sT and truncated LT proteins expressed in all MCC tumors, have been shown to be necessary for MCC cell survival [155-158]. Third, MCV sT has been found to have transformation capacity in Rat-1 cells, NIH 3T3 cells, and transgenic mouse models [159-162]. Based on these findings and the fact that MCC primarily affects older and immunosuppressed individuals, it is believed that MCV can reactivate in the absence of immune surveillance and integrate its genome into the host genome via non-homologous recombination (**Figure 5**) [148, 163]. However, if MCV genome replication still occurs after integration, the cell will be eliminated by triggering a DNA damage response. Therefore, a mutation is required to abolish the replicative capacity of the integrated MCV genome for the proto-cancer cell to survive [66]. Missense or nonsense mutations that generate C-terminal truncated LT, which lacks the helicase domain, typically eliminate the replicative capacity of LT. Interestingly, tumor-derived LTs still retain a functional DnaJ domain and Rb-binding motif, suggesting that Hsc70 and Rb binding are important for tumor outgrowth [66].

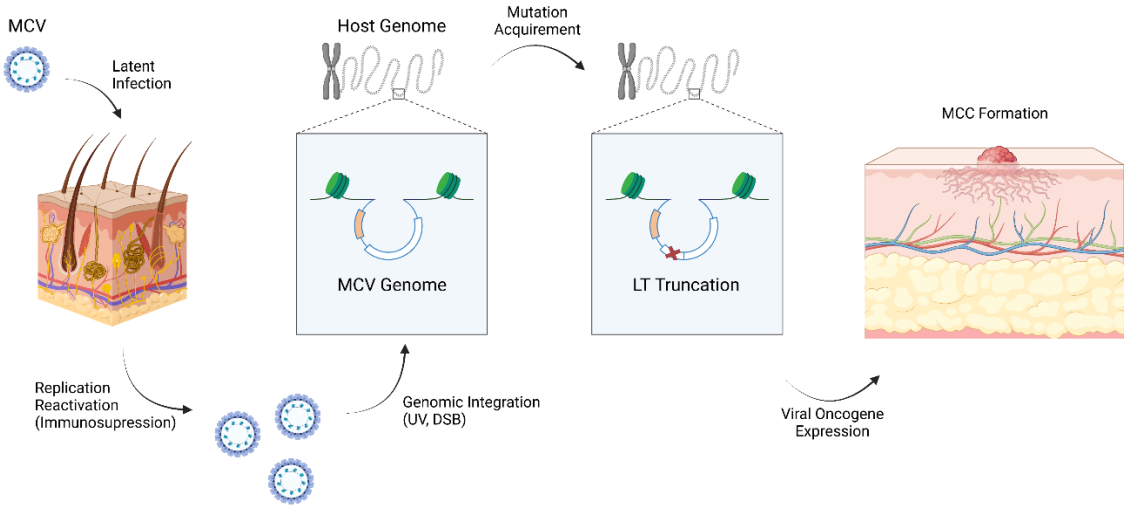


Figure 5 Hypothetical Model for MCV-induced Tumorigenesis.

MCV infection usually remains asymptomatic. However, when individuals experience immunosuppression, viral replication may reactivate. During this process, MCV integrates into the host genome, potentially through ultraviolet irradiation or other factors that induce viral DNA double-strand breaks (DSBs). After integration, selective pressure eliminates replication-competent integrated viruses containing an intact LT. Over time, the expression of viral oncogenes drives tumorigenesis, ultimately leading to the development of Merkel cell carcinoma (MCC). (Figure modified from Chang and Moore, 2012 and created with BioRender.com.)

1.3.4 Life Cycle

Similar to other polyomaviruses, the life cycle of MCV can be divided into several stages, including viral entry, early transcription, replication, late transcription, virion assembly, and release. Due to the lack of a natural infectious model system for MCV, Merkel cells were initially thought to be the natural host cells as MCC was classified as Merkel cells based on cell markers such as cytokeratin 20, chromogranin A, synaptophysin, and neuropeptides [164-170]. However, some other cell markers, such as vasoactive intestinal peptide and metenkephalin, are only identified in Merkel cells, while CD117 and CD171 are only shown in MCC cells [171-174]. In addition, as post-mitotic cells, Merkel cells are insensitive to oncogenic stimuli, which suggests that Merkel cells may not be the origin of MCC or the host of MCV. Thus, researchers have started to test the capacity of infection and replication of MCV in different cell types. Besides cancer-derived cell lines, human dermal fibroblasts, especially the dermal fibroblast underlying the basal layer of the epidermis and surrounding hair follicles, were shown to support both MCV infection and replication [138, 149, 175, 176]. Some studies have suggested that MCV can be detected in eyebrow hair bulb samples obtained from human volunteers [177-179].

Researchers have also investigated the cell receptors that MCV utilizes to attach to the host cell. Unlike most polyomaviruses, MCV attachment is reliant on sulfated glycosaminoglycans, such as heparan sulfate and chondroitin sulfate. Upon caveolar/lipid raft-mediated endocytosis, MCV is transported into the endoplasmic reticulum within acidified endosomes, utilizing microtubular transport and disassembling the capsid in a functional redox environment. However, only a small proportion of viral particles are transported into the endoplasmic reticulum, while the majority of viruses are confined to endosomes, indicating the speed-limiting step of MCV infection and a possible explanation for the virus's low infectivity [111]. Following decapsidation, MCV genomes are transported into the cytosol before being transported to the nucleus. The process of MCV genome transportation from the endoplasmic reticulum to the nucleus has not been investigated.

Upon entering the nucleus, the endoplasmic reticulum of MCV initiates the expression of T-Ags which target cell regulators and induce the process of DNA replication in the host cell [138, 162, 180]. Similar to SV40, MCV's Large T-antigen (LT) forms a head-to-head double hexamer at the replication origin and initiates replication by unwinding the MCV genome [146, 149, 181]. Following genome replication, VP proteins are expressed and aid in the packaging of the MCV genome. However, the mechanisms underlying the early-late switch and egress of MCV particles from host cells remain elusive.

1.3.5 Minicircle System for MCV Genome Production

In order to generate the genome of double-stranded circular DNA viruses, a commonly employed technique is to insert the viral genome into a bacterial plasmid backbone, followed by amplification in bacteria. The viral genome can then be released from the plasmid backbone by

restriction enzymes and circularized by DNA ligases [138, 149, 182]. Since the infectivity and replication of MCV in mammalian cells are low, this has been the primary method used to produce MCV genomes on a large scale. However, this approach is associated with certain limitations, including relatively low yields of the MCV circular genome, which is further exacerbated by concatenation during the ligation steps [183].

To resolve this issue, I adapted the minicircle technology to produce and circularize the MCV genome [184]. The minicircle technology was originally developed to produce circular DNA fragments free from nonviral, prokaryotic replication origins and antibiotic resistance genes for use in DNA vaccines and gene therapy applications. Minicircle technology utilizes integrase, recombinase, or resolvase enzymes to release and circularize the DNA fragment of interest from the bacterial backbone [185]. The first minicircle system was established in 1997 by A-M Darquet and colleagues, who used λ-integrase to excise a circular DNA fragment from a plasmid [186]. Other recombinases, such as Cre-recombinase, FLP-recombinase, and ParA-resolvase, have also been used to generate minicircles [187-191]. In 2003, Zhi-Ying Chen and colleagues developed a minicircle system using ΦC31-integrase, but this protocol was laborious and time-consuming due to the difficulty of removing the circularized bacterial backbone. To address this problem, they further refined the system by incorporating 32 I-SceI endonuclease recognition sites into the bacterial backbone. By expressing ΦC31-integrase and I-SceI in a genetically modified minicircle-producing bacterial strain, ZYCY10P3S2T, via arabinose induction, the DNA of interest is circularized while the bacterial backbone is digested by I-SceI [184, 192]. This improvement allows for easy removal of the bacterial backbone through a plasmid purification process.

1.3.6 MCV Viral Proteins

As previously described, Merkel cell polyomavirus (MCV) expresses four T-antigens (T-Ags) from the ER and two viral protein (VP) proteins from the LR during natural infection. Among the T-Ags, MCV large T antigen (LT) and small T antigen (sT) play crucial roles in viral replication and host cell signaling regulation. At the same time, the functions of 57kT and ALTO remain unknown. For the VP proteins, VP1 is the primary capsid protein, while VP2 also plays a crucial role in MCV infection in specific cell types. This chapter aims to provide a comprehensive review of the functions of all MCV proteins.

1.3.6.1 MCV Large T Antigen

The MCV LT protein is a multifunctional protein that plays a crucial role in the virus's life cycle (**Figure 6**) [193]. Similar to other polyomaviruses, the N-terminal region of the MCV LT protein contains an intact DnaJ domain and an LXCXE domain [66, 146]. In addition to these domains, MCV contains two unique regions, MUR1 and MUR2, which are not present in other polyomaviruses. MUR1 includes a β-transducin repeat-containing protein (β-TrCP) binding site and a human Vam6p (hVam6p) binding site, while MUR2 contains an S-phase kinase-associated protein 2 (Skp2) and an F-box and WD repeat domain-containing 7 (Fbw7) binding site [147, 194]. The C-terminal region of MCV LT encodes a nuclear localization sequence (NLS), an origin binding domain (OBD), a zinc finger domain, and a helicase/ATPase domain [77, 146, 149, 181].

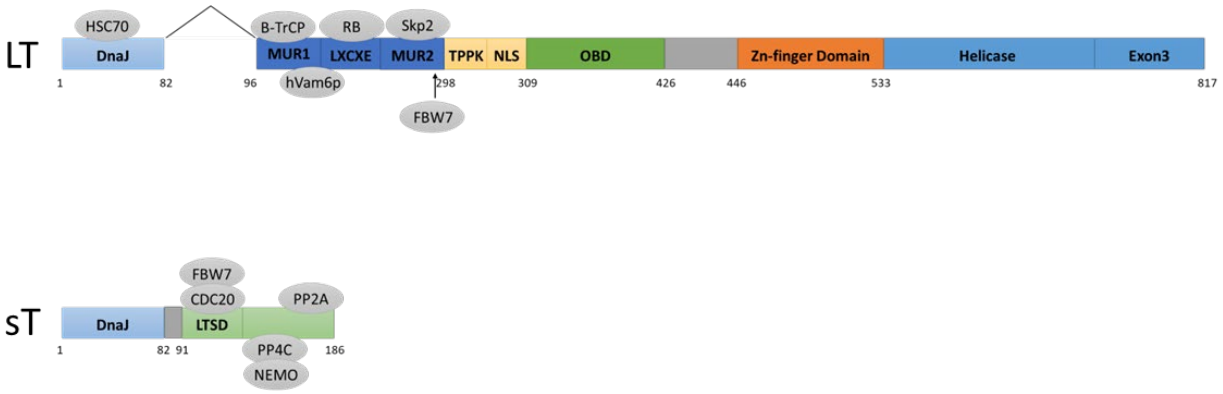


Figure 6 MCV Large/Small T antigen functional domains.

MCV large T antigen (LT) and small T antigen (sT) are the major two tumor antigens (T-Ags) which are critical for MCV life cycle and MCV-induced tumorigenesis. MCV LT contains a DnaJ domain, two MCV unique regions (MUR1/2), a LXCXE motif, a TPPK motif, a nuclear-localization sequence (NLS), a DNA-binding domain (DBD), a helicase domain and C-terminus exon3 domain. MCV sT also contains a DnaJ domain, followed by a unique LT stabilization domain (LTSD), PP4C/NEMO binding site, and a PP2A binding site. The cellular proteins which interact with MCV LT and sT through specific domains are indicated.

The DnaJ domain (amino acids 1-82) within the N-terminal portion of MCV LT contains a canonical CR1 motif (LXXLL) and a Hsc70 binding motif (HPDKGG). The interaction between MCV LT and Hsc70 through the DnaJ domain facilitates efficient viral DNA replication [146]. Additionally, as the DnaJ domain is also present in sT and 57kT, the interaction between MCV T-Ags and Hsc70 has been shown to stabilize DICER1 mRNA, which suggests MCV T-Ags may regulate host cell mRNA and miRNA through Hsc70 binding [195, 196].

Following the DnaJ domain, the MUR (amino acids 96-298) in MCV LT plays a crucial role in regulating MCV persistence [147]. β-TrCP, Skp2, and Fbw7 are members of Skp-Cullin-F box (SCF) E3 ubiquitin ligases, which mediate the ubiquitylation of target proteins to induce proteasomal degradation via the cellular ubiquitin-proteasome system [197]. Through this mechanism, LT can be degraded though SCF E3 ligases resulting in decreased effective protein

concentrations of this critical protein for viral replication. The end result may be the promotion of a latency-like status to maintain long-term persistence [147, 198, 199]. MUR also contains a hVam6p binding site, which is located close to the LXCXK motif. hVam6p is a cytoplasmic protein required for lysosomal tethering and has been shown to promote lysosome clustering and fusion in cells [200]. By interacting with hVam6p via hVam6p clathrin heavy chain homology domain, MCV LT sequesters hVam6p in the cell nucleus, depending on its NLS, and disrupts lysosomal clustering [194]. Studies have shown that this interaction inhibits MCV replication, indicating that MCV uses multiple host cell components to maintain low-level replication for its "latency" [138]. Additionally, a recent study found that MUR is critical in MCV LT-induced cellular senescence, which is also important for MCV genome persistence [201].

Adjacent to the hVam6p binding site, MCV LT contains a highly conserved LXCXE motif, similar to other polyomaviruses, which interacts with Rb and Rb family proteins [66, 77, 202]. Rb and Rb family proteins, p107 and p130, are repressors of the E2F family of transcription factors. By directly engaging with the E2F transactivation domain and associating with E2F, Rb proteins impede E2F-regulated gene transcription, which is crucial for advancing the cell cycle from the G1 to S phase [203].

SV40 LT is capable of binding to the Rb and Rb family proteins, thereby releasing the E2Fs *in vitro* [204-206]. However, interestingly, during an actual SV40 infection, only the p107-E2F and p130-E2F complexes are disrupted [207]. This implies that the cell cycle arrest mediated by p107 and p130 may pose a hindrance to the productive infection process. Different from SV40, to activate and hijack the DNA replication machinery of the host cell, MCV LT sequesters Rb, but not p107 and p130, to free E2Fs, upregulate E2F target genes, and subsequently transit cell into the S phase as described [158, 208-212]. As a result, survivin, an E2F target gene, is upregulated

by MCV LT. Inhibiting survivin expression leads to inhibition of MCC cell proliferation and extended lifespan of mice bearing MCC xenografts [208, 213].

Downstream of MUR is the NLS, OBD, zinc finger domain, and helicase/ATPase domain. Unlike other polyomaviruses, the NLS amino acid sequence of MCV LT is RKRK instead of KKKR(or K)KVEDP [214]. Upon translocation to the nucleus through its NLS, the LT's OBD (amino acids 309-426) recognizes and associates with the specific pentanucleotide repeat sequence, GAGGC, found in the replication origin of the MCV [146]. Interestingly, a commonality across polyomaviruses lies in their utilization of similar pentanucleotide sequences for LT binding [215-218]. However, the specificity of T-antigen binding is primarily determined not by these shared sequences, but by their flanking sequences and the spatial distance between the pentanucleotide sequences [219]. Importantly, while the OBD is the main DNA binding domain of MCV LT, adequate DNA binding and initiation of replication necessitate LT's phosphorylation at amino acids T297 and T299, situated between the NLS and OBD [80]. Following the recognition and binding of the pentanucleotide sequences, the LT mediated by its zinc finger domain (amino acids 446-533), assembles into an asymmetric double hexamer configuration in a head-to-head manner [181, 220]. While there are similarities with the SV40 LT, a key distinction lies in the fact that the SV40 LT forms symmetric double hexamers [221]. This difference underscores the critical principle that the precise arrangement of pentanucleotide sequences must align with the structure of each polyomavirus LT for the initiation of replication, contributing to the distinct replication mechanisms unique to each polyomavirus species [66, 218]. After LT oligomerization, the helicase/ATPase domain (amino acids 533-817) then hydrolyzes ATP and unwinds DNA, which further recruits cellular DNA replication machinery to initiate MCV genome replication [77, 146, 181].

As described, the C-terminal domains (amino acids 309-817) of MCV LT are important for replicating the viral genome. However, like other polyomaviruses, these domains can also induce a host DNA damage response (DDR) [175, 222-230]. The DDR activates two downstream pathways: the ataxia telangiectasia mutated (ATM) kinase pathway and the ataxia telangiectasia and Rad3-related (ATR) kinase pathway. While ATM is primarily triggered by double-stranded DNA breaks (DSBs), ATR responds to a wide range of DNA damage, including single-stranded DNA lesions [231]. In the course of a natural MCV infection and during whole genome transfection in host cells, both the ATR and ATM pathways are simultaneously activated. This activation subsequently leads to the co-localization of host DNA Damage Response (DDR) factors with MCV replication foci, a crucial process underpinning successful MCV DNA replication [224]. However, it is primarily the expression of MCV LT that induces the ATR pathway, implying that the activation of the ATM pathway could be attributed to other viral activities [175]. The DDR induced by the MCV LT triggers the phosphorylation of p53 and subsequent activation of its downstream target genes. This cascade leads to host cell cycle arrest in S and G2/M phase [175, 230, 232, 233]. Additionally, insights from a study on BK Virus (BKV) indicate that the virus stimulates the DDR by manipulating the mismatch repair pathway to extend the host cell's S phase and inhibit regular mitotic entry post the G2 phase [228, 229]. This approach could be representative of a strategy employed by polyomaviruses to commandeer the host cell's DDR pathway, thereby facilitating its own replication.

In addition to the previously discussed motifs, MCV LT has been shown to interact with several other cellular components, such as bromodomain protein 4 (Brd4) and ubiquitin-specific processing protease 7 (Usp7). Brd4, belonging to the Bromodomain and Extraterminal (BET) protein family, is well known in cancer research for its pivotal role in cell cycle progression,

transcriptional regulation, and chromatin remodeling [234]. The N-terminal portion of MCV LT interacts Brd4 and transports it to MCV genome replication foci, which is necessary for efficient MCV replication [235]. When both MCV LT and sT are present, Brd4 also enhances NCCR transcriptional activity [236]. Usp7, is an enzyme that belongs to the family of deubiquitinating enzymes (DUBs). Usp7 is widely recognized for its role in various cellular processes, including DNA damage repair, apoptosis, immune response, and the cell cycle [237]. MCV LT binds to Usp7 and relocates it to MCV genome replication foci, different from Brd4, where it represses MCV replication [238]. While numerous other interactors with MCV LT have been identified, the functions and implications of these additional LT interactions remain to be fully elucidated [239].

1.3.6.2 MCV Small T Antigen

MCV sT shares exon 1 with LT, containing the DnaJ domain, and thus, can regulate DICER1 expression via Hsc70 (**Figure 6B**) [195]. However, unlike LT, the DnaJ domain of sT does not affect its capacity in MCV replication and host cell transformation [146, 159, 240].

Downstream of the DnaJ domain, sT contains an LT stabilization domain (LTSD) (amino acids 91-95), identified by alanine scanning mutagenesis and found to stabilize LT by inhibiting Fbw7 (**Figure 6**) [159]. Fbw7 targets not only MCV LT but also other proto-oncogenes such as cyclin E, Myc, Jun, and Notch [241]. Thus, the Fbw7 inhibition function of sT may be involved in MCV-induced tumorigenesis.

LTSD also interacts with cell division cycle protein 20 (Cdc20), a fundamental regulator of cell division that is instrumental in the transition from metaphase to anaphase. This interaction stimulates mitogenesis and activates the cyclin-dependent kinase 1 (CDK1)/cyclin B1 (CYCB1) complex. CDK1, a key facilitator of the G2 to mitosis transition in the cell cycle, subsequently promotes the hyperphosphorylation of the eukaryotic initiation factor 4E (eIF4E)-binding protein

(4E-BP1). As a result, the cap-binding translation initiation factor eIF4E is liberated, enhancing cap-dependent translation, a process necessary for the transformation induced by sT [240, 242, 243].

MCV sT possesses the capacity to influence transcription via its LTSD. The EP400 complex, part of the nucleosome acetyltransferase of H4 (NuA4) histone acetyltransferase family, is known to be targeted by adenovirus E1A oncoprotein for cell transformation purposes [162]. MCV sT, in part through its LTSD domain, recruits MYCL, a member of the MYC transcription factor family, to the EP400 complex, subsequently activating MDM2, an E3 ubiquitin-protein ligase and a p53 suppressor, and CK1α, a kinase that activates MDM4. [244]. The activation of CK1α by MCV sT relieves the auto-inhibitory function of MDM4, another p53 suppressor, enhancing the MDM4-p53 interaction [245, 246]. MDM4 then assists in the recruitment of MDM2 to p53, leading to ubiquitination and degradation of p53 [247-249]. This mechanism underlines the p53 suppression by MCV through transcriptional regulation. Recent findings also indicate that MCV sT serves as a viral transcription activator through its LTSD. The LTSD interacts with Cux1, c-Jun, Brd9, and Cbp - all known regulators of MCV transcription - facilitating activation of MCV early region (ER) and late region (LR) transcription [250].

The LTSD of MCV sT is followed by two binding sites for PP4C/NEMO and PP2A (**Figure 6**). NEMO is an essential component of the IKK complex, which activates the transcription of proinflammatory and antiapoptotic genes in response to NF-κB signaling. NEMO, along with IKKα and IKKβ, phosphorylates IκB, which permits its degradation by the β-TrCP ubiquitin ligase complex and activation of NF-κB [251]. PP4 and PP2A are phosphoprotein phosphatases that dephosphorylate target proteins [252]. MCV sT binds to the NEMO, PP4C, the catalytic subunit of PP4, and/or PP2A Aβ, the second PP2A structural subunit, inhibiting IKKα

and IKKβ phosphorylation, thus preventing IκB phosphorylation and nuclear translocation of NF-κB. Binding to PP4C also enhances cell mobility and migration, which may contribute to the highly metastatic nature of MCC. Additionally, MCV sT binds to an iron-sulfur cluster mapped to the PP2A binding sites, enhancing LT-mediated replication independent of LTSD [253]. Interestingly, the integrity of the PP2A binding sites is not necessary for MCV sT-induced transformation [161, 240].

MCV sT also inhibits N-myc Downstream-Regulated Gene 1 (NDRG1), a tumor and metastasis suppressor, promoting MCC cell proliferation, though the mechanism remains unclear [254].

1.3.6.3 MCV 57kT and ALTO/Middle T

Compared to LT, 57kT lacks an intact OBD and helicase/ATPase domain due to splicing [66]. ALTO, also known as MCV middle T (MT), is a transmembrane protein encoded by the same frame as sT but without the sT sequence [67]. However, the functions of MCV 57kT and MT remain unknown.

1.3.6.4 MCV VP1 and VP2

VP1 and VP2 are MCV capsid proteins encoded by the LR. They play critical roles in viral genome packaging and viral attachment/entry. The attachment of MCV to the cell surface involves the interaction of VP1 capsomeres with sulfated glycosaminoglycans [255]. Additionally, for the virus to enter the host cell, VP1 must interact with sialic acid receptors [103]. Although VP1 can self-assemble in mammalian cells, VP2 is required for optimal infectivity in some cell lines [102]. Myristoylation, the addition of a myristic acid molecule to the N-terminus of VP2, is crucial for efficient MCV infection in cell lines that support VP2-dependent entry [102].

1.3.7 MCV Non-Coding RNAs

Non-coding RNAs (ncRNAs) have emerged as an exciting and expanding area of research in recent years, owing to their diverse and critical regulatory roles in the cell. Among the identified ncRNA classes are microRNAs (miRNAs), long non-coding RNAs (lncRNAs), and circular RNAs (circRNAs). Additionally, some human tumor viruses are known to encode ncRNAs, which regulate viral and host gene expression, modulate the host immune response, facilitate viral replication, or maintain viral persistence. This chapter provides a detailed discussion of the role of MCV-encoded ncRNAs, specifically miRNAs and circRNAs.

1.3.7.1 MicroRNAs

MicroRNAs (miRNAs) are a class of small non-coding RNAs that play a fundamental role in the post-transcriptional regulation of gene expression in eukaryotes (**Figure 7**). miRNAs are transcribed initially as long primary transcripts (pri-miRNAs) by RNA polymerase II and then undergo processing in the nucleus by the Drosha-DiGeorge syndrome critical region 8 (DGCR8) complex to generate shorter hairpin structures called precursor miRNAs (pre-miRNAs) [256-258]. Pre-miRNAs are then transported to the cytoplasm and cleaved by the RNase III enzyme Dicer, associated with the Tar RNA binding protein (TRBP), to form miRNA duplex intermediates [259-261]. One strand of the miRNA duplex is incorporated into the RNA-induced silencing complex (RISC), while the other strand is degraded [262, 263]. The mature miRNA within RISC recognizes and binds to specific target RNAs through the “seed sequence” within positions 2-8 nucleotides from the 5’ end of the miRNA. Upon binding, mature miRNAs can interact with target mRNAs by either inducing mRNA degradation or translational inhibition. Mature miRNAs can also attach to the coding region or the 5’ UTR of target mRNAs, leading to either the repression or activation

of translation, depending on the specific context of the interaction [264-267]. MiRNAs' biogenesis and regulatory roles are complex and highly regulated, with their expression being tissue-specific and developmentally regulated. Thus, understanding the intricate mechanisms underlying miRNA function is crucial for deciphering gene expression regulation in eukaryotes.

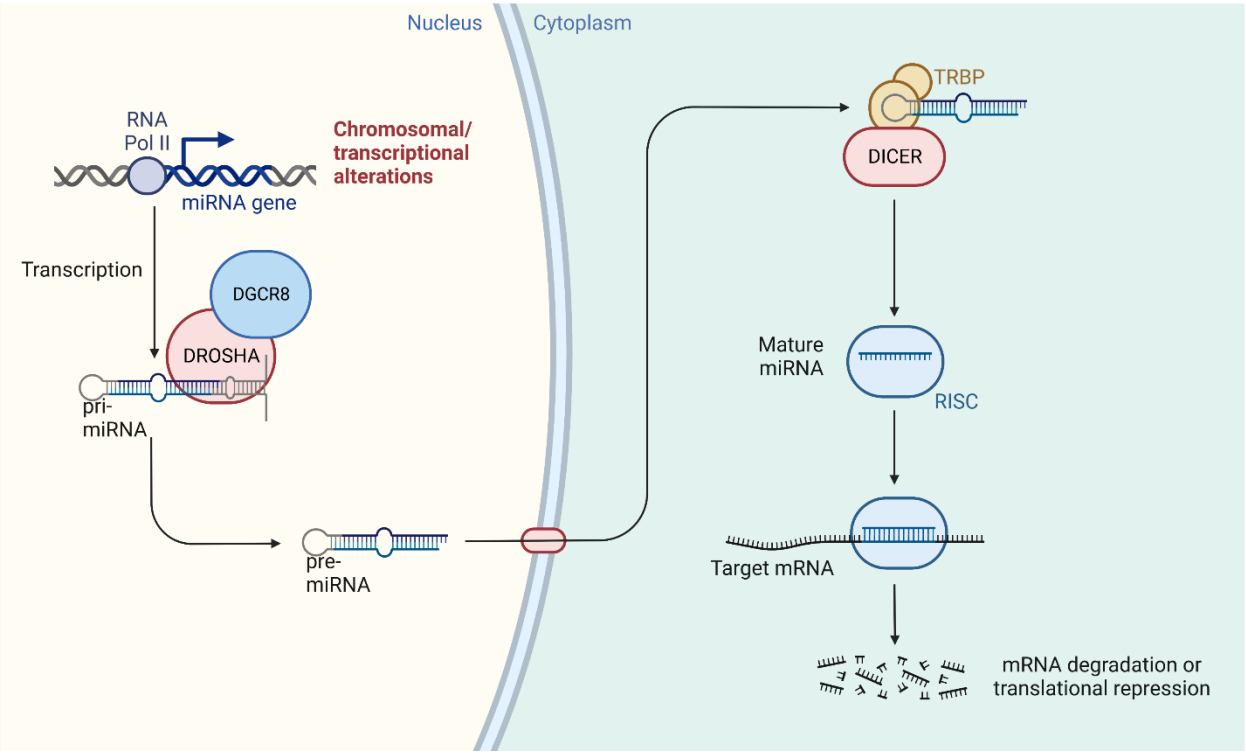


Figure 7 A Schematic of MicroRNA Biogenesis.

Pri-microRNA (pri-miRNA) is transcribed by RNA polymerase (Pol) II and subsequently processed by Drosha and DiGeorge Syndrome Critical Region 8 (DGCR8) to generate pre-miRNA. After translocating to the cytosol, pre-miRNA is cleaved by Dicer and Tar RNA-binding protein (TRBP). One strand of the miRNA duplex is then incorporated into the RNA-induced silencing complex (RISC), which carries out regulatory functions on its target mRNA. (Figure created with BioRender.com.)

In addition to eukaryotes, viruses also exploit miRNAs to regulate gene expression. Human tumor viruses Kaposi's sarcoma-associated herpesvirus (KSHV) and Epstein-Barr virus (EBV) encode multiple miRNAs that are required to repress host cell apoptotic response by inhibiting caspase activation, directly or indirectly [268-275]. KSHV and EBV miRNAs can also inhibit the

host immune response by targeting type I interferon (IFN) response and Toll-like receptor (TLR) pathway, which can eliminate host cells [14, 276-280]. Additionally, KSHV and EBV target the NK cell recognition ligand MICB through their miRNAs [281]. To further evade host immune response, KSHV and EBV utilize miRNA to repress viral protein expression to establish latency [13, 282, 283]. Besides the repression of host cell apoptosis and immune response, KSHV and EBV are able to modulate the host cell with the involvement of viral miRNAs. Viral miRNAs play roles during the modulation as well. KSHV- and EBV-encoded miRNAs directly participate in host cell transformation by interfering with Akt, MAPK, Wnt, and TGFβ pathways [13, 15, 284-291].

Two non-tumorigenic human polyomaviruses, BKV and JCV, also express miRNAs to regulate their early transcripts and target NKG2D, an NK cell recognition ligand. Similarly, MCV expresses two mature miRNAs, MCV-miR-M1-3p and MCV-miR-M1-5p, from a single pre-miRNA which may have multiple pri-miRNAs [72, 76]. A promoter region upstream of the pre-miRNA in the early region (ER) with enriched H3K4me3 has been identified, and mutation of this region only partially abolishes MCV-miR-M1 expression, suggesting that late transcripts that read through the polyadenylation signal can also serve as pri-miRNAs [76]. MCV-miR-M1 plays a critical role in repressing early gene expression by degrading early mRNAs, which further inhibits MCV replication [72, 76]. This repression could be a mechanism for MCV to maintain long-term persistence in host cells [76]. Additionally, MCV-miR-M1-5p targets the host innate immunity by directly downregulating ND10 component SP100, which inhibits CXCL8 secretion and attenuates neutrophil migration towards MCV host cells [73]. Overall, these findings suggest that MCV miRNAs play a crucial role in regulating host and viral gene expression during viral infections.

1.3.7.2 Circular RNAs

Circular RNAs (circRNAs) are formed by covalently closed loops via back-splicing, originally considered mis-splicing products (**Figure 8**). However, circRNAs are now known to have essential biological functions and play a critical role in gene expression regulation at both transcriptional and post-transcriptional levels. The production of circRNAs through back-splicing is regulated by various factors, including splicing machinery abundance, activity, and cis-acting regulatory elements. Their expression is tissue-specific and developmentally regulated. CircRNAs can bind to miRNAs, acting as miRNA sponges, and interact with RNA-binding proteins, modulating their activity and localization. Although circRNAs do not contain the 5' cap structure of mRNA, some circRNAs have been shown to act as templates for protein translation via cap-independent translation [292].

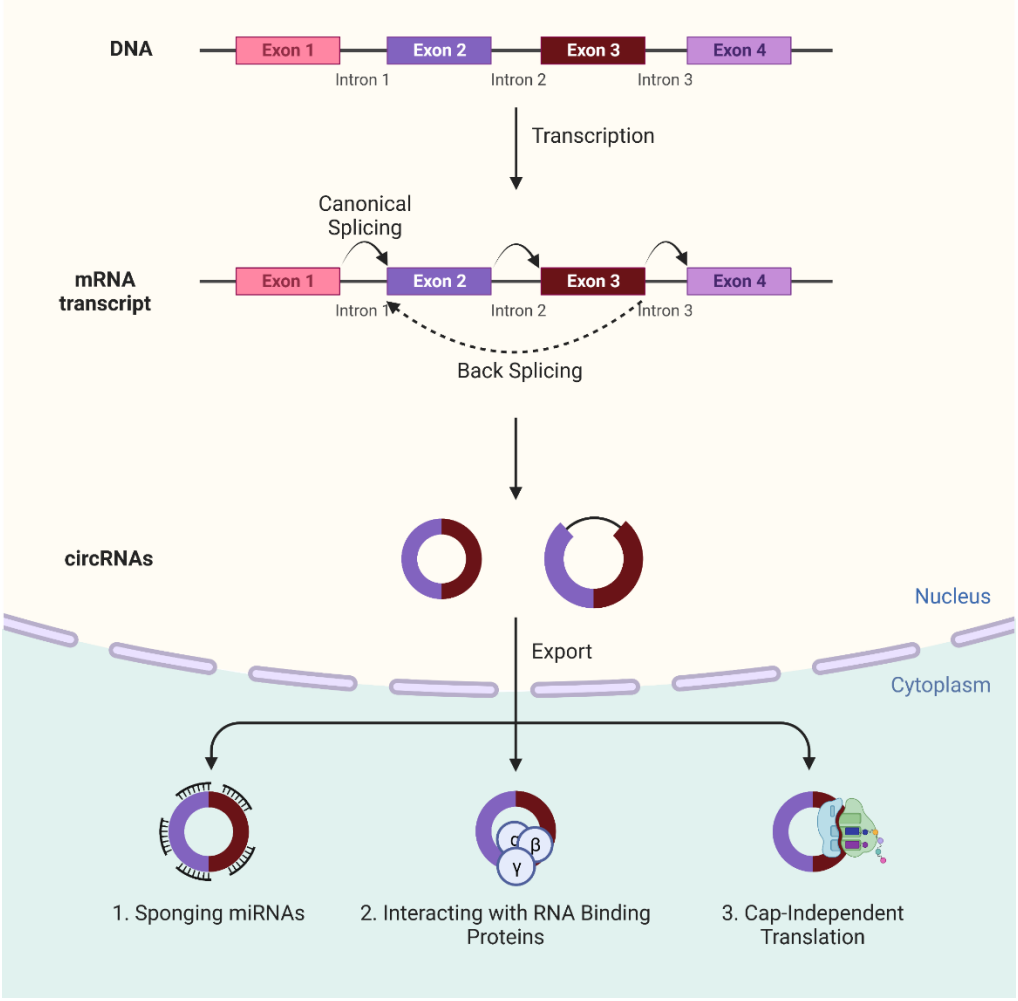


Figure 8 A Schematic of Circular RNA Biogenesis.

Circular RNAs (circRNAs) are generated through backsplicing, a process distinct from canonical (forward/linear) splicing, which results in covalently closed circRNA molecules with diverse exon and intron combinations. CircRNAs can further regulate cellular homeostasis by sponging miRNAs, interacting with RNA-binding proteins, and serving as templates for cap-independent translation. (Figure created with BioRender.com.)

In addition to cellular circRNAs, several tumor viruses, including KSHV (Chapter 2), MCV (Chapter 3), EBV, HPV, and HBV, have been identified to encode circRNAs (**Table 3**). These virally encoded circRNAs, like their cellular counterparts, have the ability to sequester cellular or viral miRNAs and engage in cap-independent translation to generate viral proteins. As

such, they are believed to contribute to the regulation of viral lifecycle and virus-associated oncogenesis. However, the functions of the majority of viral circRNAs remain poorly understood.

Table 3. Identified Human Tumor Viruses Encoded circRNAs

Viruses	CircRNAs	Locus	Reference
KSHV	circ-vIRF4IR	vIRF4	[293, 294]
	circ-vIRF4	vIRF4	[293, 294]
	circPAN/K7.3	PAN/K7.3	[293]
	kcirc3	ORF4/6	[295]
	kcirc29	PAN/K7	[295]
	kcirc38	ORF21/22	[295]
	kcirc54	ORF34	[295]
	kcirc55	ORF34/35/36	[295]
	kcirc57	ORF34/35/36/37	[295]
	kcirc97	ORF60/61/62	[295]
MCV	circMCV_1622_861 (circMCV-T, circALTO1)	ER	[150, 151]
	circMCV_2955_861 (circALTO2)	ER	[150, 151]
	circMCV_3337_861	ER	[150]
	circMCV_3913_3271	LR	[150]
EBV	circBART_1.1 (circRPMS1_E4_E2_E3a_I3a_E3b)	BART	[293, 296]
	circBART_1.2 (circRPMS1_E4_E2_E3a_ E3b)	BART	[293, 296]
	circBART_2.1 (circRPMS1_E4_ E3a_I3a_E3b)	BART	[293, 296]
	circBART_2.2 (circRPMS1_E4_ E3a_ E3b)	BART	[293, 296]
	circRPMS1	BART	[16, 297, 298]
	circRPMS1_E4_E3a	BART	[296]
	circRPMS1_E4_E2	BART	[296]
	circRPMS1_E7_E2	BART	[296]

	circRPMS1_E7_E3a	BART	[296]
	circRPMS1_E7_E5	BART	[296]
	circRPMS1_E7_E5.1	BART	[296]
	circRPMS1_E7_E5.2	BART	[296]
	circEBNA_W1_C1	EBNA	[296]
	circEBNA_W2_C1	EBNA	[296]
	circEBNA_U	EBNA	[296]
	circLMP2_E8_E2	LMP2	[296]
	circLMP2_E7_E2	LMP2	[296]
	circLMP2_E5_E3	LMP2	[296]
	circLMP2_E1_RPMS1_E3	LMP2/BART	[296]
	circLMP2A	LMP2	[17, 299]
	circLMP-2_e5	LMP2	[300]
	circBHLF1	BHLF1	[296]
	circBGLF1_BHLF1	BHLF1/BGLF1	[296]
	circBDLF1_BHLF1	BHLF1/BDLF1	[296]
	circBXLF1_BHLF1	BHLF1/BXLF1	[296]
	circBILF2_BHLF1	BHLF1/BILF1	[296]
	circBHRF1	BHRF1	[296]
HPV	circE7	E7	[301]
HBV	circRNA HBV_circ_1	pgRNA	[302]

As shown, Kaposi's sarcoma-associated herpesvirus (KSHV) has been found to express circRNAs, primarily derived from the viral interferon regulatory factor 4 (vIRF4)/K10 and PAN/K7.3 loci, with some originating from other genes at lower abundance [293-295]. Notably, KSHV circPAN, circK7.3, and cytoplasmic circ-vIRF4 can be packaged into KSHV tegument, suggesting their potential role as immediate early genes that modulate host cell signaling pathways, which will be discussed in details in chapter 2 [294]. Additionally, certain low-abundance KSHV

circRNAs, such as kcirc54, kcirc55, and kcirc97, have demonstrated the ability to regulate host cell growth and repress KSHV lytic gene expression. However, the underlying mechanisms are yet to be fully elucidated [295].

Recently, four circular RNAs (circRNAs) have been identified in MCV (Chapter 3) [150, 151]. However, the functions of MCV circRNAs remain controversial. Abere et al. demonstrated that the most abundant MCV circRNA, circMCV_1622_861, is targeted and degraded by MCV miR-M1 and does not exhibit protein-coding potential, as evidenced by the lack of association with polysomes and Western blot analysis, which will be further discussed in chapter 3 [150]. In contrast, Yang et al. reported that both circMCV_1622_861 and circMCV_2955_861 encode the ALTO protein, which can regulate the activity of specific promoters. Moreover, these two circRNAs can be enriched in exosomes and transferred to other cells [151]. The divergent findings regarding the functions and translational capacity of MCV circRNAs underscore the imperative for additional studies to elucidate the mechanisms and functional properties of these novel viral circular transcripts.

2.0 Kaposi's Sarcoma-Associated Herpesvirus-Encoded circRNAs Are Expressed in Infected Tumor Tissues and Are Incorporated into Virions

Work described in this chapter was published in mBio. 2020 Jan 7;11(1):e03027-19.

With authors Abere B., Li J., Zhou H., Toptan T., Moore P. S., and Chang Y.

B. Abere performed reverse-transcription PCR (RT-PCR), RT-qPCR, and KSHV viral particle production. B. Abere, J. Li, and T. Toptan performed the BaseScope RNA *in situ* hybridization (ISH). H. Zhou and B. Abere performed the circ-vIRF4 half-life assay. B. Abere wrote the manuscript. P.S. Moore and Y. Chang conceived the project and revised the manuscript.

As outlined in Chapter One, Kaposi's sarcoma-associated herpesvirus (KSHV), also referred to as human herpesvirus 8 (HHV-8), is a human tumor virus implicated in the development of Kaposi's sarcoma (KS), primary effusion lymphoma (PEL), multicentric Castleman disease (MCD) and KSHV inflammatory cytokine syndrome (KICS). KSHV has the capacity to infect various cell types, such as B cells, endothelial cells, and epithelial cells. Upon infection, KSHV can establish latent infection in host cells, characterized by slow viral replication and limited viral gene expression. Alternatively, the virus may enter a lytic phase, during which rapid replication and infectious virion production occur. The KSHV genome comprises double-stranded linear DNA, approximately 165-170 kb in length, and features a conserved gene cluster spanning open reading frames (ORFs) 4 to 75. Similar to other herpesvirus members, these conserved genes

typically encode proteins involved in viral replication and structural components. Furthermore, KSHV encodes a distinct set of genes, denoted with the prefix K (K1-K15), which serve multiple roles in viral infection and virus-induced diseases. The virus also produces 25 viral microRNAs (miRNAs) and several long non-coding RNAs (lncRNAs). Recent studies have identified the expression of multiple circular RNAs (circRNAs) from several KSHV genes, including viral interferon regulatory factor 4 (vIRF4)/K10, PAN, and K7.3. However, the properties and functions of these circRNAs remain largely unexplored.

To investigate this, we analyzed the expressions of circ-vIRF4, circK7.3, and circPAN in various samples. All three KSHV circRNAs were detected in all examined PEL cell lines, a subset of KSHV-associated primary tumors, and a small fraction of patient sera. Nonetheless, the expression of these circRNAs was differentially regulated. Upon induction of the KSHV lytic cycle, the levels of circPAN and circK7.3 transcripts, as well as linear vIRF4 mRNA, increased substantially. In contrast, the activation of circ-vIRF4 transcription displayed only modest enhancement or even reduction. These findings were further corroborated through *in situ* hybridization (ISH) analysis.

Considering that circRNAs may possess the capability to regulate host cells, we investigated whether KSHV-encoded circRNAs could be incorporated into the viral tegument, enabling their entry into host cells during infection and functioning as immediate early genes. We determined that the viral circRNAs circ-vIRF4, circK7.3, and circPAN were present in virus teguments, but not the intron-retention isoform of circ-vIRF4, circ-vIRF4.IR. Consequently, we assessed the distribution of circK7.3, circPAN, and circ-vIRF4 isoforms. Our findings revealed that circK7.3 and circPAN were located in both the nucleus and cytosol. However, circ-vIRF4.IR was predominantly found in the nucleus, while the exon-only isoform, circ-vIRF4, was primarily

distributed in the cytosol. This difference in localization may explain why circ-vIRF4.IR is not incorporated into viral particles. To further characterize circ-vIRF4, we measured its turnover rate and demonstrated that circ-vIRF4 has a 2.5-fold longer half-life compared to its linear counterpart.

In summary, our results indicate that circK7.3, circPAN, and circ-vIRF4 are present in PEL cell lines, KSHV-associated primary tumors, and patient sera. However, the regulation and distribution of these KSHV-encoded circRNAs differ. While circK7.3, circPAN, and circ-vIRF4 can be packaged into viral particles, circ-vIRF4.IR cannot. Furthermore, circ-vIRF4 exhibits a longer half-life than its linear transcript. Our study shows that KSHV circRNAs are present in KS tumors, and their expression is dependent on the stage of the virus's life cycle. KSHV circRNAs are also integrated into viral particles, indicating a possible role in the initial stages of infection.

2.1 Introduction

Kaposi sarcoma-associated herpesvirus (KSHV)/human herpesvirus 8 (HHV8) causes Kaposi sarcoma (KS), an angiogenic neoplasia of endothelial origin [303] and two B-cell lymphoproliferative disorders, primary effusion lymphoma (PEL) [304] and multicentric Castleman disease (MCD) [305]. Latent KSHV infection is typical for the virus in tumors and PEL cell lines, and is characterized by highly restricted viral gene transcription [306]. However, different viral transcription profiles, with less restricted gene expression and occasional virion production, have been found in different tumor types [307-313].

In addition to >90 KSHV protein-coding mRNAs, most of which are predominantly expressed during full-virus lytic replication, noncoding viral RNAs (ncRNAs) are also functionally important features of the KSHV transcriptome. These KSHV ncRNAs include 12 precursor

microRNAs (pre-miRNAs) that give rise to a total of 25 mature miRNAs [314-316]. These viral miRNAs are expressed during latency [317, 318] and regulate host and viral gene expression to promote immune evasion, viral persistence, and disease progression [319-323].

A number of KSHV long noncoding RNAs (lncRNAs) [324-328] from intergenic and noncoding regions of the viral genome are also expressed — most abundantly during lytic replication or after de novo infection [306, 308, 329-331]. Most KSHV lncRNAs other than PAN run antisense to known open reading frames (ORFs). Notable among these are the antisense-to-latency transcript (ALT), which is transcribed antisense to the major viral latency locus; T3.0 and T1.2, which are oriented opposite to replication and transcription activator (RTA/ORF50); and K7.3, which runs antisense to PAN [314, 324, 325, 329, 332].

CircRNAs constitute a class of 3′-to-5′ covalently closed, cyclized RNAs derived through back-splicing (BS) of a pre-mRNA such that a donor splice junction (SJ) ligates to an upstream acceptor site [333]. CircRNAs thus lack a 5'cap or 3'polyA tail [333, 334]. CircRNAs have been found to act as miRNA sponges [335, 336], to sequester RNA-binding proteins [337-339], and to regulate isogenic transcription and splicing [333, 337, 340, 341] and may generate protein products through internal ribosome entry site (IRES)-driven or m6A-deriven 5'-cap-independent translation [301, 342, 343]. Recent studies also suggest that cellular circRNAs modulate innate immune responses [344-346]. CircRNAs are resistant to exonucleolytic decay and therefore have long half-lives compared to linear transcripts from the same gene [333, 347]. Some cellular circRNAs have been shown to be abundant in cancer tissues and liquid biopsies and might be useful biomarkers of disease progression or prognosis [348, 349].

KSHV encodes circRNAs from the K10 locus (circ-viral interferon regulatory factor 4 [circ-vIRF4]) and from the PAN and K7.3 loci [293, 295, 350]. One of the two circ-vIRF4 RNA

molecules has intron-retention (IR) [350] of the conserved intron that is spliced from the linear vIRF4 mRNA transcript. In addition to circ-vIRF4, a cluster of multiple, bidirectional KSHV circRNAs that do not correspond to known mRNA splice junctions are expressed from the PAN/7.3 locus [293]. Each PAN/K7.3 circRNA species is individually of low abundance, but, *en masse*, the species represent a major portion of KSHV-encoded circRNAs. Additional low-abundance KSHV circRNAs have been described as originating from ORF4, ORF6, ORF21, ORF22, ORF34, ORF35, ORF36, ORF37, ORF61, and ORF62 loci in de novo infections of endothelial cells [295]. We focus on the three KSHV circRNAs most abundantly identified in sequencing of naturally infected PEL cells.

In this study, we characterized the expression profile of KSHV-encoded circRNAs in a panel of PEL cell lines, primary KSHV-associated tumor tissues, and patient-derived liquid biopsies specimens. We found KSHV circRNAs to be ubiquitously but differentially expressed in PEL cell lines. They are incorporated into KSHV virion particles produced from BJAB-rKSHV.219 cells, suggesting a function for viral circRNAs at the initial steps of primary infection that operates in a fashion similar to that seen with the tegument proteins which are involved in innate immune modulation for successful establishment of KSHV-infection [351]. Consistent with the predominantly cytoplasmic nature of KSHV virion tegument and envelope acquisition, only a cytoplasmic form of circ-vIRF4 is packaged into the viral particle. KSHV circRNAs are detected at a higher rate than the linear latency-associated nuclear antigen (LANA) transcript in primary KS biopsy specimens as well as in patient-derived sera and plasma; however, viral circRNA detection is highly susceptible to RNA degradation during tissue and liquid biopsy processing. Nevertheless, RNA decay analysis revealed that the rate of turnover of KSHV circ-vIRF4 is twice that of linear vIRF4 as revealed by RNA polymerase (Pol) II inhibition.

2.2 Materials and Methods

2.2.1 Cell Lines

KSHV- and EBV-co-infected BC1 [304], JSC1[352]; KSHV-positive BCBL1[353], BC3 [354] and BCP1 [355]; and EBV/KSHV-negative BJAB [356] cell lines were obtained from the American Type Culture Collection (ATCC). KSHV and EBV co-infected BBG1 [357] were provided by Shou-Jiang (SJ) Gao (University of Pittsburgh, PA, USA), while BJAB-rKSHV.219 cells[358, 359] (BJAB infected with a recombinant KSHV.219 virus [360]) are a generous gift of Thomas F. Schulz (Hannover Medical School, Hannover, Germany). All cell lines were maintained in Roswell Park Memorial Institute (RPMI) 1640 media (Cellgro) supplemented with 10% Fetal Bovine Serum (FBS) (VWR Seradigm). BJAB-rKSHV.219 cells were maintained with 4.2 µg/ml of puromycin selection in RPMI supplemented with 10% FBS. To induce KSHV lytic reactivation, all PEL cells were treated with 20 ng/mL of TPA and 3 mM NaB for 48 h.

2.2.2 RNA Isolation, DNase and RNase R Treatment

Total RNA was isolated from tumor samples, cell lines, virus preparations or cellular/virus preparation fractions using TRIzol or TRIzol-LS (Invitrogen) and processed with TURBO DNA-Free kit (catalog no. AM1907, Invitrogen) following the manufacturer's protocols. One µg of DNase digested RNA was then treated with 4 units (U) of Ribonuclease (RNase) R (catalog no. RNR07250, Lucigen) in 1X RNase R reaction buffer at 37 °C for 30 min in the presence of 8U of RiboLock RNase inhibitor (catalog no. EO0381, Thermo scientific); for untreated samples,

nuclease free water was added to the reaction instead of RNase R. To inactivate the RNase R, all samples were then incubated at 65 °C for 20 min.

2.2.3 Semiquantitative and Quantitative RT-PCR

A 1-µg volume of DNase-digested total RNA that had been either treated with RNase R or left untreated was reverse transcribed with random hexamers by the use of a SuperScript IV first-strand synthesis system (catalog no. 18091; Invitrogen) according to the manufacturer's protocol. Semiquantitative PCR was then performed using either Q5 high-fidelity DNA polymerase (catalog no. M0491; New England BioLabs/NEB) for circRNA detection or *Taq* DNA polymerase with ThermoPol buffer (catalog no. M0267; NEB) for linear RNA detection, according to the protocols specified for the products.

TaqMan universal master mix II with uracil-N-glycoslyase (UNG) (catalog no. 44400; Applied Biosystems) was used for quantitative PCR (qPCR) analysis of cDNA samples, while an iTaq universal SYBR green one-step kit (catalog no. 172-515; Bio-Rad) was used for direct quantitation of RNA derived from serum and plasma. Threshold cycle (C_T) values were used to calculate the mRNA fold changes according to the delta-delta C_T method. The RNase P C_T value was used as a reference. PCR primers and probes are listed in **Table 4** and **Table 5**.

Table 4 PCR primers

Primer Name	Sequence
circ-vIRF4_F (DP)	GCGCTCTCCAAGTATACGTGGCAA
circ-vIRF4 R2 (DP)	CAAATGCATGGTACACCGAATAC
vIRF4 C-term F	GAGTGGCGTTTCTGGAGTAT
vIRF4 C-term R	CTTCGCATACCGCGTCTTA

linear LANA R	GTTTAGTGTAGAGGGACCTTGGG
linear LANA F	TCTCCATCTCCTGCATTGCC
linear RTA/ORF50 F	CACCCCCCGATTCAACAGTC
linear RTA/ORF50 R	CCTCTCTTTGCTTCTCTGCTTTCG
circPAN divF1 (DP)	CGCCCACTGGTGTATCAGA
circPAN divR1 (DP)	AATCGCAGCTTTTGTTCTGC
circK7.3 divF2 (DP)	ACGGAAAACCTAGCCGAAAG
circK7.3 divR2 (DP)	CGGGTTATTGCATTGGATTC

Table 5 qPCR primers and probes

Primer/probe Name	Sequence
circ-vIRF4_qPCR_F (DP)	TGGCGATATAACGACTGAACAGAA
circ-vIRF4_qPCR_R (DP)	CAAATGCATGGTACACCGAATACC
circ-vIRF4_BSJ_P	6FAM-ATCTACCTCAGCCCCCGCGC-QSY
C-termin-vIRF4_qPCR_F2	ACGGATTCCCAATATTTTGCTC
C-termin-vIRF4_qPCR_R2	GTTCATAGCGGGAATTTGGATT
C-termin-vIRF4_TaqP2	ABY-TCAAGTGGAAGGCTGGTGGTTTGGT-QSY
POP4 (RNase P)	VIC-MGB Hs00198357_m1, Catalog number: 4331182

2.2.4 Virus Production and Histodenz Gradient Purification

The rKSHV.219 virus was produced as previously described [358, 361]. Briefly, to induce the KSHV lytic cycle, BJAB-rKSHV.219 cells were inoculated at a density of 6×10^5 cells/ml in a total volume of 800 ml culture in the presence of 3 μg/ml of anti-human IgM antibody (catalog

no. 2020-01; SouthernBiotech) and grown for 4 to 5 days in Techne spinner flasks (Cole-Parmer GmbH, Wertheim, Germany) with 60 rpm agitation. Culture supernatant was collected after low-speed centrifugation was performed to remove cells and debris and was concentrated by ultracentrifugation at 15,000 rpm for 5 h in a type 19 rotor (Beckman Coulter Inc., Brea, CA, USA). The pelleted virus was resuspended in 1× phosphate-buffered saline (PBS) overnight, layered onto a 20% to 35% Histodenz (D2158; Sigma) gradient, and centrifuged at 71,000 × g for 2 h using an SW41Ti rotor (Beckman). The desired fractions were collected, diluted in 1× PBS, and centrifuged at 29,000 × g for 2 h using an SW 32 rotor (Beckman). The pellets from each fraction were resuspended in 1× PBS and layered onto the same Histodenz gradient for a second round of purification. Each purified fraction (bottom, top, and supernatant) was then divided in two. One group was treated with 1% Triton X-100, while the other group was left untreated. Both groups of samples were then treated with 15 U/μl of micrococcal nuclease (NEB) in a 500-μl total volume and incubated at 37°C for 20 min. EGTA (5 mM) was added to each sample to stop the reaction, and each was then placed on ice for 30 min. Each sample was further divided into two tubes of 250-μl sample volume, and 750 μl of TRIzol LS reagent was added to each tube for RNA extraction.

The infectious virus titer was determined for each gradient fraction by applying serial dilutions of each preparation on HEK-293 cells as described previously [362]. Three days after infection, the number of GFP-positive cells was counted to calculate the KSHV infectious virus titer for each fraction.

2.2.5 Nuclear/Cytoplasmic Fractionation

Nuclear/cytoplasmic fractionation was performed using 1×10^7 BC1 cells with NR-PER nuclear and cytoplasmic extraction reagent (catalog no. 78833; Pierce), according to the manufacturer's protocol in the presence of RiboLock RNase inhibitor (catalog no. EO0381; Thermo Scientific). A 750-µl volume of TRIzol reagent was added to 250 µl of each fraction for RNA extraction. A 1-µg volume of total RNA from each fraction was used for cDNA synthesis, and the expression level of the indicated circRNAs in each fraction was analyzed by RT-PCR. The quality of the fractionation assay was controlled by immunoblotting for a nuclear marker, ORC-2 (catalog no. 551178; BD Biosciences), and for a cytoplasmic marker, LAMP-1 (eBioH4A3; eBioscience).

2.2.6 Tumor Samples and Sera/Plasma

Seventeen pathologically confirmed freshly frozen (FF) or optimal cutting temperature (OCT) embedded KS tissue specimens and 5 KS FFPE tissue microarrays (TMA) (1 visceral tissue microarray, 1 oral tissue microarray, and 3 skin KS tissue microarrays) were obtained from the AIDS and Cancer Specimen Resource (ACSR; San Francisco, CA, USA). Tissues were snap-frozen and stored in liquid nitrogen until use, while the TMAs were used within 2 months of sectioning. Ten archival serum samples and 4 fresh parallel serum samples and plasma samples from patients with AIDS-associated Kaposi's sarcoma were obtained as by-products of diagnostic or therapeutic procedures performed at the Columbia University College of Physicians and Surgeons and at the University of Pittsburgh Medical Center (UPMC) with the approval of the Institutional Review Board. Specimens were deidentified prior to use in our study.

2.2.7 Cell Pellet Preparation.

At 24 h after lytic induction, BC1/BJAB/BJAB-rKSHV.219 cells were harvested and pellets were fixed in 10% neutral buffered formalin and processed with HistoGel (catalog no. HG-4000-012; Thermo Scientific) for routine histology. Formalin-fixed paraffin-embedded cell pellets were then arrayed in a single block to minimized slide-to-slide staining variations. Sections (5 μm) were cut and stained according to the BaseScope protocol described below.

2.2.8 BaseScope RNA *In Situ* Hybridization (ISH)

BaseScope RNA ISH was performed using either a BaseScope duplex detection reagent kit (catalog no. 323810) or a BaseScope Detection Reagent v2-RED kit (catalog no. 323910; Advanced Cell Diagnostics) with probes targeting circRNA BSJ or linear RNAs (see **Table 6** for the list of probes used). FFPE section slides were baked in a dry oven for 1 h at 60°C, followed by deparaffinization performed with xylene twice for 5 min each time and with 100% ethanol twice for 2 min each time. Slides were then baked for 5 min at 60°C and incubated with RNAScope hydrogen peroxide for 10 min at room temperature followed by two 2-min washes in wash buffer and then with RNAScope 1× target retrieval reagent for 15 min at 98°C. Following target retrieval, slides were transferred into 100% alcohol for 3 min, dried at 60°C, and treated with RNAScope protease IV for 30 min at 40°C. After washing was performed, slides were hybridized with specific BaseScope probes (**Table 6**) for 2 h at 40°C in a HybEZ oven (Advanced Cell Diagnostics) followed by amplification comprised of incubations with AMP1 for 30 min at 40°C, AMP2 for 30 min at 40°C, AMP3 for 15 min at 40°C, AMP4 for 30 min at 40°C, AMP5 for 30 min at 40°C, AMP6 for 15 min at 40°C, AMP7 for 30 min at room temperature, and AMP8 for 15 min at room

temperature for red signal. For the green signal, after the AMP5 incubation step described above, slides were treated with AMP10 for 15 min at 40°C, AMP11 for 30 min at room temperature, and AMP12 for 15 min at room temperature. Slides were washed twice with 1× wash buffer for 2 min at room temperature between the incubation periods. Chromogenic signals were developed by incubation with either BaseScope Duplex Fast Red or BaseScope Duplex Green solution for 10 min at room temperature in the dark and by counterstaining with 50% Gill's hematoxylin I staining solution (catalog no. HXGHE1PT; American MasterTech Scientific) for 1 min at room temperature. Slides were dipped in 0.02% ammonia water (221228; Sigma-Aldrich) 2 or 3 times, washed with deionized water, and mounted with VectaMount (64742-48-9; Vector Laboratories).

Table 6 List of BaseScope detection probes purchased from ACD

Target	BaseScope Probe	Part ID/Catalog#
Circ-vIRF4-BSJ	BA-V-KSHV-vIRF4-circRNA-C2	715981-C2
CircPAN-BSJ	BA-V-HHV8-T1.1-circRNA	719981
CircPAN-BSJ	BA-V-HHV8-T1.1-circRNA-O1	719991
CircPAN-BSJ	BA-V-HHV8-T1.1-circRNA-O2	720001
CircK7.3-BSJ	BA-V-HHV8-K7-sense-circRNA-O1-C2	719951-C2
CircK7.3-BSJ	BA-V-HHV8-K7-sense-circRNA-O2-C2	719961-C2
CircK7.3-BSJ	BA-V-HHV8-K7-sense-circRNA-O3-C2	719971-C2
Linear LANA	BA-V-KSHV-LANA-3zz-st	716011

circPAN and circK7.3 were detected using a mixture of three 1zz probes targeting the three most abundant circPAN BSJs as determined by sequencing or using a mixture of three 1zz probes targeting the three most abundant circK7.3 BSJs. Circ-vIRF4 was detected with a single 1zz probe for its BSJ, and a single 3zz probe was used to detect linear LANA RNA. Control probes for the human housekeeping genes PPIB and POLR2A and a probe against the bacterial DapB gene were

used as positive and negative controls, respectively. Chromogenic signals were visualized with an Olympus AX70 microscope, and the images were acquired using a QImaging QIClick charge-coupled-device (CCD) camera and Q-Capture Pro 7 software and scored as follows: a minus sign ("–") represents a negative result, and one to four plus signs ("+" to "++++") represent positive results corresponding to the total number of the dots or clumps (aggregated dots) counted in each 1.5-mm diameter tissue microarray spot. Scoring results are presented as follows: minus sign ("–"), no dots; one plus sign ("+"), 1 to 5 dots and no clumps; two plus signs ("++"), 6 to 10 dots or 1 clump; three plus signs ("+++"), 11 to 30 dots or 2 to 5 clumps; four plus signs ("++++"), >30 dots or > 5 clumps. Background signal for the negative-control probe that targeted bacterial DapB varied greatly depending on source, fixation, and tissue processing of the microarrays. Therefore, as a conservative measure, cases were evaluated only if they (i) were assigned a score of "–" or "+" with the negative-control probe and (ii) were assigned a score of "+" to "++++" with the positive-control probe. Representative images were captured from each group.

2.2.9 RNA decay analysis of RNA polymerase II inhibition.

BCBL1 cells were inoculated at a density of 2×10^6 cells/ml using fresh medium in the presence of 2 μM flavopiridol, and then 5 ml of cell suspension was divided into aliquots and placed in 15-ml conical tubes. The tubes were loosely capped and incubated at 37°C. Cells were collected at 0, 1, 2, 3, 4, and 6 h after inhibitor treatment by centrifugation at 400 × g for 5 min. The cell pellet was then lysed in 1 ml of TRIzol reagent and stored at –80°C until the RNA extraction step. Threshold cycle (C_T) values were used to calculate the mRNA fold changes according to the delta C_T method, and the result obtained at each time point was normalized to the

0-h control within the same group. The C_T value calculated for RNase P was used as a reference to show that the cell numbers were consistent and that global RNA homeostasis was not affected by the presence of flavopiridol.

2.3 Results

2.3.1 KSHV-Encoded CircRNAs Are Ubiquitously but Differentially Expressed in PEL Cells.

Expression profiles of KSHV-encoded circRNAs were assessed in PEL tumor-derived cell lines infected with KSHV alone (BCP1, BC3, and BCBL1) or dually infected with KSHV and Epstein-Barr virus (EBV) (BC1, JSC1, and BBG1), as well as in BJAB-rKSHV.219, a BJAB cell line stably infected with a recombinant virus (rKSHV.219). BJAB cells negative for KSHV and EBV were used as a negative control. Reverse transcription-PCR (RT-PCR) was performed using junction-spanning, divergent primer (DP) pairs to detect circRNAs or conventional convergent primers (CP) to detect corresponding linear viral transcripts. Oppositely stranded circPAN and circK7.3 RNAs are generated from both overlapping and nonoverlapping transcripts [293] and no individual divergent primer pairs can be designed to detect all KSHV circPAN and circK7.3. Instead, we designed one divergent primer pair each for circPAN and circK7.3 that can detect the majority of the circRNA species.

Transcripts of all three KSHV-encoded circRNAs (circ-vIRF4, circPAN, and circK7.3) as well as linear vIRF4 transcripts were found to be expressed across all tested PEL cell lines regardless of KSHV replication status (**Figure 9A**). Reactivation of KSHV lytic replication by

treatment with 20 ng/ml tetradecanoyl phorbol acetate (TPA) and 3 mM sodium butyrate (NaB) was confirmed by expression of the master lytic switch RTA gene (ORF50) transcript as well as by increased levels of linear vIRF4 (K10). In comparison, linear LANA (ORF73) transcript levels were not affected by replication induction (**Figure 9A**). As expected, levels of linear RTA, vIRF4, and LANA RNAs were generally diminished by treatment with RNase R (4 units [U]). In contrast, levels of KSHV circRNAs were stable or were even increased by RNase R treatment, presumably due to reduced PCR template competition. No obvious systematic differences in KSHV circRNA expression levels were found in the KSHV-only PEL cell lines compared to those having EBV coinfection.

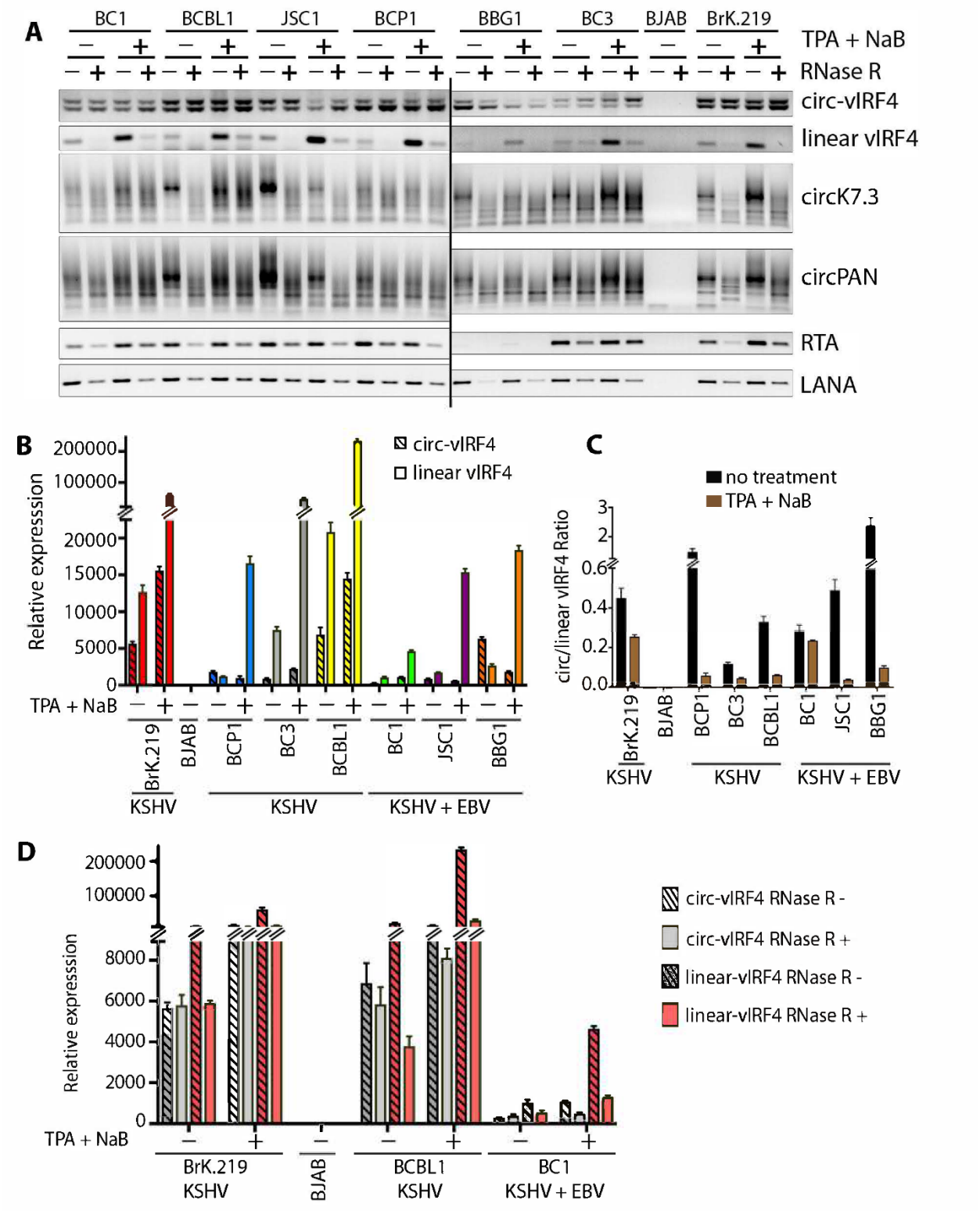


Figure 9 Expression profile of KSHV circRNAs in a panel of PEL cells.

Total RNA was extracted from PEL cells with or without induction of the KSHV lytic cycle using a combination of TPA and NaB. (A) RT-PCR was performed on RNase R-treated (+) or untreated (-) samples using divergent primer pairs corresponding to circ-vIRF4, circPANs, or circK7.3, while linear transcripts (vIRF4, LANA, and RTA) were amplified using a pair of convergent primers specific to each gene. (B) Quantitative comparison of circular and linear vIRF4 expression in cDNA from RNase R (-) samples processed as described for panel A using a TaqMan-based qPCR with a probe spanning circ-vIRF4 junction and a probe corresponding to the 3′ end of linear vIRF4, outside the circRNA coding sequence. (C) Ratio of circ-vIRF4 expression to linear vIRF4 expression in the experiment described for panel B. (D) Comparison of circ-vIRF4 expression and linear vIRF4 expression in the presence or absence of RNase R treatment. CT values were normalized with the RNase P CT value to calculate delta CT, and all samples were normalized to BJAB cells to calculate delta-delta CT (relative expression) values. Bar graphs in panels B, C, and D represent means ± standard deviations (SD) of results from three replicates.

To examine KSHV circRNA expression in greater detail, we focused on the circ-vIRF4 transcript. Quantitative RT-PCR performed using the same set of samples as those represented in **Figure 9A**, together with TaqMan probes corresponding to the circ-vIRF4 back-splice junction (BSJ) and the 3' end of linear vIRF4 transcript showed differing levels of circ-vIRF4 and linear vIRF4 expression across uninduced PEL cells (**Figure 9B**). BCBL1 cells, which show high levels of spontaneous KSHV replication [327], and BJAB-rKSHV.219 cells infected *in vitro* both expressed high levels of circ-vIRF4 and linear vIRF4. In contrast, BC1 cells, which have a 33-kb genomic duplication [363], are tightly latent [327] and showed the lowest levels of circ-vIRF4 and linear vIRF4 expression (**Figure 9B**). While most cell lines preferentially transcribed more linear than circular vIRF4 RNA (assuming similar transcript reverse transcription efficiencies), that pattern was reversed for the BCP-1 and BBG1 cells, in which circ-vIRF4 transcripts were more abundant than the corresponding linear transcripts (**Figure 9B and C**). This is consistent with

previous findings showing that most but not all cellular circRNAs are less abundant than the corresponding linear transcripts [347].

While induction of the KSHV lytic cycle markedly increased the numbers of circPAN and circK7.3 transcripts (**Figure 9A**) as well as the linear vIRF4 mRNA levels (**Figure 9A and B**) in all cell lines examined, circ-vIRF4 transcription was activated only marginally or was even reduced (**Figure 9A and B**). This is more clearly shown in **Figure 9C**; the ratio of circular transcripts to linear transcripts from the K10 locus was reduced for all cell lines, suggesting independent mechanisms for regulation of circ-vIRF4 and linear vIRF4 RNA transcriptional biogenesis. Alternatively, it is possible that circ-vIRF4s are induced by reactivation but to a lesser extent than the linear vIRF4 mRNA.

Quantitative analysis of circ-vIRF4 and linear vIRF4 levels in RNase R-treated versus untreated control samples revealed that circ-vIRF4 was resistant to the exonuclease treatment in the unreactivated cells (**Figure 9D; see also Figure 9A**). However, similar analysis of RNA from reactivated cells showed an increase in RNase R sensitivity of circ-vIRF4 whereas linear vIRF4 levels decreased dramatically in response to RNase R treatment regardless of the KSHV reactivation status (**Figure 9A, B, and Figure 10**).

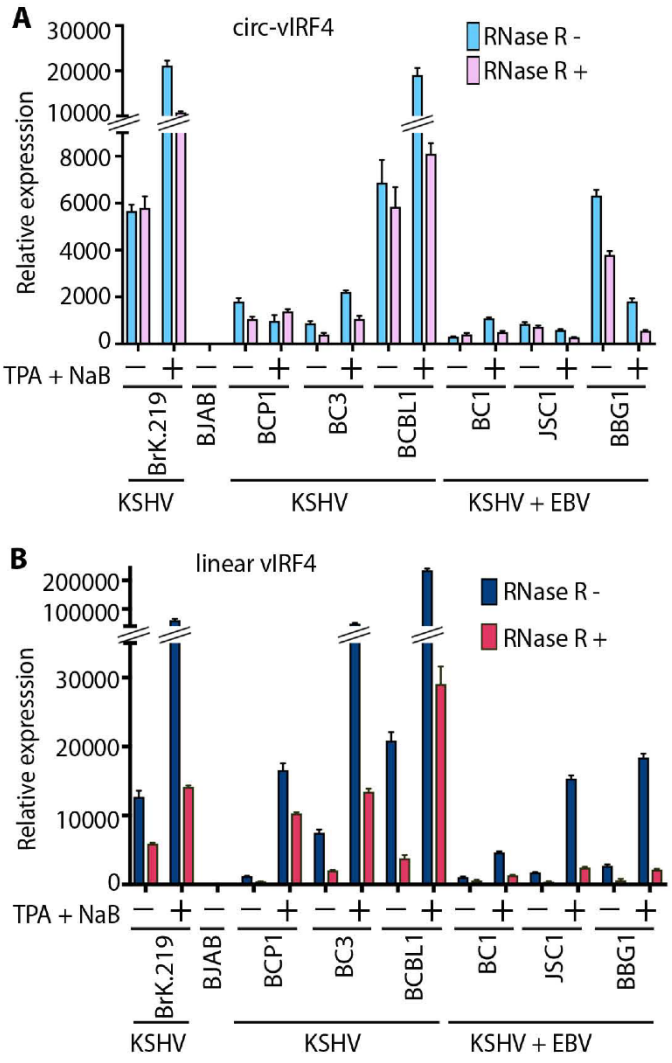


Figure 10 Expression profile of KSHV circRNAs in a panel of PEL cells.

Total RNA was extracted from PEL cells with or without induction of the KSHV lytic cycle using a combination of TPA and NaB. RT-PCR was performed on samples treated with RNase R (+) or left untreated (-) by the use of divergent primer pairs corresponding to circ-vIRF4, convergent primers corresponding to linear vIRF4, a probe spanning the circ-vIRF4 junction, and a probe corresponding to the 3′ end of linear vIRF4, outside the circRNA coding sequence. (A and B) Comparison of levels of circ-vIRF4 (A) or linear vIRF4 (B) expression in the presence or absence of RNase R treatment. CT values were normalized to the RNase P CT value to calculate delta CT values, and all samples were normalized to BJAB cells to calculate delta-delta CT (relative expression) values. The bar graphs shown in panels A and B represent means ± standard deviations (SD) of results from three replicates.

2.3.2 Virally encoded circRNAs are packaged into KSHV particles.

To determine whether KSHV-encoded circRNAs are packaged into virions, KSHV virions were produced from BJAB-rKSHV.219 cells induced with anti-human IgM antibody and purified through a 20% to 35% Histodenz gradient as schematically shown in **Figure 11A**. The rKSHV.219 molecular clone encodes a green fluorescence protein (GFP) under the control of a human elongation factor 1-α promoter from a red fluorescent protein (RFP)/GFP/PURO expression cassette inserted between ORFK9 and ORF57 [360], allowing measurement of infectious activity by fluorescence microscopy.

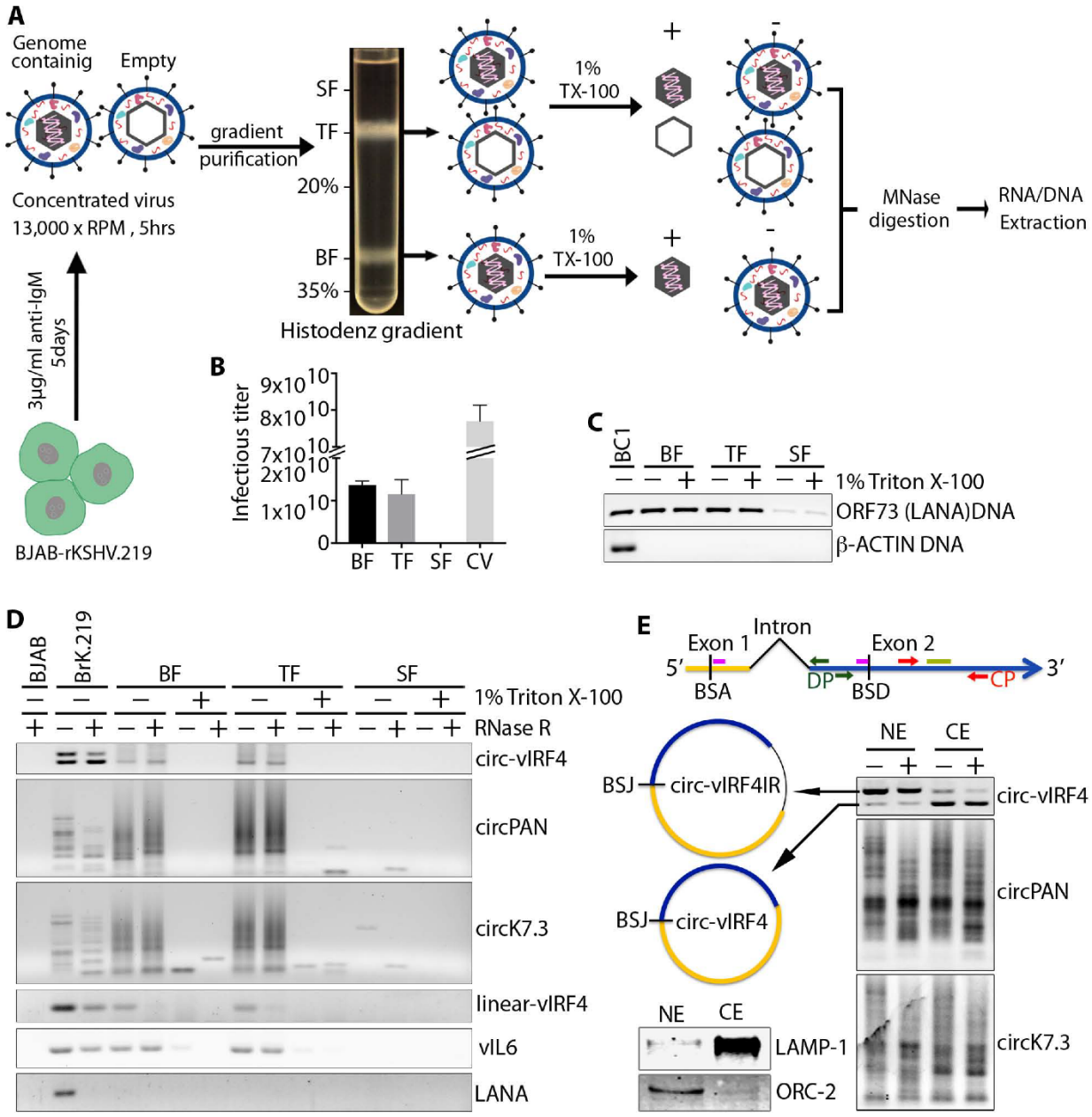


Figure 11 KSHV circRNAs packaged into viral particle.

A recombinant KSHV virus was produced from BJAB-rKSHV.219 cells and purified on a 20% to 35% Histodenz gradient. Histodenz fractions (BF, bottom fraction; TF, top fraction; SF, supernatant fraction) were treated with 1% Triton X-100 or left untreated. (A) Experimental workflow. (B) Total numbers of infectious KSHV particles from each Histodenz fraction and the unpurified crude virus (CV) preparation. Bar graphs in panel B show means ± SD of results from three replicates. (C) LANA/ORF73 PCR to detect KSHV genome in DNA extracted from the three different Histodenz fractions with and without Triton X-100

treatment. Genomic DNA from BC1 cells was used as a positive control. (D) RT-PCR detection of KSHV circRNA (circ-vIRF4, circPAN, and circK7.3) and linear RNA (vIRF4) in Histodenz fractions with and without detergent treatment. Linear viral transcript K2/vIL6 was used as a positive control and was packaged into viral particles, while LANA/ORF73 RT-PCR served as a negative control. (E) (Right panel) Nuclear/cytoplasmic fractionation in BC1 cells and KSHV circRNA detection by RT-PCR using divergent primers. (Left panel) Schematic representation of the two differentially localized circ-vIRF4 forms (dark green arrows, DP primers; pink bars, junction-spanning TaqMan probe; red arrows, CP primers; light green bar, TaqMan probe) (top) and a Western blotting control for the fractionation assay using a nuclear marker (ORC-2) and a cytoplasmic marker (LAMP-1) (bottom). NE, nuclear extraction; CE, cytoplasmic extraction. Results are representative of at least 3 independent experiments.

A fast-sedimenting bottom fraction (BF) band containing mainly virions with encapsidated viral genome, a top fraction (TF) band composed of both genome-containing and empty viral particles [364, 365]; and the supernatant containing no infectious particles were collected and analyzed (**Figure 11A**). In this protocol, the amount of encapsidated, genome-containing virions in the TF is proportional to the sedimentation time. The infectious particles from these three fractions were titrated on HEK-293 cells. The bottom and top band fractions contained similar amounts of infectious viral particles as measured by the number of GFP-positive HEK-293 cells (**Figure 11B**). The bar graphs shown in **Figure 11B** indicate the absolute number of infectious KSHV particles from each fraction. The amount of mature virions in the top fraction is consistent with the relatively short sedimentation time used. In contrast, the amount of infectious viral particles in the supernatant fraction was negligible. The three fractions were left untreated or treated with 1% Triton X-100 and subjected to micrococcal nuclease (MNase) digestion to remove any unprotected RNA or DNA outside the viral particle.

KSHV genomic DNA was detected by LANA/ORF73 PCR in DNA extracted from both Triton X-100-treated and untreated samples from the bottom and top fractions, but not in the

supernatant fraction, indicating the integrity of KSHV capsid after detergent treatment (**Figure 11C**) [364]. RT-PCR analysis of RNA extracted from these fractions using divergent primers detected all three KSHV-encoded circRNAs (circ-vIRF4, circK7.3, and circPAN) (**Figure 11D**) in virion preparations from the samples not treated with Triton X-100 but not in those from treated samples, indicating that circRNAs are likely to be present in the tegument of viral particles protected from MNase-mediated degradation (**Figure 11C**). Consistent with a previous report [366], we also detected viral linear RNA vIL6/K2 but not LANA/ORF73 in the KSHV virion gradient fractions.

2.3.3 Only The Cytoplasmic Form of Circ-vIRF4 Is Incorporated into KSHV Virions.

One of the two circ-vIRF4 RNA molecules has intron retention (IR) [350] (**Figure 9A)** of the conserved intron that is spliced from the linear vIRF4 mRNA transcript. Nuclear cytoplasmic fractionation of BC1 cells showed differential subcellular localization of the two circ-vIRF4 forms (**Figure 11E**, upper-left panel): the larger intron-containing circ-vIRF4.IR was retained in the nucleus whereas the smaller, processed, exon-only circ-vIRF4 was exported to the cytoplasm (**Figure 11E**, right panel). Sequencing of the two PCR products for circ-vIRF4 confirmed that the top band (nuclear form) contained the depicted intron whereas the bottom band did not contain the intron. CircPANs and circK7.3s were found in both cellular compartments (**Figure 11E**, right panel). The quality of fractionation was monitored by Western blot analysis of cytoplasmic (LAMP-1) and nuclear (ORC-2) marker proteins (**Figure 11E**, left-bottom panel). Consistent with the mainly cytoplasmic nature of tegument acquisition and envelopment of KSHV virions [351, 367], only the cytosolic intron-less circ-vIRF4 (**Figure 11E**) was incorporated into KSHV particles (**Figure 11D**). Circ-vIRF4.IR was excluded from virus particles. Due to the number of circRNAs

from both the PAN and K7.3 loci, we could not assign these specific circRNAs to particular subcellular compartments; therefore, we were unable to determine whether only cytoplasmic forms of these circRNAs were packaged into viral particles (**Figure 11D**). Nevertheless, for circK7.3, there appeared to be increased accumulation of smaller circular forms in the cytoplasmic fraction (**Figure 11E**).

2.3.4 KSHV CircPANs Are Inducible Whereas Circ-vIRF4 Is Constitutively Expressed in KSHV-Infected B Cells *In Situ*.

BaseScope *in situ* hybridization (ISH) probes against the unique KSHV circRNA back-splice junctions were applied to HistoGel-embedded KSHV-infected BC-1, BJAB-rKSHV.219, and KSHV-infection-negative BJAB cell pellets. This was performed using a single probe against the circ-vIRF4 BSJ, three probes to hybridize to the most abundant circPAN BSJs, and three probes against circK7.3 BSJs. KSHV circRNA hybridization results were compared to those obtained using three side-by-side probes (3zz) targeting LANA RNA (see Materials and Methods; see also **Table 6** for the list of BaseScope probes used). Formalin-fixed, paraffin-embedded (FFPE) cell pellets were arrayed in a single block to minimize slide-to-slide staining variations (see Materials and Methods for the induction protocol for the various cell lines). Quality control was performed using probes corresponding to human housekeeping genes PPIB and POLR2A to ensure RNA integrity in all cell pellets (**Figure 12**, "+ Control" column). Using the bacterial DapB probe as a negative control, none of the pellets showed positive signals in cells (**Figure 12**, "– Control" column). BJAB cells were negative for all KSHV linear and circular RNA probes. The linear LANA RNA probe showed similar staining results in the BC-1 and BJAB-rKSHV.219 cells, with no significant changes in the average number of positive-scoring signal dots per cell or in the

percentage of positive-scoring cells upon induction compared to no-treatment conditions. Similar results were seen with the circ-vIRF4 and circK7.3 probes. In contrast, the circPAN probe showed a massive increase in signal under induction conditions that obliterated all hematoxylin counterstaining in the affected cells. Of note, the circ-vIRF4 probe showed strong and multiple dot signals in virtually every cell, and although the avidity of a probe with respect to its target can vary depending on the probe, it is clear that circ-vIRF4 BSJs were abundant in the uninduced cells and showed no further increase in abundance upon induction.

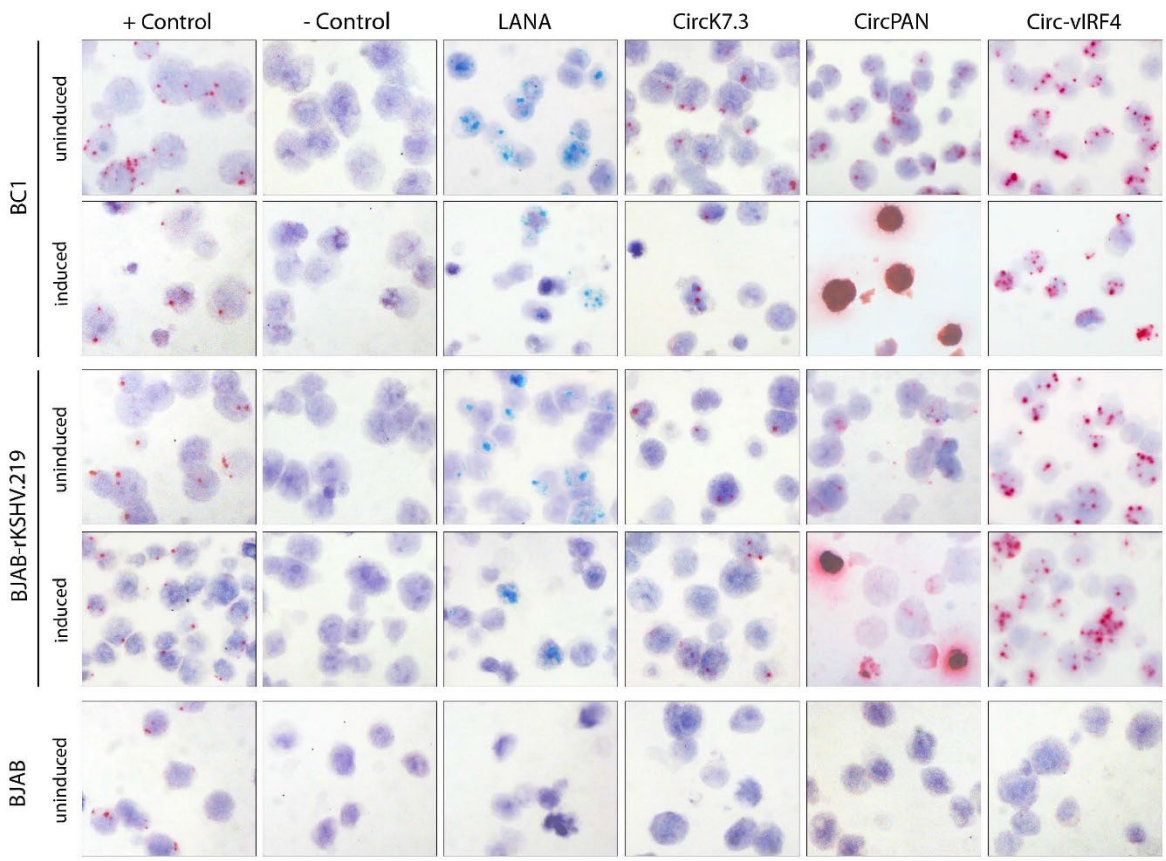


Figure 12 BaseScope RNA ISH detection of KSHV circRNAs in PEL cells.

Representative images from induced/uninduced PEL cells are shown. Uninduced BJAB cells were used as a KSHV negative control. Green signals show staining for linear RNA (LANA), while red signal represents detection of the represented KSHV circRNAs or the cellular control POLR2A. For the positive-control slide, only red signal (POLR2A) was developed, so no green signal (PPIB) is visible. Original magnification, ×40.

2.3.5 KSHV circRNAs are present in KSHV-associated primary tumors and patient sera.

RT-PCR detection of viral circRNAs in freshly frozen (FF) or optimal-cutting-temperature (OCT) embedded KS tissue samples showed that 14/17 (82.4%) samples were positive for circ-vIRF4 expression and 5/17 (29.4%) positive for circPAN/K7.3 using a divergent primer pair that can detect most circRNAs in both PAN RNA and the anti-sense K7.3 region (**Table 7**). Considering all three circRNAs together, the detection rate rose to 16/17 (94.1%) compared to only 14/17 samples positive for the linear LANA transcript.

Table 7 KSHV circRNA detection in KS patient derived tumor tissues and liquid biopsies by RT-PCR

KSHV transcript	FFT/OFT (n=17)	Serum (n=10)
circ-vIRF4	14 (82.4%)	4 (40%)
CircPAN/K7.3	5 (29.4%)	3 (30%)
CircPAN /K7.3 + circ-vIRF4	16 (94.1%)	5 (50%)
LANA	14 (82.4%)	NA

FFT: Fresh frozen tissue, OFT: OCT embedded tissue, n: number of tumors

BaseScope probes (see **Table 6**) against the unique KSHV circRNA back-splice junctions were used on archival formalin-fixed, paraffin-embedded (FFPE) KS tissues. To assess KSHV circRNA ISH, tissues were first assessed for RNA integrity and lack of nonspecific hybridization. A total of 181 archival KS skin, oral, or visceral tumors were screened. Of these, 92 were considered valid for analysis based on the tissue RNA quality determined by confirmation of both (i) positive staining corresponding to a human housekeeping gene (PPIB or POLR2A 1zz probes) and (ii) low levels of nonspecific background staining (fewer than 10 positive-scoring [+] dots per tissue for a negative-control bacterial gene [DapB 1zz probe]) (**Table 8**).

Table 8 KS tumors used for KSHV circRNA BaseScope-ISH detection

Tumor type	Total cases	False Positive	No tissue	Valid cases
KS	181	33	56	92
Non-KS Skin cancer	24	11	2	11

A total of 61 (66.3%) of 92 valid KS tissue biopsy specimens were positive for KSHV circRNAs, including 32 of 92 (34.8%) for circ-vIRF4, 49 of 92 (53.3%) for circPAN, and 28 of 92 (30.4%) for circK7.3. By comparison, 35 of 92 (38%) were positive for LANA mRNA (**Table 9**). Representative images of KS tumors are shown (**Figure 13**). The variations in the levels of KSHV circRNA expression in these tissues can be partially attributed to differences in RNA quality due to tissue processing and source; we also cannot rule out the possibility of differential expression at this point.

Table 9 KSHV circRNA detection in KS tumors by BaseScope-ISH

KSHV transcript	KS (n=92)
circPAN	49 (53.3%)
circK7.3	28 (30.4%)
circ-vIRF4	32 (34.8%)
circPAN + circK7.3 + circ-vIRF4	61 (66.3%)
LANA	35 (38.0%)

n: number of tumors

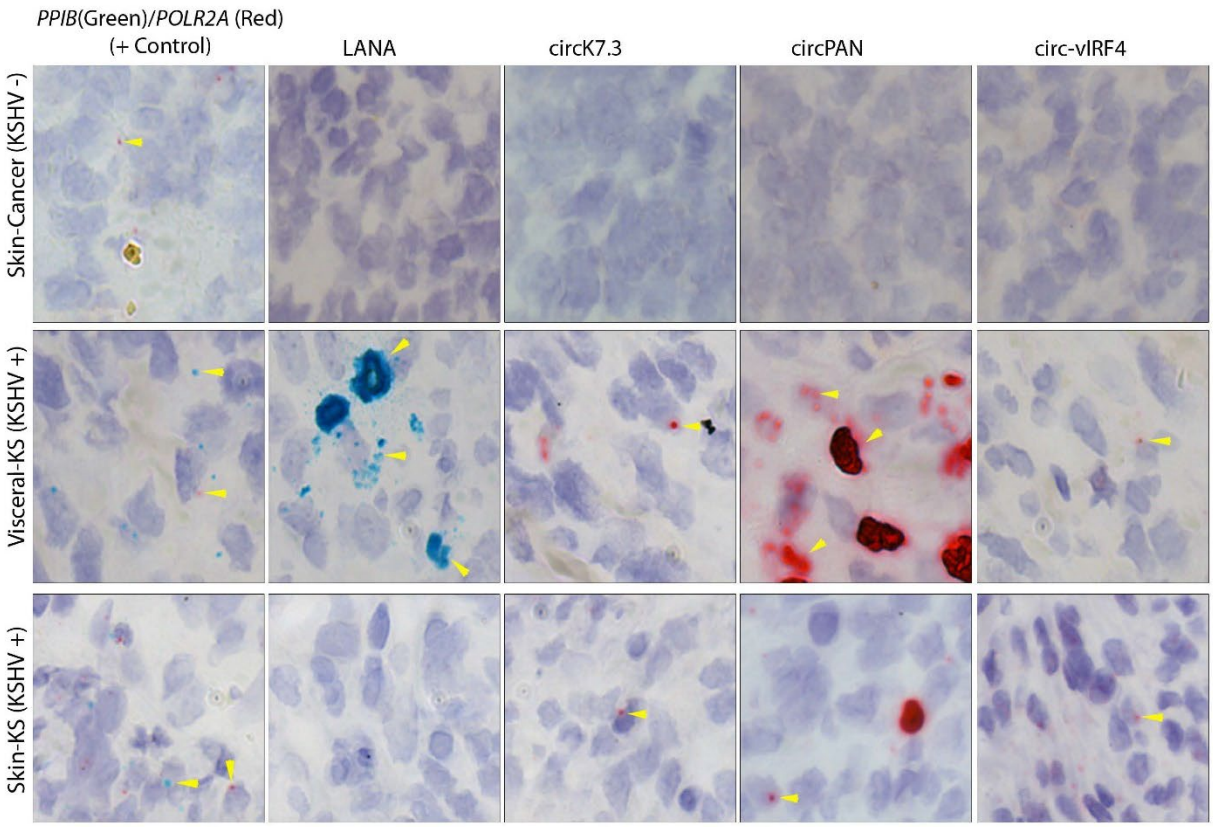


Figure 13 BaseScope RNA ISH detection of KSHV circRNAs in primary KS tumor tissues. Representative images from a visceral KS tumor tissue and a skin KS tumor tissue are shown. A malignant melanoma, non-KS skin cancer tissue, was used as negative control for the KS lesions. Green signals show staining for linear RNAs (control PPIB or LANA), while red signal represents detection of the represented KSHV circRNAs or the cellular control POLR2A. Arrowheads show positive staining from circular RNA or linear RNA. Original magnification, ×40.

2.3.6 KSHV circ-vIRF4 is more stable than its linear transcript.

To analyze circ-vIRF4 turnover, BCBL1 cells were treated with the RNA polymerase II inhibitor flavopiridol (2 μM) and RNA decay rates were determined by quantitative RT-PCR using TaqMan probes specific to the BSJ or to the linear RNA sequence outside the circRNA coding

region of circ-vIRF4. Probes were similarly designed for SMARCA5, a cellular gene control that produces both circular and linear RNAs. RNaseP linear RNA, transcribed by RNA pol III and therefore not susceptible to flavopiridol inhibition, was used for treatment control. Consistent with previous reports from studies comparing cellular circRNAs to their linear mRNA counterparts [333, 347], circ-vIRF4, with a calculated half-life of ~5 h, displayed a lower turnover rate than vIRF4 mRNA in the same samples, which had a half-life of ~2 h (**Figure 14A**). Similar values were found for cellular circSMARCA5, with a half-life of ~6 h, and linear SMARCA5 mRNA, with a half-life of ~4h (**Figure 14B**). The levels of RNase P RNA were unchanged throughout the 6-h flavopirodol treatment time course.

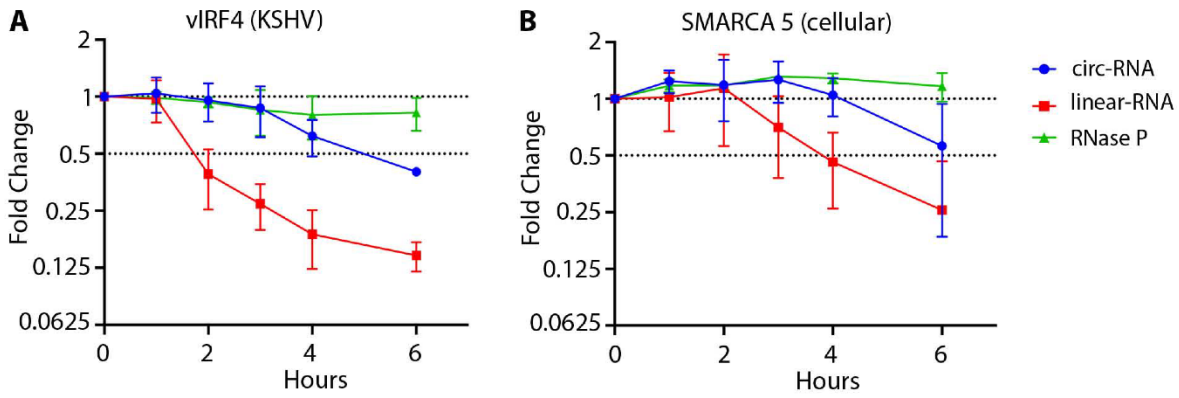


Figure 14 Circular versus linear vIRF4 RNA decay rates.

BCBL1 cells were treated with 2 µM flavopiridol, 1 × 10^7 cells were collected every hour for 6 h, and qRT-PCR was performed using a TaqMan probe spanning the circRNA junction and a probe corresponding to the 3' end of the linear transcript outside the circRNA coding sequence. (A) vIRF4. (B) A cellular control (SMARCA5). RNase P was used as a control that was not affected by RNA Pol II inhibition. Error bars show means ± SD of results from three replicates from 3 independent experiments.

2.4 Discussion

KSHV-encoded circRNAs represent a newly described category of viral RNA molecules [293, 295]; however, their function in the context of virus biology and their contribution to viral pathogenesis remain to be defined. To begin to investigate this, we characterized the expression profiles of circ-vIRF4, circPANs, and circK7.3s in a panel of PEL cell lines; measured whether these cyclized RNA molecules are packaged in KSHV virus particles; determined their stability; and assayed for their presence in patient-derived KSHV-associated tumor tissues/sera to assess their usefulness as clinical markers of infection.

KSHV circRNAs (circ-vIRF4, circPAN, and circK7.3) are ubiquitously but differentially expressed across PEL cell lines (**Figure 9**). Expression of circPANs and circK7.3s is induced upon KSHV lytic replication, which mirrors that of their linear lncRNA counterparts, PAN and K7.3. The lncRNA PAN is the best studied and has been functionally determined to interact with transcriptional and epigenetic modifiers to regulate host and viral gene expression [368-370]. The 1.08-kb PAN 3' region is polyadenylated but does not associate with polyribosomes and contains no known introns [326]. The lack of splicing in PAN transcripts suggests a novel mechanism for circular RNA formation from this viral precursor RNA. While PAN is detected at low levels in latent PEL cells [306], it becomes the most abundant viral transcript during lytic KSHV replication. Knockdown studies revealed that PAN is essential to efficient late lytic KSHV gene expression [371]. PAN mainly localizes to the nucleus but can also be found in the cytoplasm and packaged into KSHV virions, where it may be active in newly infected cells as an immediate early RNA [366, 369]. The K7.3 transcript, arising from the strand opposite PAN [327, 332], has unknown function(s). Like PAN, K7.3 RNA is induced during virus replication, albeit at much

lower abundance, and while it is assumed to be noncoding, this has not been formally examined [327].

In contrast to circPAN, circ-vIRF4 levels show no consistent changes upon induction of lytic replication, suggesting that its biogenesis is not directly regulated by transcription activity of the vIRF4 operon. KSHV vIRF4 represents one of four ORFs (vIRF1 to vIRF4) encoded by a cluster of interferon regulatory factor (IRF) homologs in the KSHV genome [363]. The vIRFs inhibit their cellular counterparts and are also involved in cell cycle regulation and oncogenic transformation [372, 373]. Expression of KSHV vIRF1, vIRF2, and vIRF4 is inducible, whereas KSHV vIRF3, which encodes a nuclear protein called LANA2, is latently expressed in all KSHV-infected B cells (65). LANA2 is detected in PEL and MCD [306, 374], but not in KS. Although expression of vIRF4 RNA and protein can be induced by TPA treatment in PEL cells [312, 375], circ-vIRF4 is detectable in unstimulated cells. However, unlike the results seen with its cognate linear transcript, circ-vIRF4 expression as evaluated by RT-PCR showed only a modest increase, or even a decrease, during lytic replication, which is consistent with latency expression (**Figure 9A, B and C**). Of note, by BaseScope *in situ* hybridization, the circ-vIRF4 probe showed strong and multiple dot signals in virtually every KSHV-infected PEL cell (**Figure 12**). Expression of circ-vIRF4 biomolecules was present at significant levels in uninduced cells and showed no further increase upon induction. The latent profile of circ-vIRF4 expression suggests a possible regulatory role for this circRNA. In most of the PEL cells, the level of circ-vIRF4 expression was lower than that seen with its linear counterpart, which is consistent with previous reports indicating that the majority of cellular circRNAs are expressed at lower levels than their cognate linear counterparts [347]. However, among the six PEL cell lines tested by RT-PCR, the BBG1 and BCP1 cells showed higher circ-vIRF4 expression than linear vIRF4 expression; also, circ-vIRF4

expression decreased after lytic induction in contrast to the linear transcript. This variation could be attributed to the differences in the origins of the cells (the BBG1 cell line was derived from a cutaneous B-cell lymphoma) or could represent a result of cell culture adaptation but could not be attributed to the number of KSHV genome copies, since BBG1 is estimated to have ~200 copies/cell and BCP1 contains ~150 copies/cell [376, 377]. Further experimentation on the kinetics of these circRNAs at different time points during KSHV lytic induction would shed more light on their biogenesis and potential function.

PAN RNA accumulates in the nucleus of infected cells during KSHV lytic replication, but it is also present in unstimulated KSHV-infected cells [308, 311]. A significant amount of PAN is localized to the cytoplasm and incorporated into KSHV virions [366, 369]. In addition to PAN, KSHV ORF59, ORF58, ORF54, ORF17, K2, K4, K5IR, K6, K7, K8.1, and K12 RNAs have been found packaged into KSHV virions in which they are directly released into newly infected cells and are presumed to act as immediate early genes early in infection [366]. To study whether KSHV circRNAs could be packaged into budding virions, recombinant rKSHV.219 was produced from stably infected BJAB cells using an anti-IgM antibody to cross-link the B-cell receptor (BCR) [358, 359]. Similar to PAN, circPANs, circK7.3s, and circ-vIRF4 were found packaged in nuclease-resistant viral particles (**Figure 11**). PAN RNA has already been implicated in modulation of the cellular immune response as well as viral gene expression to facilitate KSHV infection [368, 369, 378-381]. Similarly, viral microRNAs and proteins packaged in KSHV virions modulate innate immune functions during primary infection [351]. Experimentation will be needed to determine whether virion-packaged KSHV circRNAs might have similar roles.

Two forms of circ-vIRF4 are detectable in infected PEL cells: a larger form (632 nucleotides [nt]) generated through intron retention of the canonical linear vIRF4 intron

(102 nt) and a smaller form (530 nt) with a naturally occurring intron spliced out (**Figure 11D and E**). Our cytoplasmic/nuclear fractionation experiments showed that circ-vIRF4.IR localizes to the nucleus and was not found in virus particles whereas the exon-only form was cytoplasmic and retained in the virion tegument. This is consistent with herpesvirus tegumentation and the envelopment maturation that occurs in the cytoplasmic compartment [351, 367]. CircPANs and circK7.3s are also packaged into viral particles; however, given the high number and different sizes of circRNAs from these regions, we were unable to determine whether only cytoplasmic forms of these circRNA are incorporated into the KSHV virion.

KSHV-encoded circRNAs can be detected in patient-derived primary tumor tissues *in situ* and in liquid biopsy specimens, even in tissues that had been fixed and stored at room temperature for decades. RNA degradation occurring during tumor tissue fixation/storage as well as liquid biopsy preparation might also lead to underestimation of circRNA levels. Despite this remarkable retention of RNA in archival tissues, it is unlikely that circRNA detection can surpass the sensitivity of KSHV protein or antibody detection.

CircRNAs are more stable than the corresponding linear cognate RNAs mainly due to their covalently closed circular nature, which confers resistance to exonucleolytic decay. Estimates of the half-life of cellular circular RNAs range between 19 to 24 h, 2.5 times longer than that of their linear counterparts [333, 334, 347, 382]. Consistent with the stability of cellular circular RNAs, circ-vIRF4, with a half-life of 5 h in unreactivated BCBL1 cells, showed a half-life that was double that of its linear transcript (**Figure 14**). The stability of circ-vIRF4 relative to that of the corresponding linear mRNA points to the intriguing possibility of differential regulation of RNA biogenesis through circularization.

3.0 Merkel Cell Polyomavirus Encodes Circular RNAs (circRNAs) Enabling a Dynamic circRNA/microRNA/mRNA Regulatory Network

Work described in this chapter was published in mBio. 2020 Dec 15;11(6):e03059-20.

With authors Abere B., Zhou H., Li J., Cao S., Toptan T., Grundhoff A., Fischer N., Moore P. S., and Chang Y.

B. Abere and T. Toptan performed RNase R^+ RNA and sequencing data analysis. B. Abere performed the PCR and RPAD validation of identified MCV circular RNAs. J. Li performed the BaseScope *in situ* hybridization and polysome fractionation. H. Zhou, B. Abere, and S. Cao performed the Western blot and RT-qPCR. A. Grundhoff and N. Fischer collaborated on this project. B. Abere wrote the manuscript. P.S. Moore and Y. Chang conceived the project and revised the manuscript.

As mentioned in Chapter One, Merkel cell polyomavirus (MCV) is the most recently identified human tumor virus. In this study, we aimed to investigate whether MCV can produce circRNAs, similar to other known human tumor viruses, such as KSHV and EBV. To achieve this, we performed RNase R^+ RNA sequencing to identify RNase R resistant RNAs in MCV-positive Merkel cell carcinoma (MCC) cell lines (CVG-1 and MS-1) and 293 cells with a transfected MCV genome. We identified four circRNAs derived from the forward strand, circMCV_1622-861 (circMCV-T), circMCV_2955-861, circMCV_3337-861, and circMCV_3913-3271.

We further characterized the most abundant MCV circRNA, circMCV-T, using RT-PCR, RPAD, and *in situ* hybridization (ISH) in transfected 293 cells, MCV-positive MCC cell lines, and MCC patient-derived tumor samples. CircMCV-T was detected in 293 cells with an intact MCV genome, some MCV-positive MCC cell lines, and two out of three MCC patient-derived tumor samples. Interestingly, unlike other circRNAs, the majority of circMCV-T is not RNase R resistant. We then attempted to isolate the interactor that destabilizes circMCV-T, which proved to be the MCV-encoded miRNA, miR-M1. This miRNA has a perfect seed sequence match to circMCV-T. Co-transfection of miR-M1 greatly degraded circMCV-T, while knockout of the hairpin structure of miR-M1 increased circMCV-T stability. This suggests that circMCV-T can sequester miR-M1, the negative regulator of MCV early gene expression. Moreover, we examined whether circMCV-T can rescue MCV early gene expression and MCV replication by absorbing miR-M1. However, co-transfection of circMCV-T only induced a moderate increase in early gene mRNA levels and had no significant effect on MCV replication and protein expression. We also demonstrated that circMCV-T does not associate with polysomes, and no viral protein was detected in circMCV-T transfected cells by Western blot.

Additionally, we identified four circRNAs from RatPyV2 through RNase R^+ RNA sequencing, and one of these circRNAs is derived from the homologous region of circMCV-T. This suggests a shared posttranscriptional RNA processing mechanism between the two polyomaviruses.

In summary, we identified four circRNAs from MCV and four from RatPyV2. CircMCV-T, the most abundant MCV-encoded circRNA with a homologous counterpart from RatPyV2, was detected in 293 cells with an intact MCV genome, some MCV-positive MCC cell lines, and two out of three MCC patient-derived tumor samples. Due to miR-M1 targeting, circMCV-T is not

entirely RNase R resistant. However, circMCV-T exhibits a moderate effect on regulating MCV viral protein expression and replication and is not capable of protein expression. Thus, the precise function of circMCV-T remains to be determined. The findings of our research reveal a dynamic regulatory network involving MCV-encoded circRNA, miRNA, and mRNA. It highlights the complex functional modulation between different types of polyomavirus-encoded RNAs, which play a crucial role in controlling viral replication.

3.1 Introduction

Merkel cell polyomavirus (MCV) causes ~80% of Merkel cell carcinoma (MCC) tumors [38]. MCV is a nonenveloped, double-stranded DNA virus whose genome is divided into two bidirectionally oriented and sequentially expressed transcriptional units: the early region (ER) and the late region (LR). Generated through alternative splicing, MCV ER transcripts express tumor antigen (T-Ag) proteins, including large T antigen (LT), small T antigen (sT), 57,000-molecular-weight T antigen (57kT), and alternative large T antigen open reading frame (ALTO) or middle T-Ag (MT) [66, 67, 76]. Some of these ER-encoded proteins have characterized functions essential for MCV replication and MCC oncogenesis [66, 138, 146, 155, 159, 202, 209, 232, 235, 238, 240, 253, 383, 384]. The MCV LR produces transcripts that define structural proteins VP1 and VP2 necessary for MCV virion production [163].

MCV also generates noncoding RNAs (ncRNAs). A microRNA (miRNA) precursor is transcribed from the negative strand of the ER (miR-M1) and gives rise to two mature miRNAs, MCV-miR-M1-5P and MCV-miR-M1-3P [72, 76]. miRNAs are ~22 nucleotide (nt) molecules derived from precursor RNA, processed by Drosha and Dicer. Incorporated into RNA-induced

silencing complex (RISC), miRNA results in mRNA degradation or translation inhibition of its target transcript [385, 386]. MCV-miR-M1 shows perfect complementary sequence identity to a region in the early gene transcripts resulting in the degradation of these transcripts [72]. By regulating LT transcripts MCV-miR-M1 inhibits MCV replication, and thus contributes to the maintenance of the MCV genome as episomes in the infected host cell [72, 76].

Circular RNAs (circRNAs) are 3'-to-5' covalently cyclized RNA molecules that are produced through a back splicing mechanism [333, 334, 387, 388]. Because of their circular conformation, circRNAs are less susceptible to exonuclease activity resulting in their increased stability compared to their linear counterparts [333, 334, 347, 382]. When circRNAs were first identified in the early 1990s, they were considered a "mis-splicing" product [389]. However, in the past decade, circRNAs have been found to be critical cellular regulators during development and disease, with the potential to be used as biomarkers as a result of their differential tissue- and disease-specific expression profiles [336, 349, 390-393]. Several broad functions have been ascribed to circRNAs. They have been shown to act as miRNA sponges [336, 392, 394]. They can also regulate RNA binding protein (RBP)/RNA interactions [337-339] or code for proteins through cap-independent translation[342, 343]. CircRNAs are not limited to eukaryotic cells. Virus-encoded circRNAs were described in 2018 in the human herpesviruses Epstein-Barr virus (EBV) and Kaposi's sarcoma-associated herpesvirus (KSHV) [293, 295, 296, 350] and then subsequently in 2019 in the human papillomavirus (HPV) [301]. Attributable functions of most of these viral-encoded circRNAs are still undefined. However, HPV-encoded circE7 has been reported to be translated into an E7 oncoprotein [301] and the EBV-encoded circLMP2A may induce stemness in EBV-associated gastric cancer [395]. It has not been previously determined whether the human tumor virus MCV, in like fashion to these oncogenic DNA viruses, can encode circRNAs.

In this study, we identify circRNA backsplice junctions (BSJs) in polyomaviruses by RNase R^+ sequencing. For MCV, we characterized the expression of the most abundant MCV circRNA (designated circMCV-T) in MCV genome-transfected 293 cells as well as in MCC tumor-derived cell lines that have MCV monoclonally integrated into the host genome. These two conditions represent two different outcomes of MCV infection in human cells: that of active viral DNA replication/infection and that of replication-deficient, virus-induced transformation. CircMCV-T is consistently detectable in concert with ER linear mRNA transcripts as well as with MCV miRNA when expressed from a fully intact, nonintegrated viral genome. However, with replication-deficient integrated genomes, which occur in MCV-positive MCC cell lines, circMCV-T expression is inconsistently detectable; it is present in WaGa and CVG-1 but not in MKL-1, MKL-2, or MS-1 cells. Using MCV ER expression vectors as well as an efficient recombinase-induced DNA circularization minicircle (mc) technology to produce an MCV molecular clone, we find that the expression pattern of circMCV-T is inversely related to MCV-miR-M1 levels, and its exogenous expression stabilizes the mRNA of MCV early transcripts.

3.2 Materials and Methods

3.2.1 Tumor Samples and Cell Lines

Three freshly frozen tumor tissue specimens (MCC-1, MCC-2, and MCC-3) from patients with MCV-positive MCC were obtained from the Cooperative Human Tissue Network (CHTN). RatPyV2-positive parotid salivary gland tissue was obtained from a rat with X-linked severe combined immunodeficiency (X-SCID) [396] according to the University of Pittsburgh animal

care and use committee guidelines (protocol number 16048182). MCV-positive MCC tumor derived cell lines WaGa [397], MKL-1 [398], MKL-2 [399], CVG-1 [400], and MS-1 [401] were maintained in Roswell Park Memorial Institute (RPMI) 1640 media (Cellgro) supplemented with 10% Fetal Bovine Serum (FBS; VWR Seradigm). 293 cells (ATCC, CLR 1573) were maintained in Dulbecco's modified Eagle medium (DMEM; Corning) supplemented with 10 % FBS. Transfections were performed using either Lipofectamin 2000 (Life Technologies) or FuGENE 6 (Promega) following the manufacturers' protocols.

3.2.2 Plasmids and Constructs

Construction of the censuses pJMCV-HF plasmid for the production of circular, wild-type MCV-HF genome is detailed in reference [138]. Plasmids pCDNA3.1-ER-wt, pCDNA3.1-ER-hpko, pCDNA3.1GFP-MCV-miR-M1-wt 300, and pCDNA3.1GFP-MCV-miR-M1-hpko 300 have also been previously described [76]. For ectopic circRNA expression, we made use of the pcDNA3.1 (+) Laccase2 MCS exon vector, which models a *Drosophila* Laccase 2 miniature intron incorporating flanking inverted repeats and splice sites to facilitate circularization, The pcDNA3.1 (+) Laccase2 MCS exon vector was a gift from Jeremy Wilusz (cat. no. 69893, Addgene)[402]. The pLaccase-circMCV-T construct was generated by amplification of the circMCV-T coding region from the MCV-HF genome (GenBank: JF813003) using primers circ_MCV_LT_PacI_F and circ_MCV_LT_SacII_R with subsequent cloning into pcDNA3.1 (+) Laccase2 MCS exon vector using PacI and SacII restriction sites. pLaccase circMCV-T-hpko was similarly made using the same primer pairs but with pCMV-ER-hpko as template. The pMC.BESPX-MCV-HF plasmid used for production of MCV-HF mini-circle was synthesized (GeneScript). The MCV-HF genome was first synthesized with SmaI and BstEII restriction sites inserted in to the MCV-HF genomic

sequence between the ER and LR (between nt 3146 and 3147). The mini-circle plasmid vector pMC.BESPX (cat. no. MN100A-1, System Biosciences) [184, 403, 404], a kind gift from Mart Ustav (University of Tartu, Tartu, Estonia) and Alison McBride (National Institute of Allergy and Infectious Diseases, Bethesda, MD), was then cloned into the synthesized MCV-HF genome using restriction sites SmaI and BstEII to generate pMC.BESPX-MCV-HF. To make the hairpin mutant pMC.BESPX-MCV-HF-hpko mini-circle plasmid, a 1.6 kb fragment (nt 1152-2827) containing the MCV-miR-M1 region from the pMC.BESPX-MCV-HF plasmid was swapped with a similar fragment from the pCMV-ER-hpko plasmid using EcoRI and BamHI restriction sites. All constructs were validated and confirmed by sequencing. The list and source of all constructs are shown in **Table 10**.

Table 10 List and description of plasmids

Construct	Function**	Parental Vector	CM Plasmid #	CM Plasmid Name
MCV-HF	Amplifies MCV-HF genome in bacteria for MCV genome re-circularization by re-ligation after enzyme digestion that release MCV-HF genome from bacterial backbone.	pJ	3147	MCV-HF
MCV-wt-mc	Amplifies MCV-HF genome in bacteria for MCV genome re-circularization by recombination.	pMC.BESPX	4587	pMC.BESPX-MCV-HF
MCV-hpko-mc	Amplifies MCV-HF-hpko, which diminishes MCV miR-M1, in bacteria for MCV genome re-circularization by recombination.	pMC.BESPX	4610	pMC.BESPX-MCV-HF-hpko
pER-wt*	Expresses all early gene products, include MCV miR-M1.	pCDNA3.1	4427	pCMV2b-ER (pCMV-ER)

pER-hpko*	Expresses all early gene products, except for MCV miR-M1.	pCDNA3.1	4498	pCMV2b-ER-hpko
miR-M1 wt*	Expresses MCV miR-M1 and GFP.	pCDNA3.1	4429	pCDNA3.1GFP-MCV-miRM1 300 (pCMV-miRNA)
miR-M1 hpko*	Expresses GFP only, negative control for MCV miR-M1 expression.	pCDNA3.1	4499	pCDNA3.1-GFP-MCV-miRM1 hpko 300
circMCV-T wt	Expresses circMCV-T. This plasmid also has all elements to express MCV miR-M1 in the negative strand.	pcDNA3.1(+) Laccase2 MCS Addgene #69893	4448	pcDNA3.1(+)-Laccase-circMCV-LT
circMCV-T hpko	Expresses circMCV-T that cannot produce or interact with MCV miR-M1.	pcDNA3.1(+) Laccase2 MCS Addgene #69893	4643	pcDNA3.1(+)-Laccase-circMCV-LT-HPKO.SNPFIXED
EV	The backbone of pLaccase-circMCV-T-wt and pLaccase-circMCV-T-hpko. A negative control for circMCV-T expression.	pcDNA3.1(+) Laccase2 MCS Addgene #69893	4444	pcDNA3-Laccase-MCS (AA mut)

*These constructs were generated in the laboratories of Nicole Fischer and Adam Grundhoff, see reference [76].

**GFP, green fluorescent protein.

3.2.3 MCV-HF Recircularization and Mini-circle Production

Release and recircularization of the MCV-HF genome from pJ-MCV-HF was previously described [138]. To produce MCV genomes using the minicircle technology [184, 403, 404],

pMC.BESPX-MCV-HF or pMC.BESPX-MCV-HF-hpko plasmids were transformed into ZYCY10P3S2T (cat. no. MN900A-1; System Biosciences) competent cells, a kind gift from Mart Ustav (University of Tartu, Tartu, Estonia) and Alison McBride (National Institute of Allergy and Infectious Diseases, Bethesda, MD). Single colonies were inoculated into 5 ml of LB with 50 µg/ml kanamycin and grown at 37°C for approximately 8 hours. 100 µl of the starter culture was added into 400 ml of Terrific Broth and cultured overnight at 37°C. Next morning, when OD 600 reached between 4 and 5, an equal volume of induction mix (40 mM NaOH and 0.02% L-arabinose in a total of 400 ml LB) was added and culture was grown for additional 5 hours at 32°C to induce minicircle production. DNA was extracted by Maxi prep (Macherey-Nagel) and recombination efficiency was confirmed by restriction digestion. The resulting MCV-HF minicircle genome retains a 39 bp scar of bacterial plasmid sequence (GCCCCAACTGGGGTAACCTTTGGGCT CCCCGGGCGCGAC) between nt 3146 and 3147.

3.2.4 RNA Isolation and CircRNA Sequencing

Total RNA was extracted from 293 transfected with MCV-HF recircularized genome (5-day post-transfection), MCC cell lines (CVG-1 and MS-1), and RatPyV2-positive parotid gland tissue using TRIzol reagent (Invitrogen) and further processed with TURBO DNase kit (cat. no. AM190; Invitrogen) according to the manufacturer's instruction. RNA quality was confirmed by Agilent TapeStation (Children's Hospital of Pittsburgh of UPMC, sequencing core facility) and with an Agilent 2100 Bioanalyzer (CD Genomics). RNase R-treated samples were used for library preparation and subsequent circRNA sequencing using Illumina HiSeq platform in PE150 sequencing mode (CD Genomics).

3.2.5 Bioinformatic Analysis

Raw FastQ files were trimmed with Trim Galore (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) using the following parameters: q 25, e 0.1, and length 50, and the quality control was performed with FastQC. CircRNA prediction was conducted by CIRI2 (https://sourceforge.net/projects/ciri/files/CIRI2/) [405] with the default settings. RNA sequencing (RNA-seq) reads were aligned to MCV-HF (GenBank accession no. JF813003) or RatPyV2 (accession no. KX574453) reference genomes using BWA mapper. CLC genomics workbench (Qiagen) was used to align RNA-seq reads to MCV-HF (GenBank accession no. JF813003) and RatPyV2 (accession no. KX574453) reference genomes and to visualize additional annotation tracks using the following parameters: mismatch cost of 10, insertion cost of 3, deletion cost of 3, length fraction of 0.5, and similarity fraction of 0.9.

3.2.6 RNase R Treatment and RPAD

One microgram of DNase-digested total RNA was treated with 4 units (U) of Ribonuclease (RNase) R (cat. no. RNR07250; Lucigen) in 1× RNase R reaction buffer at 37 °C for 30 min in the presence of 8 U of RiboLock RNase inhibitor (cat. no. EO0381; Thermo scientific); for untreated samples, nuclease-free water was added to the reaction mixture instead of RNase R. RNase R was inactivated by incubating at 65 °C for 20 min. RPAD was performed as previously described [406]. Briefly, 2 µg of RNase R-treated RNA was ethanol precipitated in the presence of sodium acetate and 20 µg of glycogen as a carrier, followed by polyadenylation using the *E. coli* PolyA Polymerase I (E-PAP) kit (cat. no. AM1350; ThermoFisher) in the presence of RiboLock RNase inhibitor. Control RNA was processed in the same way without the E-PAP

buffer. Polyadenaylated RNA was then depleted using Purist MAG kit (cat. no. AM1922; ThermoFisher). cDNA was then synthesized as described below.

3.2.7 Semiquantitative and Quantitative RT-PCR

One microgram of DNase-treated total RNA in the presence or absence of RNase R treatment was reverse transcribed using random hexamers and the SuperScript IV first-strand synthesis system (cat. no. 18091; Invitrogen). Semiquantitative PCR for circRNA detection was performed using Q5 high-fidelity DNA polymerase (cat. no. M0491; New England BioLabs/NEB) and *Taq* DNA polymerase with ThermoPol buffer (cat. no. M0267; NEB) for linear RNA detection, according to the respective protocols from the manufacturers. See **Table 11** for the list of PCR primers used in this study.

Table 11 PCR primers

Primer Name	MCV Genome Position*	Sequence
circMCV-T_DP1F (DP1)	1514-1533	TACAAGCACTCCACCAAAGC
circMCV-T_DP1R1 (DP1R1)	926-946	TATTCGTATGCCTTCCCG
circMCV-T_DP1R2 (DP1R2)	1046-1067	GGACCCATACCCAGAGGAAGAG
circMCV-T_DP2F (DP2)	1409-1428	TGGTGAAGGAGGAGGATCTG
circMCV-T_DP2R (DP2)	1483-1503	GCTCTGCAAGCTCTGCTAGTT
circMCV-T_DP3F (DP3)	1079-1099	CCAGGCTTCAGACTCCCAGTC
circMCV-T_DP3R (DP3)	1057-1078	GACGCTGAGAAGGACCCATACC
circSMARCA5_F	N/A	CTCCAAGATGGGCGAAAGT
circSMARCA5_R	N/A	TTCTGATCCACAAGCCTCC

U6_F	N/A	GTGCTCGCTTCGGCAGCACA
U6_R	N/A	AAAATATGGAACGCTTCACGA
linear sT_F	397-422	AAGCTCAGAAGTGACTTCTCTATGTT
linear sT_R	613-632	TCTCCCCACGTCAGACAGTT
linear LT+57kT_R (CP-R)	3052-3080	TTATTGAGAAAAAGTACCAGAATCTTGGG
Linear β-Actin_F	N/A	CACACTGTGCCCATCTATGAGG
linear β-Actin_R	N/A	TCGAAGTCTAGGGCGACATAGC
Circ_MCV_LT_PacI F	N/A	CGCTTAATTAATTTCCCATCTAGGTTGAC
Circ_MCV_LT_SacII R	N/A	GATCCGCGGACTTACTGTTTTATTACT

* NA, not applicable.

To detect MCV-miR-M1 by quantitative stem-loop RT-PCR, 1 μg of total RNA without DNase and RNase R treatment was reverse transcribed using TaqMan MicroRNA reverse transcription kit (cat. no. 4366597; Thermo Fisher) with a MCV-miR-M1-specific stem-loop primer (SL_MCV-miR-M1), and a GAPDH-specific primer (GAPDH rev) for normalization as described before [76]. Quantitative PCR (qPCR) of MCV-miR-M1, GAPDH, MCV circMCV-T and linear transcripts was performed on cDNA input using TaqMan universal master mix II with UNG (cat. no. 44400; Applied Biosystems) in a QuantStudio 5 real-time PCR system (Thermo Fisher). See **Table 12** for list of qPCR primers and probes. Threshold cycle (C_T) values were used to calculate expression levels. C_T values for MCV-miR-M1 were normalized to GAPDH, other transcripts were normalized to RNase P and expression levels were calculated according to the $\Delta\Delta C_T$ method.

Table 12 qPCR primers and probes

Primer/Probe Name	**Sequence***

Taqman Probe circMCV-T	6FAM-AAAACAGTTGACGAGGCCCCTATATATGGG-QSY
circMCV-T_qPCR_F	ACTCCTGTTCCTACTGATTTTCC
circMCV-T_qPCR_R	TCCTCCTGATCTCCACCATTC
	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGAT
SL_MCV-miR-M1	ACGACTGTACC
Taqman probe MCV-miR-M1	ABY-CGCACTGGATACGACTGTACC-QSY
MCV-miR-M1 FW	GCATCTGGAAGAATTTCTA
Universal rev	GTGCAGGGTCCGAGGT
Taqman probe GAPDH	ABY-GTGGCGCTGAGTACGTCGTGGAGTC-QSY
GAPDH BSP FW	GGTCGGAGTCAACGGATTTG
GAPDH rev	ATGGTGGTGAAGACGCCAGT
GAPDH DNA fw	TGTGTCCCTCAATATGGTCCTGTC
Taqman Probe sT	ABY-AGCTGTAAGTTGTCTCGCCAGCATTGT-QSY
linear sT_qPCR_F	GCTAGATTTTGCAGAGGTCCT
linear sT_ qPCR_R	AAAACACTCTCCCCACGTCA
Taqman Probe T-Ag	ABY-TGGAATTGAACACCCTTTGGAGCA-QSY
linear T-Ag_ qPCR_F	TGCTCCTAATTGTTATGGCAAC
linear T-Ag_ qPCRR	AGCTTGTGGATATTTTGCTGGA
POP4 (RNase P)	VIC-MGB Hs00198357_m1, Catalog number: 4331182

*6FAM, 6-carboxyfluorescein; QSY, QuantStudioY; ABY, AppliedBiosystemsY; VIC-MGB, VIC-Minor Groove Binder.

3.2.8 BaseScope RNA *In Situ* Hybridization (ISH)

To prepare the cell pellet array, 293 cells were transfected with a pLaccase-empty vector, pLaccase-circMCV-T or pER. After 48 h, transfected 293 cells and the MCC cell line CVG-1 were harvested, and pellets were fixed in 10% neutral buffered formalin and processed with HistoGel (cat. no. #HG-4000-012; Thermo scientific) for routine histology. Five-micrometer sections of formalin-fixed paraffin-embedded cell pellets were then stained for the respective RNA species using BaseScope RNA *in situ* hybridization (ISH) with the BaseScope Duplex Detection Reagents kit (cat. no. 323810; Advanced Cell Diagnostics/ACD) using probes targeting circRNA BSJ or linear transcripts following a detailed protocol described previously [294]. In this study, probes with a red or blue signal were combined for hybridization (POLR2A probe is combined with *PPIB* probe; MCV circMCV-T probe is combined with linear T-Ag probe) and developed on the same slide sequentially. First, the red signal was amplified and developed as described previously [294]. Subsequently, slides were treated with AMP10 for 15 min at 40 °C, AMP11 for 30 min at room temperature and AMP12 for 15 min at room temperature with washing steps after each incubation period. Chromogenic signal was developed by incubating with BaseScope Duplex Green solution for 10 min at room temperature in the dark. Images were acquired using an Olympus AX70 microscope with a QImaging QIClick CCD camera and Q-Capture Pro 7 software. CircMCV-T is detected with a single 1zz probe, BA-V-MCV-gp3-circRNA-Junc-C2 (cat. no. 722851-C2; ACD), to its BSJ, while a single 2zz probe, BA-V-MCV-gp3-2zz-st (cat. no. 722841; ACD), was used to detect linear T-Ag RNA. Since circMCV-T probe is diluted with linear T-Ag probe (1:50 dilution), circMCV-T signal (red) and linear T-Ag signal (blue) are detected at the same time in two different channels. Human housekeeping genes *PPIB* and *POLR2A* were used as a positive control, while a

probe against the bacterial DapB gene was used as a negative control (BaseScope Duplex Control probe Hs-1zz; ACD cat. no. 700101).

3.2.9 Polysome Fractionation

293 cells transfected with 5µg of pLaccase-circMCV-T plasmid DNA in a 10-cm plate for 48 h or with 2 µg of MCV-HF-hpko mc in the presence or absence of 2 µg pLaccase-circMCV-T-hpko plasmid in a well of a 6-well plate for 4 days post-transfection. CVG-1 cells (2×10^7) were treated with cycloheximide (CHX; 100 µg/ml) and incubated for 15 min at 37 °C, 5% CO_2. Cells were then collected by centrifugation at 200 × g for 3 min, washed with cold PBS containing 100 µg/ml CHX and pelleted by centrifugation at 200 × g for 3 min. One milliliter lysis buffer (100 mM KCL, 5 mM $MgCl_2$, 10 mM HEPES [pH 7.2-7.4], 0.5% NP-40, 100 µg/ml CHX, 1.5 mM DTT, 200 U/ml Riboblock RNase inhibitor) was added to the cell pellet and lysed at 4 °C for 10 min. Cell lysates were centrifuged at maximum speed (17,000 × g) for 10 min at 4 °C, the supernatant was laid over a sucrose gradient (1.2 ml 50%, 0.9 ml 40%, 0.9 ml 30%, 0.8 ml 20% and 0.7 ml 10% sucrose, bottom to top) in ultracentrifuge tubes (cat. no. 344057; Beckman) and centrifuged at 35,000 RPM for 3 h at 4 °C using AH-650 ultracentrifuge rotor (Sorvall) and Discovery 90SE ultracentrifuge (Sorvall). Samples were then subjected to fractionation on ISCO density gradient fractionation system (Teledyne Technologies) using the following pump program: collection in micro tubes; last tube, 15 (15 fractions); fraction by drops, 20 drops (300 µl); flow delay, 0 seconds; event, 25% pump output; event time, 0 seconds; and chart speed, 150 cm/h. Eight hundred microliters of TRizol LS reagent (cat. no. 10296010; Invitrogen) was then added to each fraction and RNA was extracted according to the manufacturer's protocol. A random hexamer was

used for cDNA synthesis, and qPCR was performed for circMCV-T and GAPDH using the Taqman chemistry as described above.

3.2.10 Immunoblotting

Transfected 293 cells were collected in 1% SDS buffer (1% SDS, 10 mM Tris-HCl [pH 8.0], 1 mM EDTA [pH 8.0]) and sonicated on ice four times for 5 s each time, and 100 µg of protein was used for SDS-PAGE. Protein was transferred to nitrocellulose membranes and incubated with 1:500 dilution of CM2B4 (MCV LT and 57kT antibody), a 1:10 dilution of CM7B1 (MCV ALTO antibody), a 1:500 dilution of CM9B2 (MCV VP1 antibody) and a 1:1,000 dilution of mouse α-tubulin antibody (12G10; DSHB). A 1:10,000 dilution of IRD800 anti-mouse antibody (cat. no. 926-32210; Rockland) was used as fluorescent secondary antibody. Images were acquired on ChemiDoc Imaging System (BioRad).

3.2.11 MCV Replication Assay

293 cells transfected with either a wt or hpko mutant MCV-HF minicircle genome were collected 2- and 4-day post-transfection using DNA lysis buffer (0.1 M NaCl, 10 mM Tris·HCl [pH 8.0], 25 mM EDTA [pH 8.0], 0.5% SDS, 200 µg/ml Proteinase K) and incubated at 45 °C overnight. DNA was extracted using standard phenol:chloroform (cat. no. P3803; Sigma) and precipitated with sodium acetate and ethanol. DNA was treated with BamHI and DpnI (NEB) overnight. Five nanograms of treated DNA was used for qPCR. Viral genomes were quantified with primers (Genomic VP1_F and Genomic VP1_R), spanning three DpnI restriction sites and normalized to GAPDH (primers: GAPDH DNA fw and GAPDH-rev).

3.3 Results

3.3.1 Identification of Merkel Cell Polyomavirus Encoded CircRNAs by RNase R^+ RNA Sequencing.

To investigate whether polyomaviruses, specifically MCV, can encode circRNAs, RNA sequencing was performed on RNA isolated from two models of MCV infection: 293 cells transfected with an *in vitro*-recircularized wild-type (wt) MCV-HF genome (GenBank accession no. JF813003) [138] and CVG-1 or MS-1 MCC derived cell lines with an integrated viral genome and ER polymorphisms that abolish genome replication. We treated extracted RNA with RNase R to deplete linear transcripts and to enrich circular forms of RNA. Since our goal for sequencing was to identify circular RNA molecules, we did not process RNase R untreated samples, nor did we perform sequencing repeats. MCV circRNA back splice junction (BSJ) candidates were identified by the CIRI2 circRNA prediction algorithm [405] using the consensus genome MCV-HF (GenBank accession no. JF813003) as a reference. The structure and arrangement of MCV early T antigens (T-Ags) and late transcripts as well as miR-M1 is shown in **Figure 15A**. Total RNase R-resistant reads and reads mapping to the MCV genome for each cellular model of MCV infection are shown in **Table 13**. In comparison to the replicating model of MCV-HF-transfected 293 cells, those of the MCC cell lines MS-1 and CVG-1 displayed at least 2 orders of magnitude-lower numbers of RNase R-resistant viral reads despite the total numbers of reads being comparable between the two models. This is in agreement with a previous report where the numbers of viral sequence reads from poly(A) sequencing of MCC-derived cell lines MKL-1 and WaGa were significantly lower than for PFSK-1 cells transfected with MCVsyn, a circularized, wild-type genome identical to MCV-HF [76]. A total of seven distinct MCV BSJs were detected

from MCV-HF-transfected 293 cells by RNase R^+ sequencing (**Table 14**), while there were no BSJ reads detected from the MCC derived cell lines CVG-1 and MS-1. This is consistent with the low number of viral transcripts produced in these cell lines, which have contracted viral transcriptional programs compared to those of replication-competent viral genomes in the 293 transfection model. The contribution of total combined positive MCV circRNA BSJ reads to the total RNase R^+ sequenced reads from the MCV early region (ER) is very low (< 0.019 reads per million [RPM]) (**Table 14**) compared to the circRNA reads from the EBV (~37 to 440 RPM for circBARTs) and KSHV (~200 RPM for circ-vIRF4) genomes [293, 295, 296]. We observed depletion of sequence coverage corresponding to the 5'-most part of the putative circMCV-T coding region. The sequencing coverage plot also displayed RNase R-resistant peaks corresponding to regions in the first exon of LT, sT, as well as parts of the late genes (**Figure 15B**, bottom).

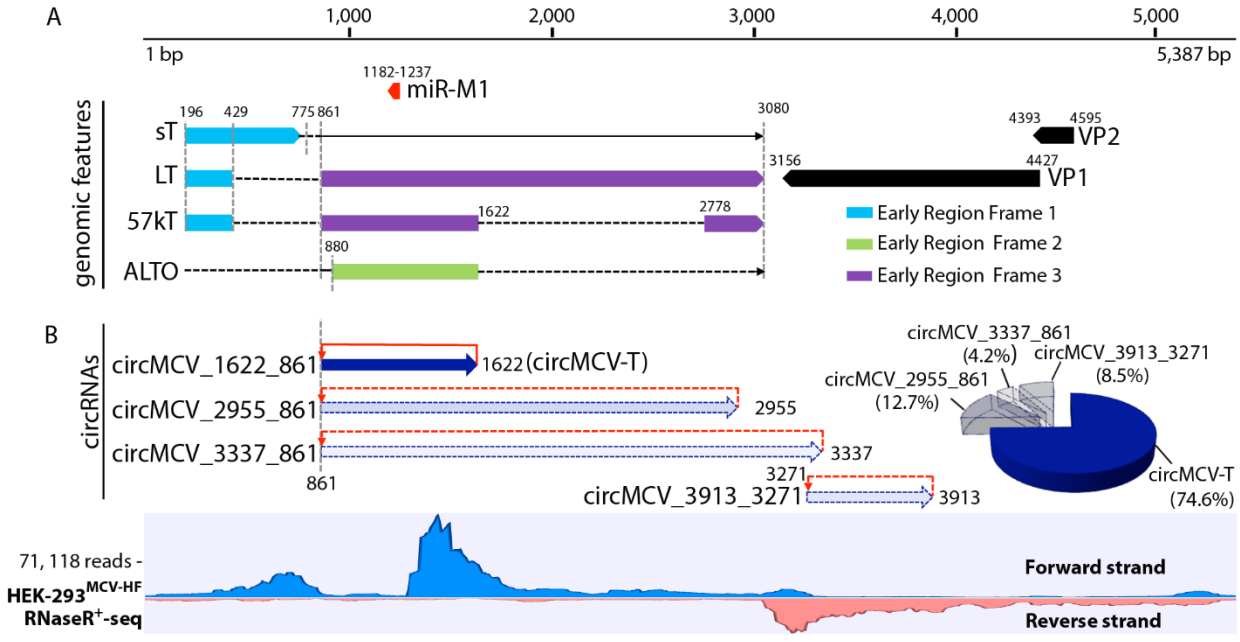


Figure 15 Identification of MCV-encoded circRNAs by RNase R+ RNA sequencing.

(A) Schematic representation of the MCV genome organization. Structures of known MCV early transcripts are shown. Positions of splice donor or acceptor sites are shown by vertical dashed lines and/or numbers. Known mRNAs are represented with solid arrows, while protein products are shown as boxes. Dashed lines show sequences that are spliced out of the indicated RNA transcript. (B, bottom) RNase R^+ RNA sequence coverage of MCV-HF in transfected 293 cells after 5 days of transfection. Sequence reads from the forward strand (blue) and reverse strand (red) are shown. (Top) Schematic representation of the four putative MCV-encoded circRNA backsplice junctions (BSJs) mapped to the MCV-HF genome. Red arrows show backsplicing directions. The percentage of each putative circRNA BSJ count from the ER is shown in a pie chart (right). The sequencing result is from one experiment.

Table 13 Summary of RNase R^+ RNAseq reads obtained from MCC cell lines, MCV-HF transfected 293 cells, and RatPyV2 infected rat parotid gland.

	# total reads	# unmapped	# mapped	%mapped reads
CVG	85,674,624	85,673,165	1,459	0.0017
MS1	80,138,038	80,137,423	615	0.0008
293-MCV-HF	77,848,404	77,462,185	386,219	0.4961
RatPyV2-parotid	82,621,490	81,875,309	746,181	0.9031

Table 14 Sequenced BSJs from MCV-HF

MCV BSJ Start (nucleotide number)	MCV BSJ End (nucleotide number)	Strand	BSJ reads	RPM
861	1622	+	53	0.014
861	2955	+	9	0.002
861	3337	+	3	0.001
3271	3913	+	6	0.002
		Subtotal Forward	71	0.019
1142	5308	-	92	0.024
1142	5119	-	25	0.006
1142	4642	-	4	0.001
		Subtotal Reverse	121	0.031

MCV-HF= re-circularized genome transfected in to 293 cells, GenBank ID JF813003)

Out of the seven different BSJs found in the MCV-HF transfected 293 cells, four were detected from the viral ER (**Figure 15B and Table 14**). Among these four putative circRNA BSJs, circMCV_1622-861 or circular MCV-T (named circMCV-T) was the most abundant, comprising

74.6% of the total 71 BSJs detected from the forward strand. The circMCV-T BSJ is generated through backsplicing of the 3' donor at nt 1622 [66, 67, 76] to an upstream 5' acceptor at nt 861, circularizing the entire exon II of T-Ag transcripts to produce a 762-nt circRNA (**Figure 15A and B**). Although the splicing coordinates for circMCV-T corresponds to those of exon II generated by canonical forward splicing to produce the 57kT transcript, in fact, all known T-Ag transcripts (LT, sT, 57kT, and ALTO) can be affected by the usage of these splice sites because the annotated splice acceptor and donor sites are required for the splicing of all naturally occurring alternative transcripts (**Figure 15A**). Two other low-abundance putative circRNA BSJs from the ER, circMCV_2955-861 (12.7% of 71 BSJs) and circMCV_3337-861 (4.2% of 71 BSJs), similarly use the same canonical splice acceptor at nt 861 but differ in their splice donor sites located at nt 2955 and 3337, respectively. These splice donor sites have not been previously annotated to be recruited in forward splicing events (**Figure 15B and Table 13**); however, these sites contain a splicing donor sequence (/GTAA) similar to that of circMCV-T. A fourth low-abundance BSJ (circMCV_3913-3271 [8.5%]) is also detected from the forward strand. Unlike the three circRNA that are same-stranded in relation to T-Ag coding transcripts and share the same back-splice acceptor coordinate, this fourth BSJ has unique splicing coordinates and is complimentary to the VP1 coding transcripts (**Figure 15B and Table 13**).

Three of the seven MCV BSJ reads were detected from the late region (LR) (**Table 13**). These reads all contain a putative common splice donor at nt 1142, which has previously been identified as a low-frequency splice site, [76] coupled with a acceptor site at nt 4642, used by the VP1 transcript, nt 5119, used by the VP2 transcript, or nt 5308, another low-frequency acceptor site [76]. However, these BSJ-simulating reads may not actually represent an authentic circRNA-forming backsplicing event; instead, these reads are likely to arise during leader-to-leader forward

splicing of multigenomic LR RNA precursors formed by the action of RNA polymerase II (Pol II) multiply circuiting the viral genome. This process is believed to account for the accumulation of late viral mRNA transcripts during late stages of viral replication [119, 120, 407].

3.3.2 MCV CircRNA Validation and Characterization

The putative circMCV-T is the most abundant BSJ comprising 74.6% of the total 71 BSJs reads from the early region of the MCV genome (**Figure 15B and Table 14**). Therefore, further characterization and analysis were directed at circMCV-T. Divergent primer (DP) pairs, amplifying either fragments or the full circle spanning the BSJ, were designed for reverse transcription (RT)-PCR analysis (**Figure 16A**). Exact nucleotide positions of primers are provided in **Table 11**. Because of the overlapping nature of the three putative circRNAs from the MCV ER, all DP pairs have the potential to amplify each of these circRNAs, although differing in amplification product sizes. Primer pairs DP1 (DP1.F coupled either with reverse primer DP1.R1 or DP1.R2) are calculated to produce bands of 195 or 316 bp, respectively, from circMCV-T. The DP2 primer pair will produce a 707-bp product, while the back-to-back divergent primer pair DP3 will produce a 762-bp product representing the entire circMCV-T (**Figure 16B**). RT-PCR using primer pairs DP2 and DP1.R2 detected both the large and small fragments (at the predicted sizes) of circMCV-T in WaGa [397], CVG-1 [400], and 293 cells transfected with MCV-HF in non-RNase R-treated samples (**Figure 16C, red outline**), while no products of predicted sizes were detected from MKL-1, MKL-2 and MS-1 cells (**Figure 16C**). An additional PCR product with a larger size was also detected in WaGa cells for both the DP2 and DP1.R2 divergent primer pairs. Sequencing identified them to represent circMCV_2955_861 with the 57kT intron spliced out. As seen previously, linear LT and 57kT transcript levels vary widely between cell lines infected with

different strains of MCV; however, linear viral sT transcript levels are relatively constant. Therefore, to confirm RNA quality and RNase R treatment efficiency, sT transcript levels along with cellular β-actin transcript levels were used as controls. The absence of detectable circMCV-T products from MKL-1 cells is consistent with the absence of the splice donor at nt 1622 due to a naturally occurring 46-nucleotide deletion (nt 1612 to 1657) in this particular integrated viral genome strain [38, 66]. The MKL-2 viral strain contains a 2 bp deletion at nt 3082-3083, while the MS-1 viral strain has a 40-bp deletion at nt 1912-1951 [38, 66]. Genomic polymorphisms in these two virus strains leave the circMCV-T splice donor (nt 1622) and the splice acceptor (nt 861) intact and, therefore, should not preclude formation of circMCV-T. However, other as-yet-undefined sequence-related features might affect splicing in these genomes. Remarkably, the presence of circMCV-T expression in these MCC derived cell lines mirrors that of 57kT transcript expression (**Figure 16C**). PCR using an intron-spanning convergent primer pair (DP2-F and CP-R) results in two PCR products corresponding to LT (1597 bp) and 57kT (441 pb). The larger fragment representing the LT transcript is observed only in the positive-control MCV-HF-transfected 293 cells, while the smaller fragment corresponding to the 57kT transcript can be detected only in WaGa and CVG-1 cells, but not in MKL-1, MKL-2 or MS-1 cells. As with circMCV-T, the deletions in MKL-2 and MS-1 strains also should not affect 57kT splicing.

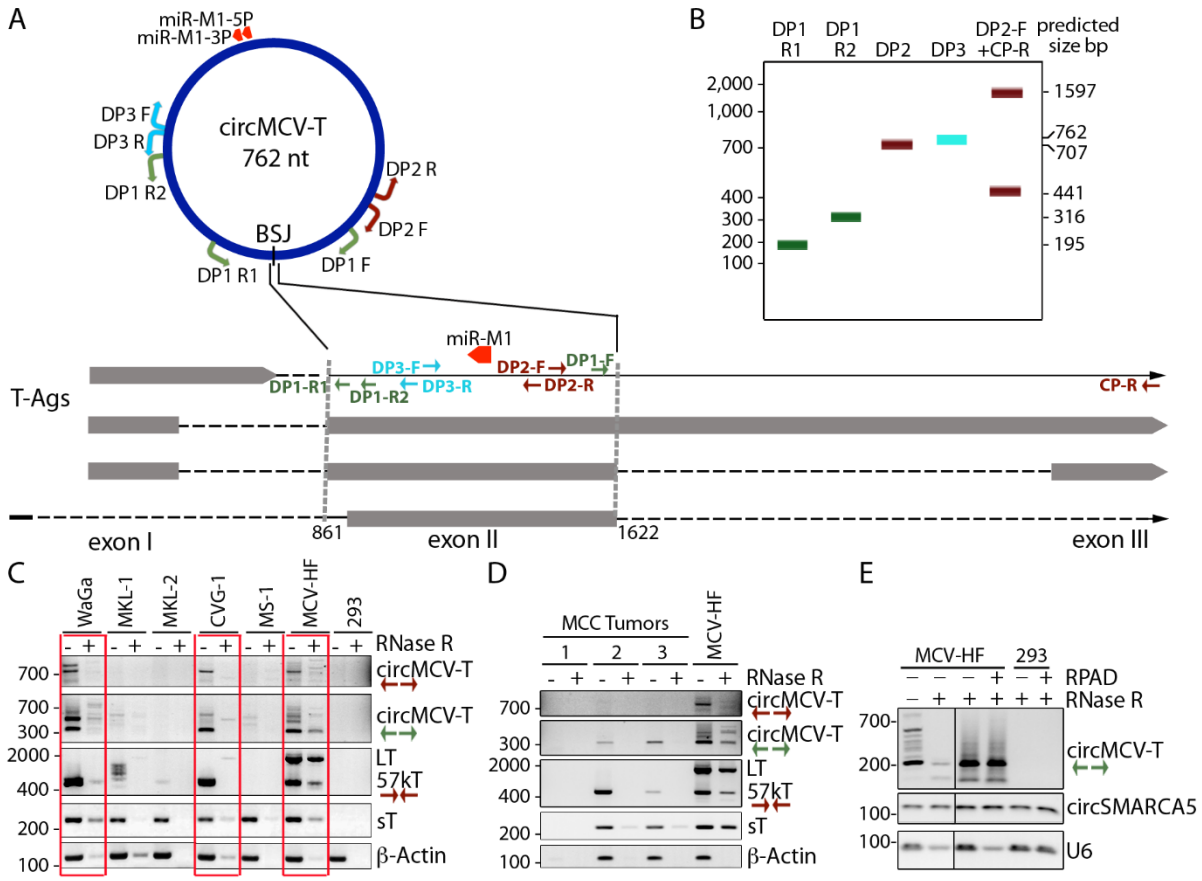


Figure 16 Validation and characterization of circMCV-T in MCC-derived cell lines and primary tumor tissues.

(A) Schematic representation of MCV T-Ag transcripts and circMCV-T. Divergent primer pairs used for circMCV-T RT-PCR are shown with color codes. (B) Schematic representation of expected PCR products using each primer pair in panel A. (C and D) Expression profiles of circMCV-T in MCC cell lines (Waga, MKL-1, MKL-2, CVG-1, MS-1) and MCV genome (MCV-HF)-transfected 293 cells (C) and in patient-derived tumor samples (MCC 1 to 3) (D) in the presence and absence of RNase R treatment are detected by RT-PCR using the indicated divergent primer pairs (DP) for circMCV-T and conventional convergent primer pairs for linear transcripts (sT, 57kT, LT, and β-actin). Untransfected 293 cells were used as a negative control, and linear RNAs were used as an RNase R digestion control. See Table 11 for the location of each primer on the MCV genome. (E) RPAD was performed on RNase R-treated RNA from 293 cells transfected with the MCV-HF genome for 5 days. The cellular circRNA SMARCA5 is used as a control that

should not be depleted after RPAD treatment, while U6 was used as linear RNA control that should be depleted by RPAD. Each result represents a single experiment. Numbers at the left of the blots are molecular sizes in base pairs.

BSJ-containing fragments of circMCV-T were also detected from two out of three MCC patient-derived tumor samples tested (MCC-2 and MCC-3). In the MCC-1 sample, from which no circMCV-T was found, linear viral (LT, 57kT, sT) and cellular control (β-actin) were also not detectable, suggesting that the quality of the extracted RNA was inadequate (**Figure 16D**). The DP1.R2 primer pair detected the predicted 316-bp product in tumor samples with intact RNA as well as MCV-HF infected 293 cells. The DP2 primer pair, which can amplify almost the entire circMCV-T, detected the predicted 707-bp band only from the MCV-HF infected 293 cells but not from MCC tumor samples. The PCR products from **Figures 16C and D** were confirmed to contain circMCV-T BSJ by sequencing. The three low-abundance BSJs from the RNase R^+ sequencing are not reliably detectable in naturally infected tissues or in cell lines by RT-PCR. Their functional significance is difficult to assess.

Unlike with circRNAs, whose nature is reportedly exonuclease resistant [333, 347], the majority of circMCV-T is RNase R sensitive in a fashion similar to that of viral and cellular linear transcripts (**Figure 16C and D**). Therefore, to further confirm that completely cyclized circMCV-T forms occur, we applied RNase R treatment, polyadenylation, and poly $(A)^+$ RNA depletion (RPAD) [406] on RNA from MCV-HF-transfected cells. The RT-PCR results from RPAD-processed and -unprocessed samples (**Figure 16E**) shows no further depletion of circMCV-T from RNase R-treated samples after RPAD (**Figure 16E, top, 3rd and 4th lanes**), indicating the presence of true circular forms of circMCV-T. Cellular circSMARCA5 was assayed as a control that is not depleted by RNase R and RPAD, while the linear RNA U6 was used as a control that can be depleted by RPAD.

3.3.3 CircMCV-T Detection *In Situ*

To investigate the presence of circMCV-T *in situ*, BaseScope RNA hybridization was applied to cell pellet arrays comprised of CVG-1 and 293 cells transfected with either the MCV ER expression construct, a pLaccase-circMCV-T construct, or an empty pLaccase vector control. BSJ-spanning 1z probe (red) to circMCV-T was detected in all samples except empty-vector-transfected negative-control cells (**Figure 17A**). The 2zz probe (blue) to linear T-Ag at a position outside the circRNA-coding region was detected in MCV ER-transfected 293 cells and in the MCC-derived CVG-1 cell line, but not in pLaccase-circMCV-T- or pLaccase empty-control-transfected cells. As expected, the highest level of circMCV-T expression was found in pLaccase-circMCV-T- and MCV ER control-transfected cells, where the red signal manifested as extensive clumps obliterating cytologic details rather than as individual dots. Signals detected as clumps occur under conditions of high levels of ectopic RNA expression and cannot be quantified. Furthermore, the BaseScope *in situ* hybridization protocol is much longer and harsher on tissue section morphology; therefore, loss of crispness (e.g., nuclear details and hematoxylin counterstaining, for example) is not unexpected and cannot be construed to be an effect of circMCV-T expression. The signal and number of cells positive are almost entirely dependent on transfection efficiency in the controls. However, in CVG-1 cells, circMCV-T expression is observed as red dots of low occurrence 4.7% of 402 cells counted versus 85.3 % for linear T-Ag (**Table 15**). Cellular linear transcripts PPIB (blue) and POLR2A (red) were used as controls for RNA quality. A probe against the bacterial DapB gene is used as a negative control.

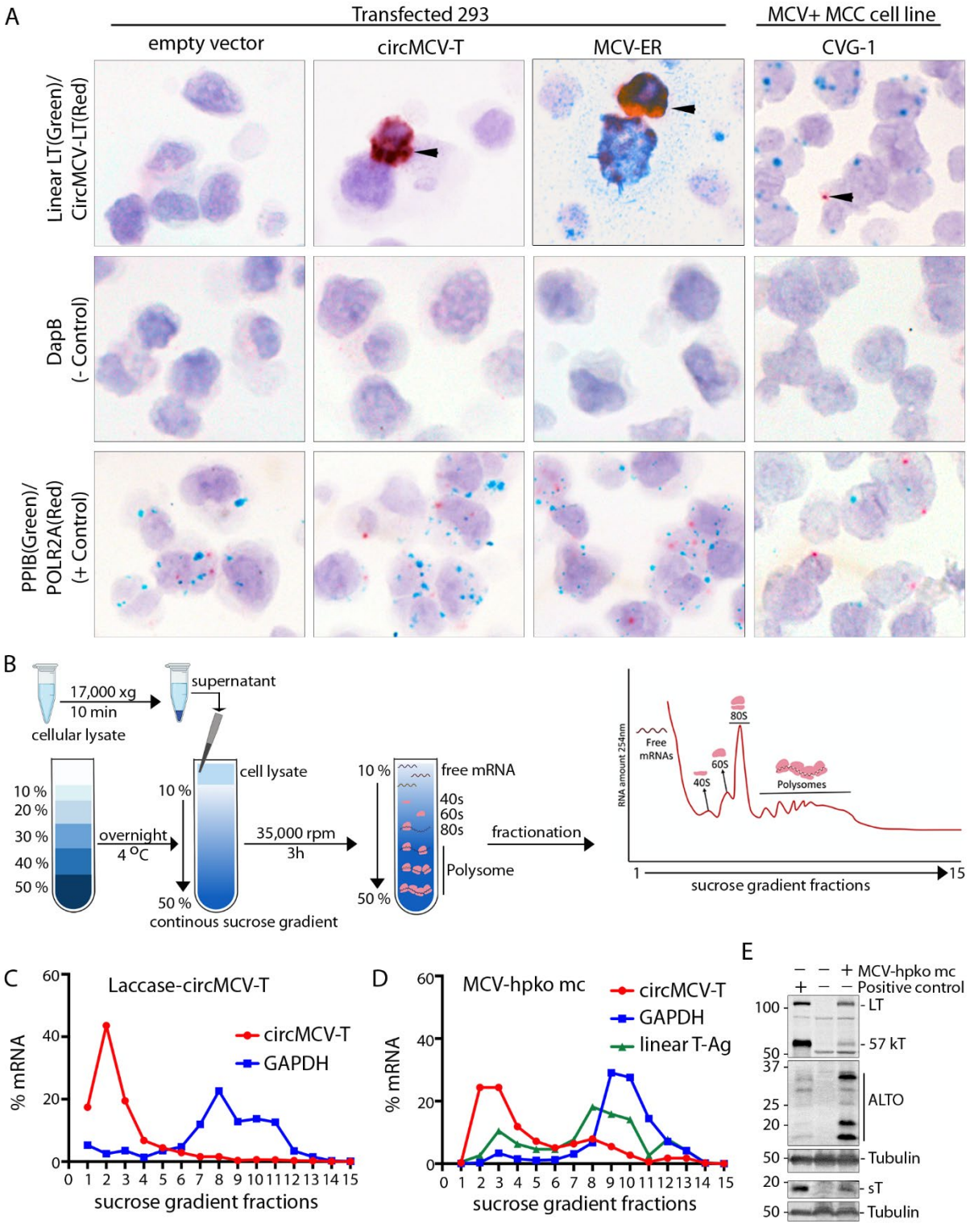


Figure 17 *In situ* detection and polysome fractionation of circMCV-T.

(A) Detection of MCV circMCV-T BSJ by BaseScope *in situ* hybridization. Representative images from 293 cells transfected with an expression vector for circMCV-T or MCV ER (early region) are shown. CVG-1, an MCV$^+$ MCC cell line with circMCV-T detectable by RT-PCR, shows an abundance of linear T-Ag transcripts (blue) compared with circMCV-T transcripts (red). 293 cells transfected with an empty vector were used as a negative control. The red signal represents the detection of circMCV-T, the cellular control POLR2A, or the bacterial negative-control gene DapB, while the blue signal represents staining for linear T-Ag RNA or the cellular control gene PPIB. Images were originally acquired at a ×40 magnification. (B) Polysome fractionation assay workflow created using BioRender.com. (C and D) RT-qPCR of circMCV-T (red), linear T-Ag (green), and GAPDH (blue) transcripts performed on polysome fractions from (C) pLaccase-circMCV-T-transfected (no linear T-Ag is produced from this expression construct and was therefore not assessed) and (D) MCV-HF-hpko mc-transfected 293 cells. (E) Western blot of MCV-encoded proteins from the MCV early (LT, 57KT, ALTO, and sT) regions from MCV-HF-mc-transfected 293 cells on which polysome fractionation is depicted in panel C with α-tubulin used as a protein internal control. Experiments whose results are represented in panels C, D, and E were each performed at least two times.

Table 15 CircMCV-T BaseScope IH detection in CVG-1 cells

	Total	Linear T-Ag Positive cells	circMCV-T Positive cells
Number of cells	402	343	19
%	100	85.3	4.73

3.3.4 CircMCV-T Is Unlikely to Be Translated into A Protein Product.

Some circRNAs have been shown to code for proteins through internal ribosome entry site (IRES) or N^6-Methyladenosine (m^6A)-mediated, 5'-cap-independent translation initiation [301, 342, 343]. To test whether circMCV-T has the potential to code for a protein product, we performed polysome fractionation (**Figure 17B**) on 293 cells transfected with pLaccase-circMCV-T construct or MCV-HF mc molecular clone with a mutation to knock out miR-M1 expression (MCV-hpko mc). The miR-M1 hpko mutation has been previously shown to increase viral

replication [76]. Polysome fractionation showed that while the positive control linear RNA GAPDH (glyceraldehyde-3-phosphate dehydrogenase) is found associated with polysomes in fractions 7 through 12, circMCV-T is detected only in the earlier fractions that do not contain polysomes in cells transfected with pLaccase-circMCV-T (**Figure 17C**). In cells transfected with MCV-hpko mc, the slight increase in detection around fraction 8 is of undetermined significance (**Figure 17D**). A Western blot control for viral protein expression from MCV-hpko mc is shown in **Figure 17E**.

We also assayed for the protein-coding capacity of circMCV-T by Western blot analysis using two antibodies with epitope targets covering two frames of possible protein translation in exon II: the CM2B4 antibody, traditionally used to detect LT and 57kT (ER frame 3), and the CM7B1 antibody, which detects ALTO (ER frame 2) (**Figure 18A**). The third frame (ER frame 1) contains a single start codon and multiple stop codons that can theoretically code for a very small protein product of ~9 kDa. With this caveat and within the limitations of immunoblotting sensitivity, and as well as with the specific binding efficacies of these antibodies, we were unable to detect a protein product from 293 cells transfected with pLaccase-circMCV-T (**Figure 18B**).

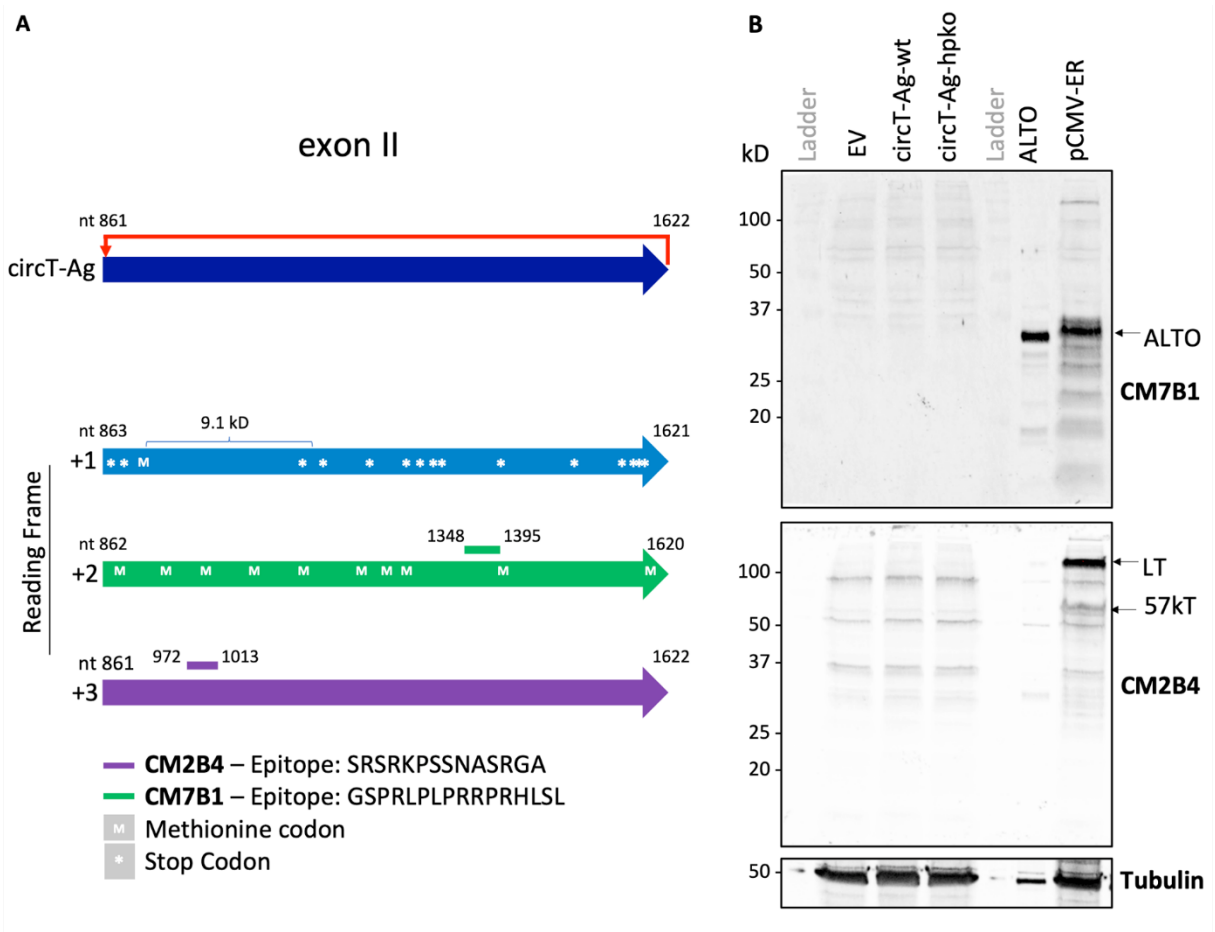


Figure 18 Assay for the circMCV-T-encoded protein product.

(A) Schematic showing all three potential open reading frames present within circMCV-T aligned to the consensus MCV-HF genome (GenBank accession no. JF813003). Only the +1 and +2 reading frames encode methionine amino acids, and only the +1 reading frame encodes stop codons. The only peptide potentially encoded by the +1 frame has a size of 9.1 kDa. (B) 293 cells transfected with plasmids expressing circMCV-T and blotted with antibodies to epitopes within the +2 and +3 reading frames show no protein translation to the limit of detection. The same membrane was incubated with both antibodies, with CM7B1 blotting preceding CM2B4 blotting.

3.3.5 The Absence of miR-M1 Expression Increases MCV CircMCV-T Level.

The MCV ER contains the coding sequence for MCV miR-M1 complementary to the T-Ag coding region of the viral genome. The resulting mature miRNA seed sequences miR-M1-5P and -3P can bind the reverse strand with 100% complementarity (**Figure 15A and 19A**). This complete complementarity results in the degradation of linear T antigen transcripts, leading to suppression of viral replication [76]. Because circMCV-T also has the MCV miRNA binding sites, we hypothesize that the RNase R sensitivity of circMCV-T is likely to be due to its linearization by miR-M1, leading to the apparent low number of BSJ reads detected from the MCV genome (**Table 14**). To test this, we compared the levels of circMCV-T as well as of linear LT and sT transcripts from the MCV ER construct containing the wt sequence for miR-M1 or with mutations in the pre-miR-M1 seed sequence resulting in a hairpin knockout (hpko) that abolishes mature MCV miR-M1 expression [76]. As previously described [76], we saw an increase in linear sT and LT transcripts in the hpko mutant MCV ER compared to the wt construct; circMCV-T levels were also detected at a higher level in the hpko mutant MCV ER, as detected by RT-qPCR using a BSJ-spanning Taqman probe and a pair of divergent primers on either side of the BSJ (**Figure 19A and B**). In contrast, miR-M1 is undetectable in the hpko mutant, unlike in the wt construct, as detected by stem-loop RT-PCR analysis. As with previous reports [76], protein expression levels of LT, 57kT, and ALTO are also increased as a result of the miR-M1 hpko mutation (**Figure 19C**). To better understand the functional interaction of circMCV-T with miR-M1, we cotransfected pLaccase-circMCV-T with a wt or hpko mutant miR-M1 expression vector and analyzed circMCV-T levels by RT-PCR and qPCR. Note that the wt pLaccase-circMCV-T expression construct can potentially express miR-M1 from the opposite-strand DNA sequence due to the presence of a miR-M1 promoter described previously [76]. Unlike with cells not transfected with

miR-M1, cells cotransfected with the wt miR-M1 expression vector exhibited decreased circMCV-T levels after RNase R treatment, as detected by the DP1 primer pair (**Figure 19D, top**). The presence of miR-M1 completely abolished the detection of the full-length circMCV-T using a back-to-back primer pair (DP3) that can amplify the entire 762-bp circle. In the absence of miR-M1 expression, circMCV-T is enriched after RNase R treatment (**Figure 19D, bottom**). Additionally, expression of miR-M1 from a wt expression vector depleted circMCV-T levels compared to those after cotransfection of an hpko mutant miR-M1 construct (**Figure 19E**). Together, these results support the notion that miR-M1 mediates linearization of circMCV-T, resulting in its RNase R sensitivity.

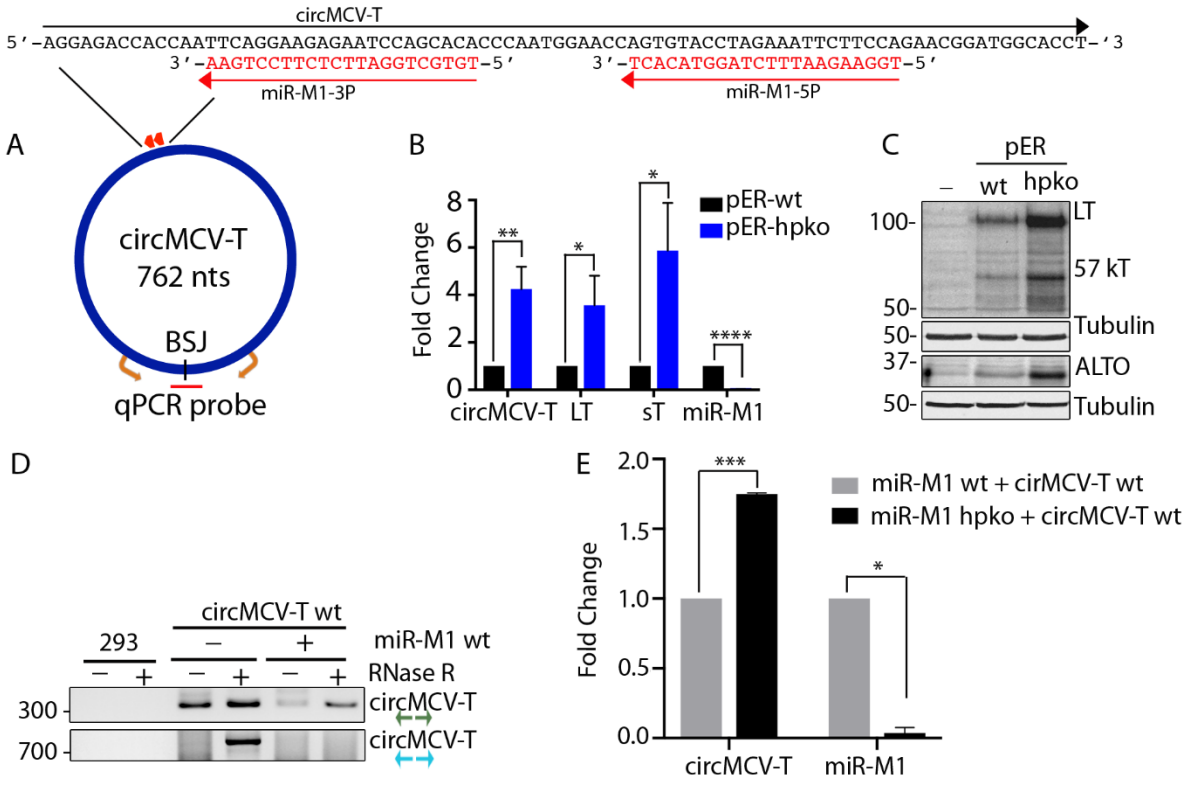


Figure 19 The presence of MCV miR-M1 decreases circMCV-T levels.

(A) Schematic representation of MCV circMCV-T and miR-M1 features. Mature MCV-miR-M1-5P and -3P binding sites are shown as red arrows, and sequence features at the miRNA binding site are represented at

the top. Divergent primers (orange) and a BSJ spanning TaqMan probe (red) are used for circMCV-T qPCR amplification. (B) Expression of circMCV-T and T-Ag RNA from MCV early region expression constructs containing either a wt or hpko mutation for MCV miR-M1. (C) Expression of T antigen proteins (LT, 57kT, and ALTO) from panel B. (D) RT-PCR detection of circMCV-T expressed from a pLaccase-circMCV-T construct in the presence and absence of MCV-miR-M1 using the indicated divergent primer pairs 48 h after transfection of 293 cells. (E) RT-qPCR quantification of circMCV-T levels from a pLaccase-circMCV-T construct in the presence of either wt or hpko mutant MCV miR-M1 48 h post-transfection of 293 cells. Results in panels C and D are representative of three independent experiments. C_T values were normalized to RNase P (B) or GAPDH (E) C_T values, and samples were normalized to pER-wt (B)- or wt miR-M1 (E)-transfected cells to calculate fold change. Error bars in panels B and E represent means ± standard deviations (SD) from three independent experiments. Statistical analysis was performed on $\Delta\Delta C_T$ values using an unpaired t test. *, P < 0.05; **, P < 0.01; ***, P < 0.001; ****, P < 0.0001.

3.3.6 Role of CircMCV-T in Early Gene Expression and Viral Replication

We have shown that, as with the linear T-Ag transcripts, circMCV-T can be targeted by miR-M1. This predicts that circMCV-T plays a role in aiding viral replication by competing for miR-M1 binding with T-Ag linear transcripts early in the viral life cycle. To test this hypothesis, the wt MCV ER or that harboring the hpko mutation was cotransfected with wt or hpko mutant pLaccase-circMCV-T or an empty vector control, followed by RT-qPCR assessment of circMCV-T, miR-M1-5P, and linear transcripts. In addition to circMCV-T, the wt pLaccase-circMCV-T expression construct makes miR-M1 from the antisense strand (**Figure 20A and B**). **Figures 20C and D** show that consistent with the previous experiment (**Figure 19B**), circMCV-T is expressed from both the wt and hpko mutant MCV ERs. Coexpression of the wt circMCV-T vector further increased circMCV-T levels both in the wt and hpko MCV ER constructs, while an hpko mutant pLaccase-circMCV-T further amplified the increase in circMCV-T expression levels. This

increase in circMCV-T expression was accompanied by an increase in the level of sT and LT transcripts in both the wt and hpko mutant MCV ER constructs. As expected, miR-M1 levels were diminished in cells transfected with hpko mutant constructs. While the increase in miR-M1 expression from the MCV ER construct cotransfected with the wt circMCV-T is expected due to additive miRNA expression from a previously described promoter [76] present in the pLaccase circMCV-T construct (**Figure 20B**), we have also observed a marked increase in miR-M1 expression from MCV ER upon cotransfection of hpko mutant pLaccase-circMCV-T. In addition, there was also an increase in miR-M1 expression from the wt pLaccase-circMCV-T construct when cotransfected with the hpko mutant MCV ER construct (**Figure 20C**). This may suggest regulation of the miR-M1 promoter by the increased level of T-Ag proteins from the hpko mutant MCV ER. To investigate whether the increase in linear T-Ag transcripts correlates with T-Ag protein expression and viral replication, the MCV-HF mc genome (MCV mc) with either the wt or hpko mutant miR-M1 sequence was cotransfected with wt or hpko mutant pLaccase-circMCV-T, and viral replication was assessed by Western blotting of MCV proteins and qPCR analysis of replicated viral genomes. **Figure 20E** shows that the overexpression of wt or hpko mutant circMCV-T does not result in a marked change in any of the MCV-encoded proteins from either the wt or the hpko mutant MCV mc genome. Further analysis of the level of replicated (DpnI-resistant) MCV genomes in these cells revealed only a slight increase at day 2 and day 4 post-transfection for hpko and wt MCV mc-transfected cells, respectively (**Figure 20F**). However, the hpko mutation in the viral genome increased both viral protein expression and genome replication compared to levels in the wt MCV-HF mc genome (**Figure 20E and F**).

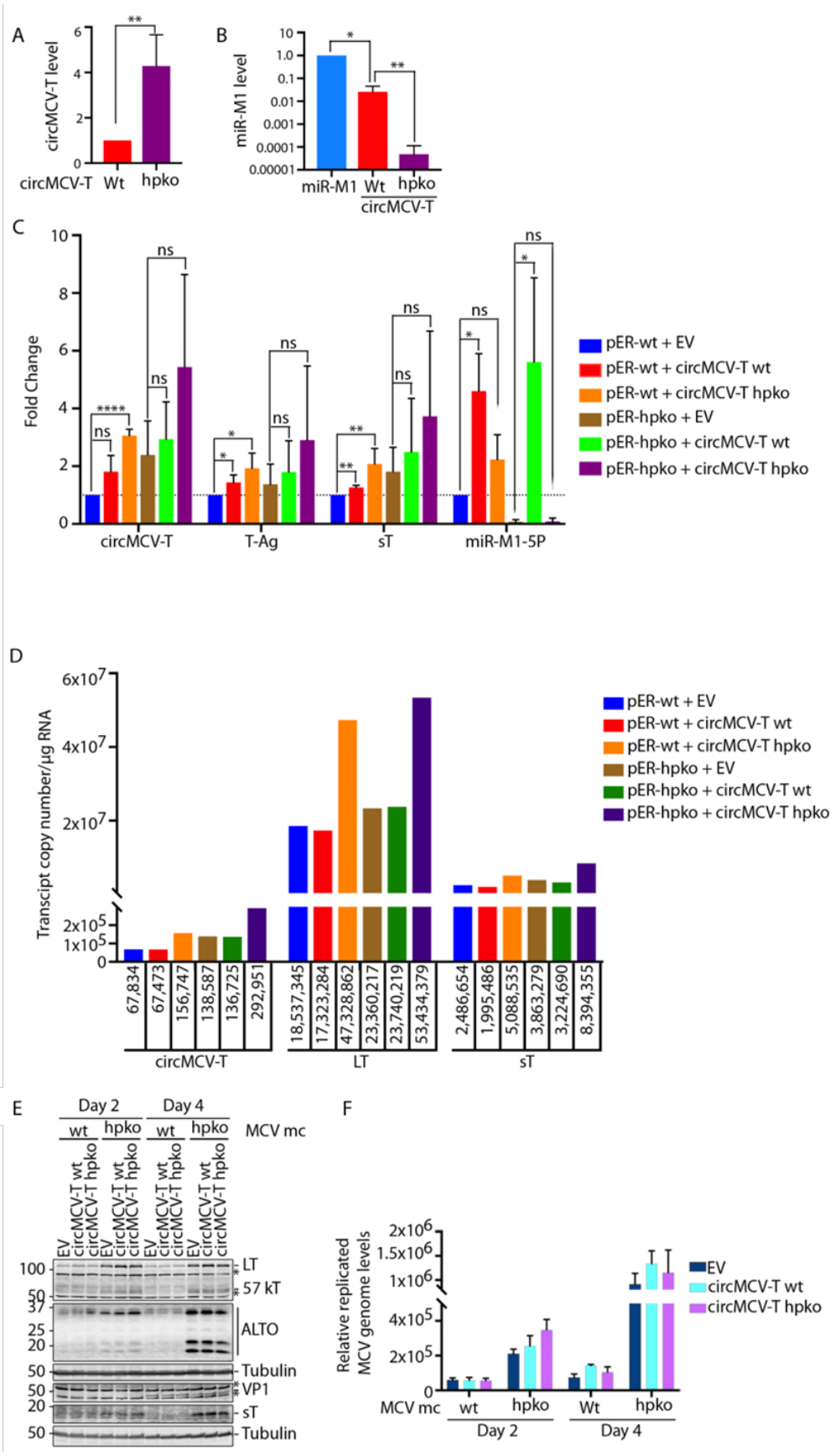

A
circMCV-T level
circMCV-T
Wt
hpko
B
miR-M1 level
miR-M1
Wt
hpko
circMCV-T
C
Fold Change
circMCV-T
T-Ag
sT
miR-M1-5P
pER-wt + EV
pER-wt + circMCV-T wt
pER-wt + circMCV-T hpko
pER-hpko + EV
pER-hpko + circMCV-T wt
pER-hpko + circMCV-T hpko
D
Transcipt copy number/µg RNA
67,834
67,473
156,747
138,587
136,725
292,951
18,537,345
17,323,284
47,328,862
23,360,217
23,740,219
53,434,379
2,486,654
1,995,486
5,088,535
3,863,279
3,224,690
8,394,355
circMCV-T
LT
sT
E
Day 2
Day 4
wt
hpko
MCV mc
EV
circMCV-T wt
circMCV-T hpko
LT
57 kT
ALTO
Tubulin
VP1
sT
Tubulin
F
Relative replicated MCV genome levels
MCV mc
Day 2
Day 4
EV
circMCV-T wt
circMCV-T hpko

Figure 20 Effect of circMCV-T on MCV early transcript expression and viral replication.

(A and B) 293 cells were transfected with the MCV miR-M1 or the circMCV-T wt or hpko mutant expression vector for 48 h, and RT-qPCR was performed for circMCV-T (A) and miR-M1-5P (B). (C and D) 293 cells were transfected with pCMV-ER wt or hpko constructs together with the circMCV-T wt or hpko mutant expression vector or an empty vector (EV) control for 48 h. RT-qPCR was performed for circMCV-T and for linear transcripts sT and LT as well as miR-M1-5P. C_T values were normalized to RNase P or GAPDH C_T values, and samples were normalized to EV-transfected cells for both the wt and hpko MCV ER construct to calculate fold change (C). Or C_T values were used to calculate copy numbers per microgram of total RNA of each transcript shown from a standard curve. Actual copy numbers for each transcript are shown at the bottom of the graph (D). Bars and error bars in panel C indicate means ± SD from three independent experiments. Statistical analysis was performed using an unpaired t test in panel A, an unpaired t test with Welch's correction in panel B, and ordinary one-way analysis of variance (ANOVA) in panel C. Significance levels: *, P < 0.05; **, P < 0.01; ***, P < 0.001; ****, P < 0.0001; ns, not significant. (E and F) 293 cells were cotransfected with a wt or hpko mutant MCV-HF mc construct together with a circMCV-T wt or hpko mutant expression vector or an EV control. Cells were collected after 2 and 4 days of transfection. (D) Western blot analysis of MCV-encoded proteins (LT, 57kT, ALTO, sT, and VP1; α-tubulin was used as an internal control). (E) Quantification of the replicated (DpnI-resistant) MCV genome by qPCR analysis. Genomic GAPDH was used to normalize C_T values, and relative levels of the MCV genome were calculated according to the ΔΔCT method. Bars and error bars represent means ± SD from 3 replicates. Results are representative of three independent experiments.

3.3.7 Identification of Rat Polyomavirus 2 (RatPyV2) Encoded CircRNAs by RNase R^+ RNA Sequencing.

We performed RNase R^+ sequencing on parotid gland RNA of a naturally infected SCID Rat supporting active rat polyomavirus 2 (RatPyV2) viral replication [396]. This model provides the advantage of an episomally replicating viral system under natural infection conditions in

comparison to the MCV artificial transfection system in 293 cells. We identified 4 circRNAs; circRatPyV2_4112_4468 (the region homologous to circMCV-T), circRatPyV2_1653_1998, circRatPyV2_850_4468, and circRatPyV2_786_1266, from the ER using the RatPyV2 genome (NCBI accession no. KX574453.1) as a reference (**Figure 21 and Table 16**). There were also two (RatPyV2_561_3856 and RatPyV2_561_3308) BSJ reads from the RatPyV2 LR region; however, although the RatPyV2 LR transcript is not fully annotated, these reads are likely to also result from leader-to-leader forward splicing events, in like fashion to the MCV BSJ-simulating reads from the MCV LR, instead of from a bona fide backsplicing process. As with MCV, the contribution of BSJ reads from RatPyV2 to the total RNase R resistance reads in this system is also very low (<0.001 RPM). Nevertheless, the most abundant putative RatPyV2 circRNA from the ER, circRatPyV2_4112_4468, uses backsplicing sites (splice acceptor AG/GT and splice donor /GTAAGT) homologous to the circMCV-T (**Table 14**), suggesting a common mechanism of post-transcriptional RNA processing between the two polyomaviruses.

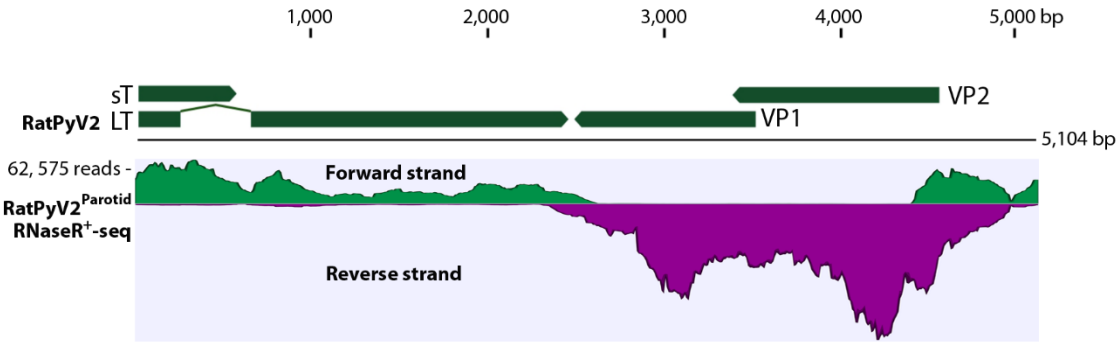


Figure 21 Identification of RatPyV2-encoded circRNAs by RNase R^+ RNA sequencing.

Schematic representation of the RatPyV2 genome organization (top) and RNase R^+ RNA sequence coverage of RatPyV2 in the parotid gland of an infected X-SCID rat (bottom). Sequence reads from the forward strand (green) and reverse strand (magenta) are shown.

Table 16 Sequenced BSJs from RatPyV2

RatPyV2 BSJ Start (nucleotide number)	RatPyV2 BSJ End (nucleotide number)	Strand	BSJ reads	RPM
561	3308	+	8	
561	3856	+	4	
786	1266	-	3	
850	4468	-	3	
1653	1998	-	3	
4112	4468	-	11	0.001

Tissue= RatPyV2 infected parotid gland from an X-SCID Rat, GenBank ID KX574453.1

3.4 Discussion

CircRNAs are a new class of closed circular transcripts whose biological relevance is becoming increasingly appreciated. Circular RNA transcripts have recently been recognized to be an important component of the transcriptome profiles of the human oncogenic DNA viruses EBV, KSHV, and HPV [293, 295, 296, 301]. In this study, we show that another human DNA tumor virus, MCV, encodes circRNAs from its early region (ER). Further, transcriptome analysis of RatPyV2 identified BSJs in virus-infected rat tissues harboring an actively replicating episomal RatPyV2 genome [396], supporting the notion that expression of circRNA can reasonably be generalized to the family of polyomaviruses, not just human polyomaviruses or oncogenic polyomaviruses.

MCV circRNA BSJs were detected by RNase R sequencing only from cells supporting active viral replication. The absence of circMCV-T detection by RNase R sequencing from MCC-derived cell lines, which do not support active viral replication, can be attributable in part to the low level of viral transcripts expressed in these cells as well as viral genome integration, which

may disrupt continuous ER RNA transcripts needed as the templates for backsplicing. Follow-up studies using more sensitive and directed PCR detection methods on a panel of MCC cell lines, all with nonreplicating integrated viral genomes and expressing tumor-specific signature T antigen truncations, show that some (CVG-1 and WaGa) do express circRNAs. Nevertheless, even with the increased sensitivity of direct PCR detection, several MCV-infected MCC cell lines (MKL-1, MKL-2, and MS-1) do not **(Figure 16C**). The circMCV-T BSJ was detected by RT-PCR in two out of three patient MCC tumors tested. The third clinical sample did not have intact RNA after extraction. Expanded testing of fresh tumor samples will clarify whether circRNAs are expressed in all tumors or in only a subset, similar to what is seen with MCC-derived cell lines.

The most abundant MCV circRNA, circMCV-T, incorporates the entire exon II from early region pre-mRNA through backsplicing. Consistently, we find that circMCV-T expression patterns mirror 57kT transcript expression in MCC-derived cell lines. Both circMCV-T and the 57kT transcripts were detected in CVG-1 and WaGa but not in MKL-1, MKL-2, or MS-1 cells (**Figure 16C**). Like other MCC-derived cell lines and primary tumors, CVG-1 expresses a truncated LT protein. In the case of CVG-1, the truncated product is due to a premature stop codon encoded by a single A-T mutation at nt 1617 [400]. This tumor-associated mutation does not change the splicing motif at either the donor (at nt 1622) or the acceptor (at nt 861) site required for circMCV-T backsplicing. WaGa cells, derived from ascites fluid [397, 408], similarly express a truncated LT protein due to a single C-T mutation at nt 1461, which also is not predicted to affect circMCV-T formation. Additionally, both of the circMCV-T-positive cell lines WaGa and CVG-1 contain a relatively high number of integrated viral genome copies per cell, in contrast to MKL-2 and MS-1 cells, which do not express circMCV-T [400, 408]. This may explain the increased circMCV-T and 57kT levels detected in these cell lines (**Figure 16C**). On the other hand, MKL-1

cells, which display integrated viral genome copy numbers comparable to those of CVG-1 cells [400], also did not express detectable levels of circMCV-T (**Figure 16C**); however, a 46-nt deletion between nt 1612 and 1657 abrogates the circMCV-T backsplice donor site at position 1622 in these cells. Using a single primer pair that can amplify and distinguish between both linear LT and 57kT transcripts in the same PCR, the full-length LT mRNA product is detected only in MCV-HF recircularized-genome-transfected cells and not in either the MCC-derived cell lines or primary tumor samples (**Figure 16C and D**). In contrast, the sT transcript is detected at a relatively consistent level across cell lines and primary tumors tested. This suggests that transcripts that give rise to sT proteins may be differentially regulated from other T antigen transcripts.

Some circRNAs have been shown to code for a protein product through an IRES or m^6A-mediated cap-independent translation initiation [301, 342, 343, 409]. However, we did not find either endogenous or overexpressed circMCV-T to be significantly associated with polysomes, and thus it is unlikely to encode protein(s) (**Figure 17C and D**). Although the circMCV-T sequence contains the start codon for the ALTO translation frame, there is no stop codon incorporated in the circle from the same reading frame (**Figure 18A**). Western blot analysis, using antibodies that recognize the LT reading frame (CM2B4) or the ALTO reading frame (CM7B1) in the circMCV-T coding region, also did not detect a protein product from circMCV-T (**Figure 18B**).

In contrast to most circRNAs, which are resistant to exonuclease digestion due to the absence of free RNA ends [333, 347], MCV-encoded circMCV-T is sensitive to RNase R treatment, suggestive of linearization. Nevertheless, RPAD shows the presence of true circular forms of circMCV-T, albeit at low abundances (**Figure 2E**). We considered that this might be explained by the presence of MCV miRNA. The MCV genome in the same region that produces

circMCV-T also elaborates an miRNA, miR-M1 from the antisense strand, which is processed into two mature miRNA seed sequences, miR-M1-5P and miR-M1-3P, that are 100% complementary not only to linear T-Ag transcripts but also to circMCV-T [72, 76]. Perfect sequence complementary between a miRNA and its target sequence is predicted to result in RISC-mediated endonucleolytic cleavage of the target RNA. Recent studies [388, 410] have shown that such an interaction between miR-671 and a circRNA target CDR1-as results in Ago2-mediated cleavage and degradation of the circRNA. In support of this notion, MCV miR-M1 has been shown to target and cleave T-Ag transcripts to regulate MCV replication and promote episomal persistence [72, 76]. Consistently with what occurred after miR-M1-mediated linearization of circMCV-T, exogenous expression of miR-M1 led to decreased circMCV-T levels accompanied by increased RNase R sensitivity (**Figure 19D and E**), and hpko mutagenesis, which removes miR-M1 expression from the MCV ER, increased circMCV-T levels compared to those of a wt construct (**Figure 19B and 5C**). miR-M1 has been shown to dominate the spectrum of miRNAs expressed in cells harboring replicating MCV episomes, in contrast to MCC-derived cells and primary tumors, which expresses miR-M1 only at very low levels [76, 411]. This might explain the sparse amounts of viral circRNA BSJs detected from cells with replicating MCV genomes in our current study. The provocative directionality of the depletion of sequence coverage at the circRNA locus (**Figure 15B**) may be attributable to the 3'-to-5' exonuclease activity preference of RNase R. It is also possible that circMCV-T itself forms a hairpin structure that can be processed by Drosha in a fashion similar to that of the pre-miR-M1 RNA from the opposite strand, which would in the end lead to its RNase R sensitivity. Future experimentation using Drosha knockout cells may shed light on this.

Another intriguing consequence of circMCV-T and MCV miRNA interaction is the effect of circMCV-T on its targeting miRNA. Although miRNAs are usually stable due to protection by the Argonaut protein complex from exonucleolytic degradation, accumulating evidence suggests that extensive complementarity between a target RNA and an miRNA may lead to target RNA-directed miRNA degradation (TDMD) [410, 412-414]. Since mature miR-M1 products can bind circMCV-T with 100% complementarity, this may suggest a mechanism that promotes MCV miR-M1 for depletion during active viral replication.

Expression of circMCV-T increased T-Ag (sT and LT) transcripts only to a modest level (**Figure 20C**). This can be, in part, due to additional expression of miR-M1 from the antisense strand of the pLaccase-circMCV-T expression construct (**Figure 20B**) and, thus, to further degradation of linear T-Ags canceling out the effect of circMCV-T expression. An additional mechanism for a circMCV-T-mediated increase in linear T-Ag transcripts independent of circMCV-T's role as a decoy to miR-M1 is shown by the increase in the levels of these transcripts in cells cotransfected with hpko mutant pLaccase-circMCV-T, which lacks miR-M1 binding sites (**Figure 20A**). Consistently, coexpression of the hpko mutant of the MCV ER together with an hpko-circMCV-T resulted in the highest level of T-Ag transcripts, although this increase is not statistically significant. Regardless, exogenous expression of circMCV-T promotes MCV replication to only a limited extent (**Figure 20E and F**), while abrogating miR-M1 expression from the viral genome alone substantially increased T-Ag protein expression and viral replication, consistent with the results of previous reports [76]. This reciprocal interaction between two different classes of viral RNAs reveals an entirely new level of complexity in the interplay between viral gene products. In the presence of circMCV-T, miR-M1 is functionally sequestered from linear T-Ag transcripts that code for replication proteins, and therefore, MCV replication is

favored; at low levels of circMCV-T, miR-M1 binds and degrades linear T-Ag transcripts and thus suppresses active viral replication to promote persistent infection. Finally, competitive splicing may also be an attractive concept for circMCV-T function. Because circMCV-T uses the same splice sites as canonical linear splicing events for LT and 57kT, its production may come at the expense of these linear transcripts. This may be a mechanism that can shift the balance of T-Ag expression toward sT. In this scenario, the presence of noncoding circMCV-T molecules that present miR-M1 sites may be an additional factor that protects sT Ag transcripts. CircMCV-T may be part of a program that includes sT transcript regulation to enhance lytic virus replication; however, miR-M1 expression suppresses lytic replication to establish MCV latency. Our findings suggest that MCV-encoded circRNAs represent another layer of regulation to fine-tune MCV replication through a balanced titration of three classes of MCV transcripts: circMCV-T, miR-M1, and mRNAs (**Figure 22**).

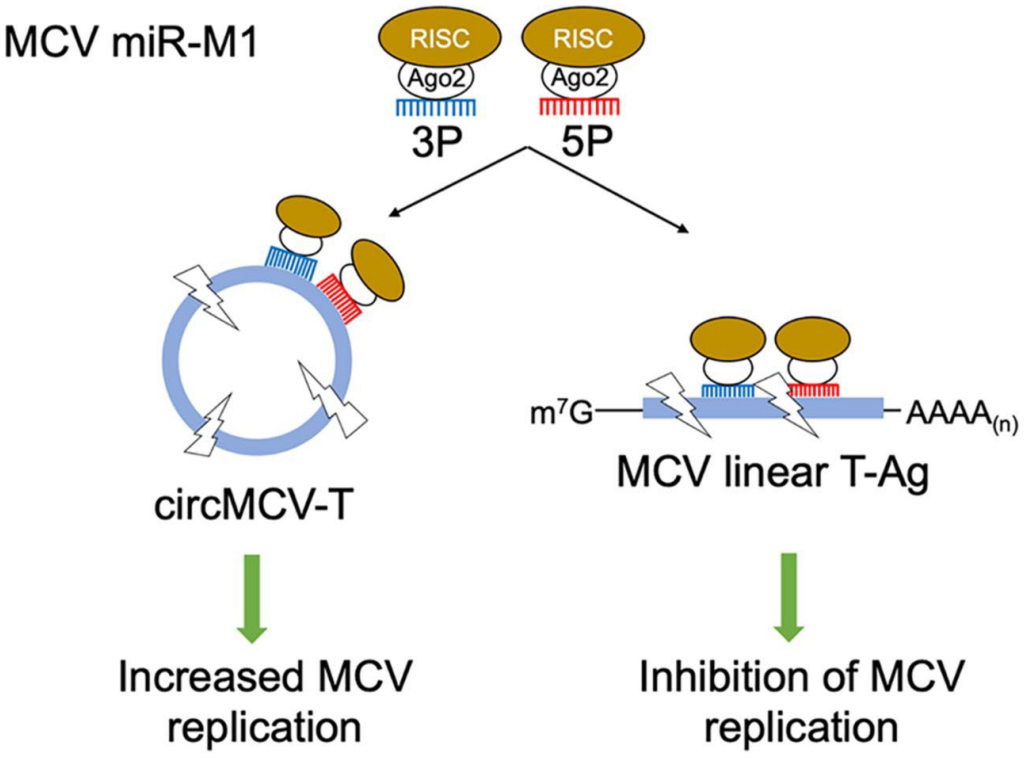


Figure 22 Working model for MCV circMCV-T regulation of MCV T-Ag expression.

MCV miR-M1 is incorporated into the RISC complex, which will target both circMCV-T and linear T-Ag transcripts. If circMCV-T absorbs the miR-M1-loaded RISC complex, linear T-Ag transcripts will be stabilized and T-Ag expression will increase. In the absence of circMCV-T, linear T-Ag transcripts will be degraded by the miR-M1-induced RISC complex, which leads to inhibition of T-Ag expression and, therefore, MCV replication.

4.0 Replication Kinetics for a Reporter Merkel Cell Polyomavirus

Work described in this chapter was published in Viruses. 2022 Feb 25;14(3):473.

With authors Abere B., Zhou H., Shuda M., Stolz D.B., Rapchak K., Moore P. S., and Chang Y.

B. Abere and H. Zhou contributed equally to this work.

H. Zhou constructed MCV minicircle plasmids, and performed Western blot, qPCR, flowcytometry, pseudovirus-packaged virion preparation, and the reporter MCV infection. B. Abere performed the immunofluorescent assay. B. Abere, M. Shuda, and K. Rapchak performed the native MCV minicircle infection. M. Shuda and K. Rapchak performed the native virion preparation and Southern blot. D.B. Stolz performed the electron microscopy of MCV virions. B. Abere wrote the manuscript with the help of H. Zhou. P.S. Moore and Y. Chang conceived the project and revised the manuscript.

As discussed in Chapter One, the study of Merkel cell polyomavirus (MCV) and its pathogenesis has been hindered by the lack of a robust viral infection/transfection system. Historically, the field has relied on a ligation technique to produce circularized viral genomes by inserting the MCV genome into a bacterial backbone and replicating the plasmid in bacteria. After purification, plasmids are digested by restriction enzymes to remove the bacterial backbone, and then recircularized using T4 ligase treatment. However, this system has limitations, including low yields of the MCV circular genome and a tendency for concatenation during the ligation steps.

In this chapter, we adapted minicircle technology introduced in Chapter One. The MCV genome was synthesized and integrated into the minicircle vector between recombination sites. The recombination was induced by arabinose, and the recircularized MCV genome was purified using a plasmid extraction kit. To verify that the MCV minicircle (MCVmc) recircularized by ΦC31, which contains a 39 bp residual recombination site between the early and late regions, has the same capacity to express viral proteins and replicate as the wild-type (wt) MCV genome recircularized by ligation, we employed qPCR, Western blot (WB), and immunofluorescence assay (IFA) to quantify MCV genome copy number, viral protein levels, and VP1 positive cell numbers at 2- and 4-day post-transfection. Our data demonstrate that the 39 bp non-viral sequence does not impact MCV protein expression and replication. Moreover, we assessed the capacity of MCVmc to produce infectious virions and establish transmission *in vitro* using immunofluorescence assay. Utilizing the MCV minicircle system, we successfully introduced mutations into the viral genome that had previously been found to have functional consequences and measured the replication of these mutants using IFA. These mutant infectious clones validated previous functions attributed to LT-host cell protein interactions and confirmed these as host cell restriction factors of viral replication.

To further utilize minicircle technology, we inserted an mScarlet coding cassette at the 3' end of the VP1 gene, which can express an mScarlet fused VP1. In parallel, we constructed VP1-mScarlet with a P2A linker between VP1 and mScarlet, resulting in ribosome skipping during translation and dissociating the VP1 and mScarlet proteins. Both versions expressed viral proteins and replicated at similar levels as wt MCVmc, as validated by WB, qPCR, and IFA. We employed these reporter viruses to study MCV replication kinetics after transfection using flow cytometry and WB. Although LT, representing early gene expression, peaked at days 3-6, the VP1 positive

cell population continued to increase. We also sought to establish an infection model using reporter MCVs; however, only wt MCVmc produced benzonase-resistant genomes, indicating that the presence of additional peptides on the C-terminus of VP1 abolishes its packaging ability. Consequently, we employed pseudovirus packaging, co-expressing wt VP1 and VP2, to package reporter MCV genomes. The pseudovirus-like reporter MCVmcs demonstrated the capacity to infect 293 cells and express mScarlet after infection.

In summary, we adopted a recombinase-based minicircle technology to recircularize bacteria-produced MCV genomes and demonstrated that MCVmc has the same capacity to express viral proteins, replicate, form viral particles, and infect human fibroblast cell lines as its wt counterpart. We also introduced multiple mutations into MCVmc to validate previously ascribed functions of LT-host cell protein interactions, confirming these as host cell restriction factors of viral replication. Additionally, we constructed two reporter MCVmcs with an mScarlet fused to the C-terminus of VP1, either with or without a P2A linker. Using these reporter MCVmcs, we were able to employ flow cytometry to study the replication kinetics of MCV. We also successfully packaged reporter MCVmcs with co-expression of wt VP1 and VP2 and infected 293 cells. Our study demonstrates that MCV possesses multiple, self-encoded viral restriction mechanisms that promote viral latency over lytic replication. These mechanisms can now be effectively investigated using recombinase technology.

4.1 Introduction

Merkel cell polyomavirus (MCV) causes most cases of Merkel cell carcinoma (MCC), an uncommon human skin cancer associated with immunosuppression [415] and aging [128]. MCC

is highly aggressive with a 54% 5-year survival rate [416]. MCV is a near-ubiquitous component of viral skin flora; it causes MCC if the viral genome becomes integrated into the host genome and acquires mutations ablating replication functions [38, 66]. Unlike virus-positive MCC, virus-negative MCC shows high levels of UV-induced somatic mutagenesis that phenocopy MCV oncoprotein functions [417-421].

MCV is a nonenveloped, double-stranded DNA virus with a 5.4 kb genome partitioned into early (ER) and late regions (LR) by a non-coding control region (NCCR) [38, 146]. In addition to the replication origin, the NCCR contains promoter elements regulating early and late gene expression [147]. The ER encodes several proteins designated tumor antigens (T-Ags), such as large T-Ag (LT), small T-Ag (sT), 57kT [66], and alternative LT open reading frame (ALTO) [67]. The ER also expresses an auto-repressive miRNA [72, 76] as well as circular RNAs that antagonize miRNA activity [150] and may express alternative T antigens through cap-independent translation [151]. The LR expresses at least two viral structure proteins, VP1 and VP2, which comprise the viral capsid [102]. Despite the simplicity of its two-gene genome (ER and LR), actual MCV gene expression and replication is complex due to alternative splicing, alternative translation initiation, promoter regulation, proteostatic regulation, and miRNA-circRNA feedback loops [66, 67, 147, 150].

MCV is widely detected by PCR on skin and by serologic assays in the blood of individuals world-wide, providing evidence for persistent life-long infection [46, 163, 422, 423]. Despite this, direct visualization of the virus or viral molecules in non-neoplastic human skin has not been achieved, presumably due to low copy number and viral latency. Furthermore, studies on the MCV life cycle have been hampered by an inability to achieve permissive replication and infection in tissue culture [102, 138, 149, 424, 425]. Recombinant MCV genomes can be circularized and

transfected into producer cells to harvest viruses for infection studies in human dermal fibroblasts [176], but this approach is technically complex.

The replication lifecycle of MCV is similar, but not identical to replication in other polyomaviruses [138, 149, 426]. Based on *in vitro* virus generation and virus-like particle (VLP) studies, MCV capsids bind to heparin sulfate moieties on host cell surfaces followed by secondary interaction with sialylated glycans [103, 424]. MCV utilizes caveolin/lipid raft-mediated endocytosis to transit the host cell plasma membrane but only minor populations of the internalized MCV virions are able to reach the endoplasmic reticulum, which has been suggested to be a bottleneck step during MCV infection [111, 427]. It is not clear how MCV traffics from the endoplasmic reticulum to the nucleus.

After the MCV virion unpacks its genome in the nucleus, T-Ags are expressed that facilitate genome replication. LT is the only viral protein essential for MCV DNA replication [138, 146]. It binds as a multimeric complex, presumably in a head-to-head double hexameric configuration, to GAGGC-like pentanucleotide sequences in the NCCR at the origin of replication (MCVori) and serves as an ATP-dependent DNA helicase [146]. Similar to SV40, formation of the replication initiation complex is presumed to involve the co-operative interaction of MCVori-bound LT, cellular DNA polymerase alpha/primase and replication protein A (RPA) [428-431]. While LT is absolutely required for replication, MCV sT is not, but it nevertheless serves as an accessory protein to enhance MCV replication [138, 146]. By interacting with the SCF E3 ligase Fbw7 complex via its LT stabilization domain (LTSD), sT inhibits Fbw7-related ubiquitin-mediated degradation of LT to markedly increase LT protein and MCV replication [159]. The contributions of 57kT and ALTO to MCV replication remain unknown.

Similar to other polyomaviruses [432, 433], MCV LT appears to also exhibits transcriptional activity to auto-repress early gene transcription possibly by binding to as yet unmapped regions in the NCCR [147]. This places a natural transcriptional brake on MCV replication to inhibit the full lytic viral lifecycle. Multiple cellular factors are also known to interact with LT to restrict MCV genome replication. E3 ligases Skp2, β-TrCP and Fbw7 interact with specific phosphorylated residues to promote LT ubiquitylation and proteasomal degradation [147]. The evolutionarily conserved inhibition of MCV replication by degradation of preformed MCV LT is described as proteostatic viral latency [434]. Other restriction factors that affect LT's replication activity include interaction with the deubiquitnase ubiquitin-specific protein 7 (Usp7) [238] and the vacuolar sorting protein hVam6p (Vps39) [138, 194]. While Skp2-LT interactions may be shared with other polyomaviruses [434], it is largely unknown whether MCV LT-specific restriction factors are applicable to other polyomaviruses.

In addition to protein restriction pathways, MCV also encodes its own microRNA (miR-M1) from the negative strand of its ER [72, 76, 435]. MCV miR-M1 restrains MCV genomic replication by degrading LT antigen mRNA. A positive-strand circular RNA produced by backsplicing from exon 2 of the LT transcript antagonizes miR-M1, to enhance MCV replication [150]. The finely tuned choreography and interplay of these cellular and viral repression/amplification factors is likely to determine whether the virus enters into lytic virion production or remains as a latent episome in cells. Although this repertoire of replication control circuits is likely to provide a sophisticated cell type- and cell environment-specific mechanism for the binary decision to replicate or remain latent, the actual measurement of this interplay is limited by the absence of a permissive replication model.

To overcome these challenges, we adapted a recombinase-mediated minicircle system [184] to produce an MCV minicircle (MCVmc), which dramatically increases total circularized genomic DNA yield (from micrograms to milligrams) on a bacterial culture per volume basis. We show that mutagenesis of specific LT and miR-M1 sites robustly increases virus production, including support of the Fbw7-affected residue at LT serine 293 that has been confirmed [436] but challenged as a functional site [437]. By introducing a fluorescent tag fused to the C-terminus of VP1, we can track MCV replication after transfection using flow cytometry and can generate a biosafe, one-round infection/replication model to measure viral replication kinetics.

4.2 Materials and Methods

4.2.1 Cells

BJ-hTert cells were established by retroviral transduction of Babe hTert-puro (pBabe-hTert-puro: a gift from Roderick J. O'Sullivan, Hillman Cancer Center, University of Pittsburgh, Pittsburgh, PA, USA) into primary BJ foreskin fibroblast (ATCC CRL-2522) and selected with 2 μg/mL puromycin. To establish 293 TRE-sTco cells, 293 (ATCC CLR-1573) were transduced with pLenti TRE MCV sT [400] and selected with 2 μg/mL puromycin. The 293, 293 TRE-sTco, BJ and BJ-hTert cells were maintained in Dulbecco's Modified Eagle Medium (DMEM; Corning, Manassas, VA, USA) supplemented with 10% FBS, while HFF-1 (ATCC SCRC-1041) cells were maintained in DMEM supplemented with 15% FBS. Human primary bone marrow or adipose tissue-derived mesenchymal stem cells (MSC-bm and MSC-a) [438], kindly provided by Shou-

Jiang Gao (Hillman Cancer Center, University of Pittsburgh, PA, USA), were maintained in MSC medium (MSCM; ScienCell Research Laboratories, Carlsbad, CA, USA).

4.2.2 Plasmids and Constructs

The construction of pJ-MCV-HF, pMC-MCV and pMC-MCV-hpko plasmids have been previously described [138, 439]. The MCV minicircle plasmids generated by restriction enzyme digestion cloning are listed in **Table 17**. To generate pMC-MCV-VP1-mS, a fragment containing the C-terminus of VP1 fused to mScarlet and part of the pMC backbone was produced by overlapping PCR using primers VP1 (2440-2466) F, VP1-mScarlet R, VP1-mScarlet F and mScarlet-attBVec R and was subsequently cloned into the pMC-MCV plasmid using PacI and XmaI restriction sites. pMC-MCV-VP1-P2A-mS was produced by restriction digestion of a fragment containing the C-terminus of VP1 fused to P2A-mScarlet and part of the pMC backbone generated with a GeneArt™ Seamless Cloning and Assembly Kit (cat. no. A13288; Thermo Fisher, Carlsbad, CA, USA) using primers VP1-P2A_FW, VP1-P2A_RV, P2A-mScarlet_FW and P2A-mScarlet_RV and subsequent cloning into the pMC-MCV backbone using PacI and XmaI restriction sites. The list and sources of all constructs are shown in **Table 18**. Primers used for plasmid construction are shown in **Table 19**.

Table 17 Construction summary of pMC-MCV harboring mutations.

Final Constructs	Mutation Sites	Insert Source	Restriction Enzymes
pMC-MCV-Rep-	C44A	pJ-MCV-HF-Rep-	AvrII, SacI
pMC-MCV-hVam6p-	T1251G, G1252C	pJ-MCV-HF-hVam6p-	BamHI, EcoRI
pMC-MCV-β-TrCP-	T1065G	pJ-MCV-HF-β-TrCP-	AvrII, BamHI
pMC-MCV-Skp2-	T1284G	pJ-MCV-HF-Skp2-	BamHI, EcoRI
pMC-MCV-Fbw7-	T1341G	pJ-MCV-HF-Fbw7-	BamHI, EcoRI

Table 18 List and description of plasmid constructs.

Construct	Function	Parental Vector	ChangMoore Plasmid #	ChangMoore Plasmid Name	Cat#
pSMART-MCV-HF	Amplifies MCV-HF genome in bacteria for MCV genome re-circularization by re-ligation after enzyme digestion which release MCV-HF genome from bacterial backbone.	pSMART	4556	MCV-HF	N/A
pJ-MCV-Rep-	The donor of Rep- mutation (C44A) for minicircle cloning.	pJ	3148	MCV-Rep-	N/A
pJ-MCV-hVam6p-	The donor of LT hVam6p binding deficent mutation (T1251G, G1252C) for minicircle cloning.	pJ	3149	MCV-hVam6p-	N/A
pJ-MCV-β-TrCP-	The donor of LT β-TrCP binding deficent mutation (T1065G) for minicircle cloning.	pJ	3987	MCV-HF/LT.S147A	N/A
pJ-MCV-Skp2-	The donor of LT Skp2 binding deficent mutation (T1284G) for minicircle cloning.	pJ	3988	MCV-HF/LT.S220A	N/A
pJ-MCV-Fbw7-	The donor of LT Fbw7 binding deficent mutation (T1341G) for minicircle cloning.	pJ	3989	MCV-HF/LT.S239A	N/A
pMC-MCV	Amplifies MCV-HF genome in bacteria for MCV genome re-circularization by recombination.	pMC.BESPX	4587	pMC.BESPX-MCV-HF	N/A
pMC-MCV-Rep-	Amplifies replication deficient MCV-HF genome in bacteria for MCV genome re-circularization by recombination.	pMC.BESPX	4672	pMC.BESPX-MCV-HF-Rep-	N/A
pMC-MCV-hVam6p$^-$	Amplifies MCV-HF genome that expresses hVam6p binding deficient LT in bacteria for MCV genome re-circularization by recombination.	pMC.BESPX	4673	pMC.BESPX-MCV-HF-hVam6p$^-$	N/A
pMC-MCV-β-TrCP$^-$	Amplifies MCV-HF genome that expresses β-TrCP binding deficient LT in bacteria for MCV genome re-circularization by recombination.	pMC.BESPX	4674	pMC.BESPX-MCV-HF-β-TrCP$^-$	N/A
pMC-MCV-Skp2$^-$	Amplifies MCV-HF genome that expresses Skp2 binding deficient LT in bacteria for MCV genome re-circularization by recombination.	pMC.BESPX	4675	pMC.BESPX-MCV-HF-Skp2$^-$	N/A
pMC-MCV-Fbw7$^-$	Amplifies MCV-HF genome that expresses Fbw7 binding deficient LT in bacteria for MCV genome re-circularization by recombination.	pMC.BESPX	4676	pMC.BESPX-MCV-HF-Fbw7$^-$	N/A
pMC-MCV-VP1-mScarlet	Amplifies MCV-HF genome that expresses mScarlet-fused VP1 in bacteria for MCV genome re-circularization by recombination.	pMC.BESPX	4764	pMC.BESPX-MCV-HF-mScarlet	N/A
pMC-MCV-VP1-P2A-mScarlet	Amplifies MCV-HF genome that expresses individual mScarlet and VP1 using a P2A sequence in bacteria for MCV genome re-circularization by recombination.	pMC.BESPX	4765	pMC.BESPX-MCV-HF-P2A-mScarlet	N/A
pwM	Expresses MCV VP1 protein in mammalian cells.	pGwf	2972	MCV VP1 (PWM)	Addgene# 22515
ph2m	Expresses MCV VP2 protein in mammalian cells.	phGf	2973	MCV VP2 (PH2M)	Addgene# 22518

pmScarlet_C1	Express mScarlet in mammalian cells.	pC1	4738	pmScarlet-C1	Addgene# 85042
pEGFP-N1	Express EGFP in mammalian cells.	pEGFP-N1	2437	pEGFP-N1	Clontech# 6085-1

Table 19 List of primers used for overlapping PCR and Gibson assembly.

VP1 (2440-2466) F	TGACACATTGCAGATGTGGGAGGCAAT
VP1-mScarlet R	GCCTCGCCCTTGCTCACCATTAATTCTTGTGTTTGGCTTT
VP1-mScarlet F	AAAGCCAAACACAAGAATTAATGGTGAGCAAGGGCGAGGC
mScarlet-attBVec R	TCCCCGGGCGCGACAAATAATTCTCACTTGTACAGCTCGT
VP1-P2A R	GTCTCCAGCCTGCTTCAGCAGGCTGAAGTTAGTAGCTCCGCTTCCTAA TTCTTGTGTTTGGCTTTCTTTTTGAGAGGCC
P2A-mScarlet F	AGCCTGCTGAAGCAGGCTGGAGACGTGGAGGAGAACCCTGGACCTAT GGTGAGCAAGGGCGAGGCA
VP1-P2A_FW	GCCAAGCTTGCATGCCGTTCTTTAATTAATGTTCATTATT
VP1-P2A_RV	CCAGCCTGCTTCAGCAGGCTGAAGTTAGTAGC
P2A-mScarlet_FW	TGCTGAAGCAGGCTGGAGACGTGGAGGAGAAC
P2A-mScarlet_RV	AATTCGAGCTCGGTACTCCCCGGGCGCGACAAATAATTCT

4.2.3 Recircularization of MCV Genome by *In Vitro* Ligation and Mini-Circle System

Recircularization of MCV-HF genome (GenBank accession #JF813003) by *in vitro* ligation was previously described [138]. To produce the recombinant MCV clone, the MCV-HF genome was cloned into the pSMART-LC-Amp vector (cat. no. 400300-2; Lucigen) using an EcoRI site present only once in the MCV genome. Since the resulting clone pSMART MCV-HF has two EcoRV blunt cutter enzyme sites immediately outside the EcoRI sites, a double digest using EcoRI and EcoRV results in a blunt-end vector fragment and an MCV-HF fragment with EcoRI cohesive ends for self-recircularization. pSMART-MCV-HF digested with EcoRI and EcoRV was recircularized by T4 DNA ligase at a low concentration (2.5 ng/μL) to reduce the formation of MCV concatemers.

Recombinase-mediated MCV recircularization was recently reported [439]. In brief, MCV-HF sequences with or without mutations (see plasmids and constructs, **Table 18**) were cloned into the pMC.BESPX plasmid which contains AttB and AttP sites. Constructs were

transformed into ZYCY10P3S2T (cat. no. MN900A-1; System Biosciences, Palo Alto, CA, USA) competent cells, a kind gift from Mart Ustav (University of Tartu, Tartu, Estonia) and Alison McBride (National Institute of Allergy and Infectious Diseases, Bethesda, MD, USA). Transformed bacteria were cultured overnight in Terrific Broth (cat. no. T0918; Sigma-Aldrich, St. Louis, MO, USA) to an OD600 of 4-6 and induced by using an equal volume of buffer (0.04% L-arabinose and 40 mM NaOH in LB). After 6 h of induction, bacteria were harvested and MCVmc DNA genomes were purified using a Maxi prep kit (cat. no. 740416; Macherey-Nagel, Düren, Germany).

4.2.4 MCVmc Transfection and Replication Assay

In 6-well plates, 293 cells were plated and transfected with 1 µg of MCV-ligated, MCVmc, MCVmc-hpko, MCVmc-Rep^{-}, MCVmc-Skp2^{-}, MCVmc-Fbw7^{-} using FuGENE (cat. no. E2311; Promega, Madison, WI, USA), following the manufacture's protocol. After 2- or 4-day post-transfection, cells were collected for protein and DNA extraction.

Total genomic DNA was isolated using DNAzol reagent (cat. no. 10503027; Thermo Fisher, Warrington, UK), according to the manufacturer's instructions, and resuspended in 0.1× TE buffer (1 mM Tris·HCl [pH 8.0], 0.01 mM EDTA [pH 8.0]). Then, 5 µg of total genomic DNA was digested overnight using DpnI and BamHI restriction enzymes to remove transfected DNA produced in bacteria and linearize the MCV genome consecutively, after which 5 ng of MCV or 50 ng of GAPDH-digested DNA was used for quantification by qPCR.

4.2.5 Quantitation of MCV Genome Copy by Real-Time PCR

Quantitative PCR (qPCR) was performed using PowerUp™ SYBR™ Green master mix (cat. no. A25778; Thermo Fisher) together with primers MCV DNA Fw: 5'-AAAACACCCA AAAGGCAATG-3' and MCV DNA Rev: 5'-GCAGAGACACTCTTGCCACA-3' to quantify MCV genome copy numbers and GAPDH DNA Fw: 5'-TGTGTCCCTCAATATGGTCCTGTC-3' and GAPDH Rev: 5'-ATGGTGGTGAAGACGCCAGT-3' to amplify endogenous control GAPDH DNA. Thermal cycling was performed on a QuantStudio™ 3 Real-Time PCR machine. Threshold cycle (C_T) values were used to calculate DNA replication levels and were normalized to GAPDH. MCV genomic DNA replication levels were calculated according to the $\Delta\Delta C_T$ method.

4.2.6 Immunoblotting

Total protein was extracted by lysing cells in 1% SDS buffer (1% SDS, 10 mM Tris-HCl [pH 8.0], 1 mM EDTA [pH 8.0]) and subsequent sonication at 20% Amp for 5 s four times on ice. Protein concentration was then quantified using the DC Protein Assay Kit (cat. no. 5000116; Bio-Rad, Hercules, CA, USA). Then, 100 µg of total protein was separated by SDS-PAGE and transferred onto a nitrocellulose membrane. Membranes were then incubated with primary mouse monoclonal antibody to MCV LT and 57kT (CM2B4), MCV VP1 (CM9B2), MCV ALTO (CM7B1), MCV sT (CM5E1) or mScarlet (cat. no. 6g6-100; Chromotek, Planegg-Martinsried, Germany) followed by 1:10,000 dilution of IRD800 conjugated goat anti-mouse secondary antibody (cat. no. 926-32210; LI-COR Biotechnology, Lincoln, NE, USA) in combination with a 1:20,000 dilution of Rhodamine conjugated anti-Tubulin antibody (cat. no. 12004163; Bio-Rad,

Hercules, CA, USA). Signals were detected on a ChemiDoc imaging system (Bio-Rad, Hercules, CA, USA).

4.2.7 Immunofluorescence Assay (IFA)

Five days after transfection, U2OS or 293 cells grown on coverslips were fixed with 4% paraformaldehyde in PBS for 15 min at room temperature, permeabilized with 0.1% Triton X-100 in PBS for 15 min at room temperature and incubated in blocking solution (5% normal Goat serum: cat. no. 9023; Sigma-Aldrich, St. Louis, MO, USA) for 1 h at room temperature. Coverslips were then incubated with primary mouse antibody to MCV LT (CM2B4) or MCV VP1 (CM9B2) for 1 h at 37 °C in a humidified chamber and washed three times in 1× PBS for 5 min at room temperature followed by incubation with AF-488 conjugated goat anti-mouse secondary antibody (cat. no. 11006; Invitrogen) for 30 min at 37 °C. For LT-VP1 co-staining in infected cells, CM2B4 antibody was conjugated with AF-488 using Alexa Fluor 488 antibody labeling kit (cat. no. A20181; Invitrogen, Eugene, OR, USA) and coverslips were sequentially stained first with CM9B2 primary antibody for 1 h at 37 °C followed by AF-568 conjugated goat anti-mouse secondary antibody (cat. no. 11004; Invitrogen, Eugene, OR, USA) for 30 min at 37 °C and then with AF-488 conjugated CM2B4 primary antibody to LT for 1 h at 37 °C. All antibody incubations were performed in a humidified chamber in the dark, with three washes in between each incubation period. After the final wash, cells were stained with 300 nM of DAPI in 1× PBS for 5 min, washed three times and coverslips were mounted on a glass slide. Images were acquired using an Olympus AX70 microscope with a QImaging QIClick charge-coupled device (CCD) camera and Q-Capture Pro 7 software for qualitative anlysis or a Cytation5 Imaging reader and Gen5 image analysis software (BioTek, Santa Clara, CA, USA) for quantitative analysis.

4.2.8 Virion Production

All wild-type MCV genomes used in this study conform to the HF strain (GenBank accession no. JF813003) [138]. The following day, 293 TRE-sTco cells were seeded at a density of 5×10^6 cells per 10-cm dish and transfected with 10 µg of MCVmc or MCV-ligated DNA using Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA). One day after transfection, cells were passaged to two T75 flasks and treated with 500 ng/mL Dox on day 3 to induce sTco expression. Transfected cells were expanded from two T75 flasks to two or three T175 flasks in the presence of Dox treatment and harvested 10 days after transfection for MCV virion isolation. Cells were collected in PBS containing 9.5 mM MgCl2, 25 mM ammonium sulfate, 0.5% TritonX, 0.1% Benzonase, 1 mM ATP and 0.1% ATP-dependent DNase and incubated overnight at 37 °C to degrade unpackaged virus genomes. Nuclear fraction was then isolated by the addition of 720 mM of NaCl into the lysate. To purify MCV virion, supernatant from the lysate was overlayed on a 2.1 mL discontinuous Opti-Prep (iodixanol) (cat. no. D1556; Sigma-Aldrich, St. Louis, MO, USA) gradient (0.7 mL of 27%, 0.7 mL of 33% and 0.7 mL of 39%) and subjected to ultracentrifuge at 226,354× g for 2.5 h at 16 °C in an AH-650 swing rotor (Sorvall). Thirteen fractions (~400 µL) were collected from the top of the ultracentrifuge tube using a pipette and the fractions were used for immunoblotting, qPCR and infection assay.

4.2.9 Electron Microscopy

Suspensions of viruses were adhered to glow-discharged (Cressington 108Auto) Formvar-coated copper grids for 10 min. Extra solution was wicked away, then the samples were negatively stained with 1% aqueous uranyl acetate and solution was immediately wicked away. Samples were

imaged on a JEOL JEM 1400 Flash transmission electron microscope (JEOL, Peabody, MA, USA) fitted with a BIOSPR12 bottom mount AMT camera (Advanced Micrscopy Techniques, Danvers, MA, USA). Images were taken at 80 kV.

4.2.10 MCV Infection Assay

Fibroblast cells (BJ, BJ.hTert, HFF-1) and MSC cells (MSC-bm, MSC-a) at a density of 2 × 10^5 cells per well of a 6-well plate were infected with 4 × 10^5 genome copy equivalent of MCV-ligated or MCVmc virion per cell in infection media, F12/DMEM medium containing 20 ng/mL EGF (cat. no. 78006; Stem Cell Technologies, Vancouver, BC, Canada), 20 ng/mL bFGF (cat. no. 78003; Stem Cell Technologies, Vancouver, BC, Canada), 3 µM CHIR9901 (cat. no. S2924; Sellechem, Houston, TX, USA) and 0.025 mg/mL collagenase IV (cat. no. 17104019; Thermo Fisher, Waltham, MA, USA), and seeded in a 6-well plate. After 3 days of incubation with the infection media, FCS was added at a final concentration (volume/volume) of 20% and cells were incubated for another 2 days. After 5 days of infection, cells were seeded in a 48-well plate, fixed and processed for the following day, as described in the previous section. The remaining cells were passaged to 6-well plates and harvested after 10 days of infection for Southern hybridization.

For single round infection with VLP-packaged reporter viruses, 293 TRE-sTco cells were seeded at a density of at 2.5 × 10^5 cells per well in 12-well plates and infected with 2 × 10^3 genome copy equivalent of VLP-packaged MCV virions per cell in DMEM with no FBS. After 24 h, FBS was added at a concentration of 10% (volume/volume) and 500 ng/mL Dox was added to the medium after 3 days of infection. Starting at day 4 post infection, medium with Dox was replenished every day and imaging was performed at day 10 post infection.

4.2.11 Southern Blot

MCV-ligated or MCVmc-infected cells harvested after 10 days of infection were suspended in genomic DNA lysis buffer (10 mM Tris, 25 mM EDTA, 0.5% SDS, 100 mM NaCl) containing 0.1 mg/mL proteinase K and incubated overnight at 37 °C. The lysate was treated with 0.1 mg/mL RNase A for 1 h at 37 °C and genomic DNA was extracted by phenol-chloroform and ethanol precipitation. Genomic DNA (3 µg) was then digested with DpnI and EcoRI overnight and subsequently electrophoresed in a 0.8% agarose TAE gel. The DNA in the agarose gel was depurinated in 0.25 N HCl, treated with denaturation buffer (0.5 M NaOH, 1.5 M NaCl) and capillary-transferred onto a Hybond-N^+ Nylon Membrane (cat. no. RPN303B; GE Healthcare, Chicago, IL, USA) overnight with 10× SSC (1.5 M NaCl, 150 mM Sodium Citrate [pH 7.0]). After UV crosslinking (Stratalinker, Stratagene, San Diego, CA, USA), the membrane was treated with prehybridization buffer (5× SSPE [diluted from 20× solution containing 3 M NaCl, 0.2 M sodium phosphate monobasic monohydrate, 0.2 M EDTA], 2% SDS, 1× Denhardts [diluted from 50× Denhardts, cat. no. 750018; Thermofisher, Waltham, MA, USA], 10% dextran sulfate, sonicated salmon sperm (10 µg/mL) for 4 h at 65 °C. A denatured, biotin-11dUTP-labeled MCV DNA probe generated with the Bio-prime Array CGH kit (cat. no. 45-0048; Invitrogen, Waltham, MA, USA) using a full-length MCV genome as template was then added to the prehybridization buffer and incubated overnight at 65 °C. After washing twice with 2× SSPE and once with 0.1× SSPE at 60 °C, membranes were incubated for 1 h with IR800 Dye-conjugated streptavidin (cat. no. 926-32230; Li-COR, Lincoln, NE, USA), and the hybridization signal was detected with a Li-COR Odyssey Infrared Imaging System (Li-COR, Lincoln, NE, USA).

4.2.12 MCV Kinetics Assay

For the MCV kinetics assay, 293 cells were plated and transfected with MCVmc, MCVmc.VP1-mS or MCVmc.VP1-P2A-mS with FuGENE was as described before. The medium was refreshed every day and cells were collected from day 0 through day 10 post-transfection for analysis by immunoblotting (see previous section) or flowcytometry.

4.2.13 Flow Cytometry

Harvested cells were fixed in 4% PFA on ice for 15 min, stored in PBS with 0.02% NaN_3 at 4 °C protected from light and were directly used to quantitate mScarlet-positive cells by flow cytometry on a BD LSR Fortessa cell analyzer (BD Biosciences, San Jose, CA, USA).

4.2.14 MCV Packaging Assay

To assay for the packaging efficiency of MCV fluorescent reporter viruses, cell lysate from 293 TRE-sTco cells that were transfected and harvested as described in virion production were overlayed on 9 mL discontinuous gradient (3 mL of 27%, 3 mL of 33%, 2 mL of 39% and 1 mL of 60% iodixanol) and subjected to ultracentrifuge at 41,000 rpm for 6 h at 16 °C in SW 41 Ti swing-bucket rotor (Beckman, Indianapolis, IN, USA). Twelve fractions (1 mL each) were collected from the top of the tube using a pipette. Fractions were used to perform immunoblotting, qPCR (in the presence or absence of Benzonase treatment) and an infection assay.

To assay for the packaging efficiency of the pseudovirus system, 293 TRE-sTco cells were co-transfected with 18 µg MCV genome DNA or reporter plasmid pEGFP-N1, 15 µg pWM

plasmid (expressing VP1) and 5 µg ph2m plasmid (expressing VP2) in T75 flasks. Three days after transfection, cells were harvested and virions were purified as described above.

4.3 Results

4.3.1 MCV Genome Recircularization by Site-Specific Recombination

We generated an MCV molecular clone using site-specific recombination, i.e., mini-circle (mc) technology [184] (**Figure 23A, B, and C**) that has been used to produce HPV [403, 404] and HBV [440] mc genomes. The MCV-HF genome (GenBank accession no. JF813003) [38] is linearized at nucleotide position 3146/3147 and cloned between the attB and attP sites in the pMC.BESPX vector to generate the parental pMC-MCV plasmid (**Figure 23B**). Transformation into *Escherishia coli* strain ZYCY10P3S2T allows arabinose-induction of bacteriophage ΦC31 integrase [184] that mediates recombination between the attB and attP sites, while I-SceI endonuclease, which is also induced by arabinose, digests the excised bacterial backbone containing 32 I-SceI recognition sites (**Figure 23B**). This produces covalently closed circular MCV minicircle (MCVmc) genomes, each with a 39 bp remnant scar sequence from the recombination (**Figure 23B and C**). Comparison of T4 DNA ligase (**Figure 23A**) and ΦC31 integrase (**Figure 23C**) recircularization shows that the minicircle system is not only more efficient than T4 DNA ligase-mediated recircularization but also generates a single copy of the MCV genome, while ligation generates partial and multimeric forms, including forms with a vector backbone that might interfere with virus replication studies.

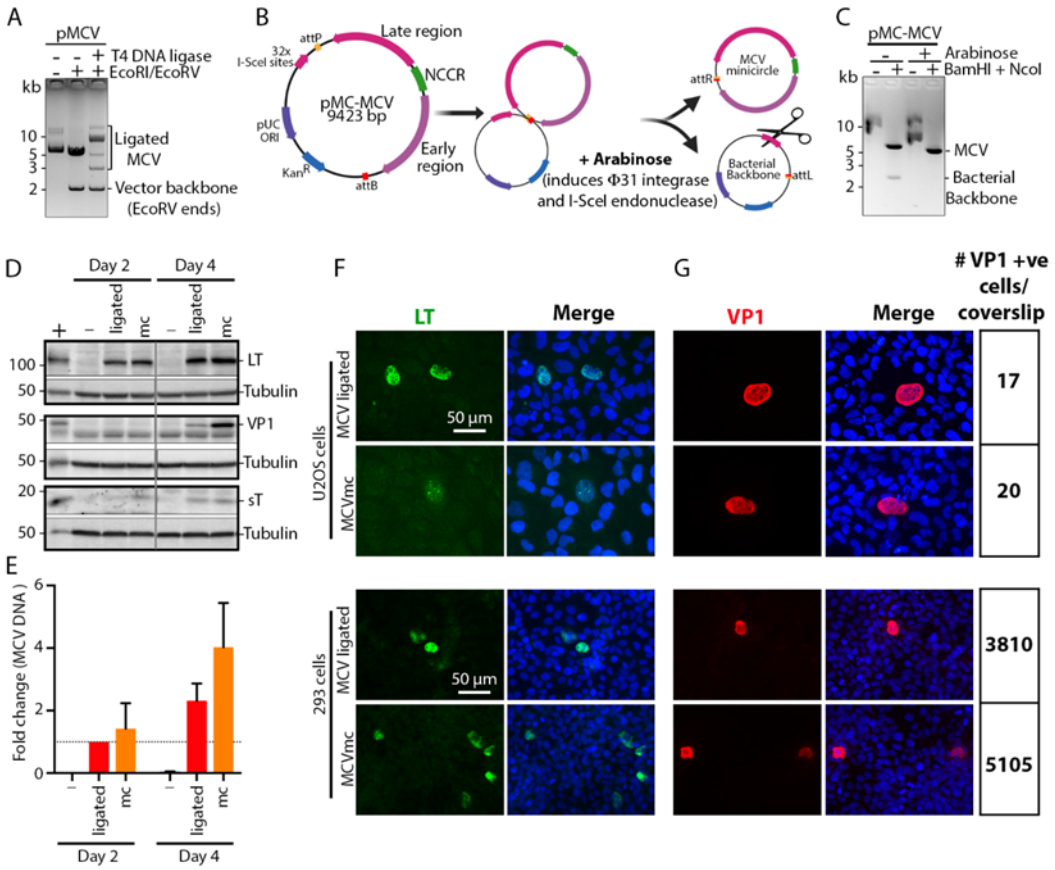


Figure 23 Development of MCV molecular clone using minicircle technology.

(A) Electrophoresis of MCV molecular clone re-circularized by T4 DNA-ligase. DNA size markers are shown at the left of the gel image. (B) Schematic of recombinase-mediated recircularization (minicircle technology) of MCV molecular clone (MCVmc). (C) Electrophoresis of MCV DNA extracted before and after recombination. DNA size markers are indicated at the left of the gel image. (D) Western blot of MCV-encoded proteins (LT, sT, and VP1) from MCV ligated and MCVmc transfected 293 cells 2- and 4-day post-transfection. Un-transfected 293 cells were used as a negative control. LT and VP1 expression construct transfected in 293 cells were used as positive control (+), while α-tubulin was used as an endogenous protein control. Protein molecular weight markers are shown at left. The result is representative of three independent experiments. (E) Real-time PCR (qPCR) of DpnI-resistant MCV genome from MCV ligated or MCVmc transfected 293 cells 2- and 4-day post-transfection. Un-transfected 293 cells were used as negative control (-). ΔΔC_T method was used to determine MCV genome copy number fold change; GAPDH was used as endogenous control while MCV ligated 2-day post-transfection group was used as the experimental control.

Error bars indicate the ± SD of three independent experiments. (F and G) Immunofluorescence of LT-AF488 (pseudo color green), VP1-AF488 (pseudo color red), and DAPI (blue) in MCV ligated and MCVmc transfected 293 or U2OS cells 5-day post-transfection. Number of VP1 positive cells per coverslips were counted by Cytation 5 cell imaging multi-mode reader and shown on the right. Images were originally acquired at 40× magnification.

4.3.2 MCVmc Gene Expression

MCVmc and *in vitro* ligated MCV clones were compared for viral gene expression and replication capacity in 293 cells. With equal amounts of transfected genomic DNA, MCVmc expresses the LT protein at a comparable level to MCV-ligated DNA (**Figure 23D**). Expression of sT was slightly higher for MCVmc compared to MCV-ligated at 2-day post-transfection. However, this difference in sT levels was minimal at 4-day post-transfection (**Figure 23D**). The VP1 late gene product was detectable only at 4-day post-transfection in both MCVmc and MCV-ligated-transfected 293 cells. The level of VP1 was slightly higher in MCVmc transfected cells compared to MCV-ligated-transfected cells (**Figure 23D**). Consistent with immunoblot results, MCVmc genomic DNA replicated at a higher rate compared to the MCV-ligated genome, as shown by qPCR of DpnI-resistant DNA (**Figure 23E**); comparable or higher expression for MCVmc was confirmed by immunofluorescence using antibodies to LT (**Figure 23F**) and VP1 (**Figure 23G**) on transfected 293 or U2OS cells, although transfection efficiency and viral gene expression was markedly lower for U2OS cells (**Figure 23G**). Notably, expression of VP1 protein, which serves as a surrogate marker for the complete cycle progression of DNA replication [119, 120] in the MCVmc system, indicates that the "scar" of 39 additional non-MCV nucleotides extra sequences at the recombinase site does not interfere with MCV replication.

4.3.3 MCVmc *In Vitro* Transmission

sT is an important replication accessory protein that stabilizes and increases LT accumulation. To maximize virion production, we transfected MCVmc or MCV-ligated into 293 TRE-sTco cells with a stably integrated, doxycycline (Dox)-inducible, codon-optimized MCV sT cassette and purified cell lysates on OptiPrep (iodixanol) gradients (**Figure 24A**). VP1-positive fractions 10 through 13 were pooled and 1.0×10^{10} genome copy equivalent of viral particles were used to infect a panel of primary cell lines (**Figure 24B-E**). Human foreskin fibroblast (HFF-1), BJ-hTert, bone marrow-derived human mesenchymal stem cells (MSC-bm) and adipose-derived MSCs (MSC-a) show infection with both MCV-ligated and MCVmc. MCVmc virions had reduced infection compared to MCV-ligated as measured by an IFA of LT and VP1 expression (**Figure 24B, C, and E**). This may be due to the incomplete digestion of unpackaged DNA that interferes with accurate MCV copy quantification or a difference in DNA backbones. This was confirmed by Southern blotting for MCV genome DNA in BJ-hTert cells (**Figure 24D**).

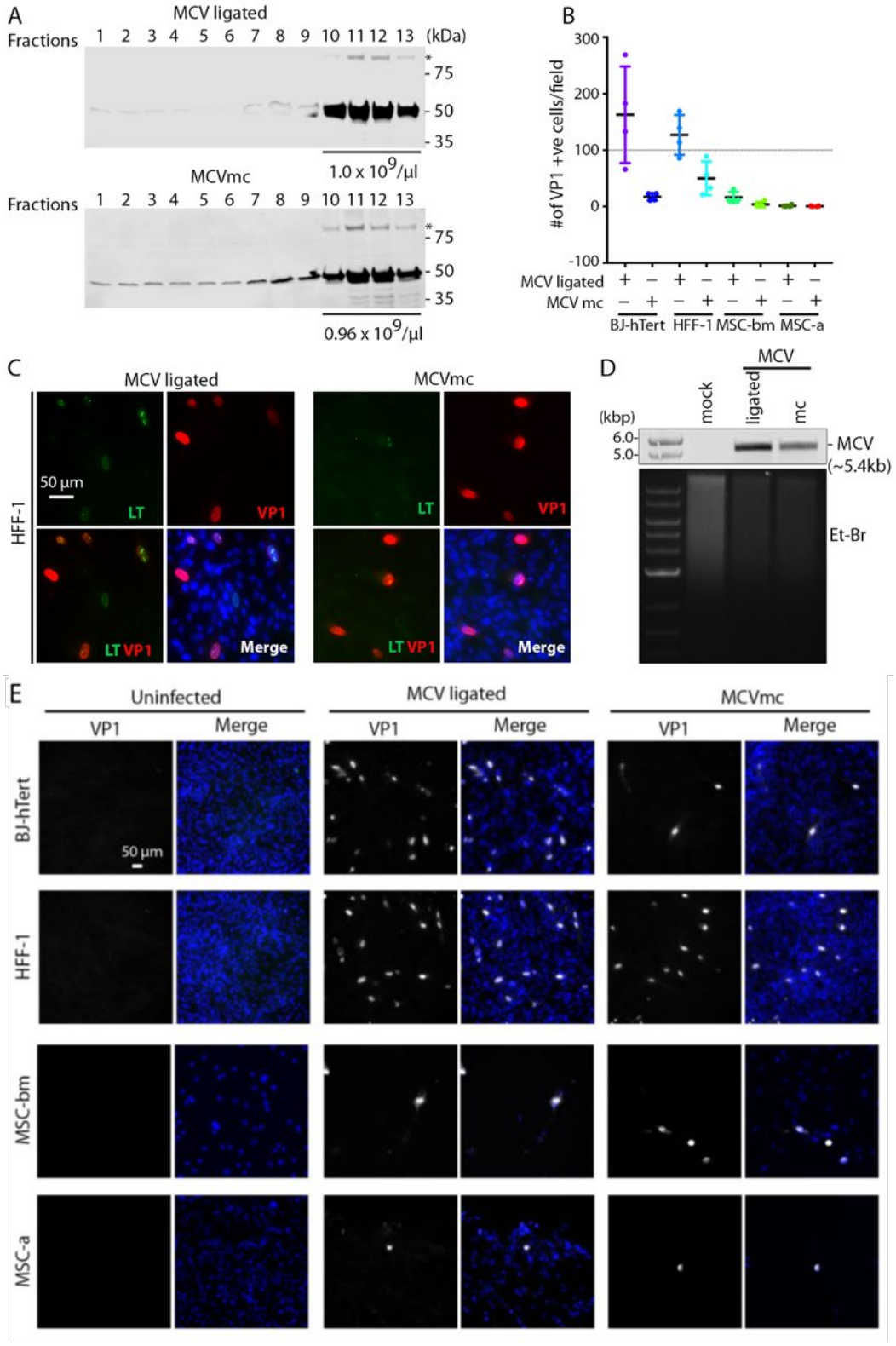


Figure 24 MCVmc is infectious in primary cells.

(A) Western blot of MCV-ligated and MCVmc virions fractions from Opti-Prep (iodixanol) gradient purification. Fractions 1-13 (top to bottom) are indicated. * Indicates VP1 dimer. Fractions 10 through 13 were mixed and MCV genome copy numbers per µL are shown at the bottom of virion-positive fractions. Protein molecular weight markers are shown on the right. Results are representative of three independent experiments. (B) Quantification of the number of VP1-positive cells per 4× magnification field in BJ-hTert, HFF-1, MSC-bm and MSC-a cells infected with MCV-ligated or MCVmc virus from (A), 6 days post infection, by IFA of VP1. Four fields from one cover slip for each cell line were counted for MCVmc and MCV-ligated. (C) Representative immunofluorescence images showing LT-AF488 (pseudo color green), VP1-AF568 (pseudo color red), together with DAPI (blue) in MCV-ligated and MCVmc-infected HFF-1 cells 6 days post infection. Images were originally acquired at 40× magnification. Images (C) were acquired and the number of VP1 positive cells (B) was counted using a Cytation 5 cell imaging multi-mode reader. Results represent three independent experiments. (D) Southern blot of MCV genome from BJ-hTert cells infected with MCV-ligated or MCVmc. Mock-infected cells were used as a negative control. The relative amount of total DNA loaded is shown by EtBr staining and DNA size markers are shown on the left. (E) Immunofluorescence of VP1-AF488 (white) in MCV ligated and MCVmc infected BJ-hTert, HFF-1, MSC-bm, or MSC-a cells 6-day post-infection. Images were originally acquired at 40× magnification. DAPI counter staining is shown in blue. Results represent three independent experiments.

4.3.4 Mutagenesis of the MCVmc genome and cell-specific effects on replication

In a proof-of-principle experiment, previously reported mutations known to affect MCV replication were introduced into the parental pMC-MCV plasmid (**Table 17**). The MCVmc mutants generated include MCVmc-Rep^{-} which has a single nucleotide (C44A) mutation in the NCCR viral origin at a LT-binding pentanucleotide that abrogates viral genome replication [138, 146]. Mutations to eliminate LT binding to hVam6p (MCV-hVam6p^{-}; LT W209A) [138, 194] and interactions with the Skip-Cul-Fbox (SCF) E3 ligases β-TrCP, Fbw7 and Skp2 [147] were also

made. Additionally, we generated MCVmc.hpko, a virus with mutations to the hairpin loop required to produce miR-M1 [76] (**Figure 25A**).

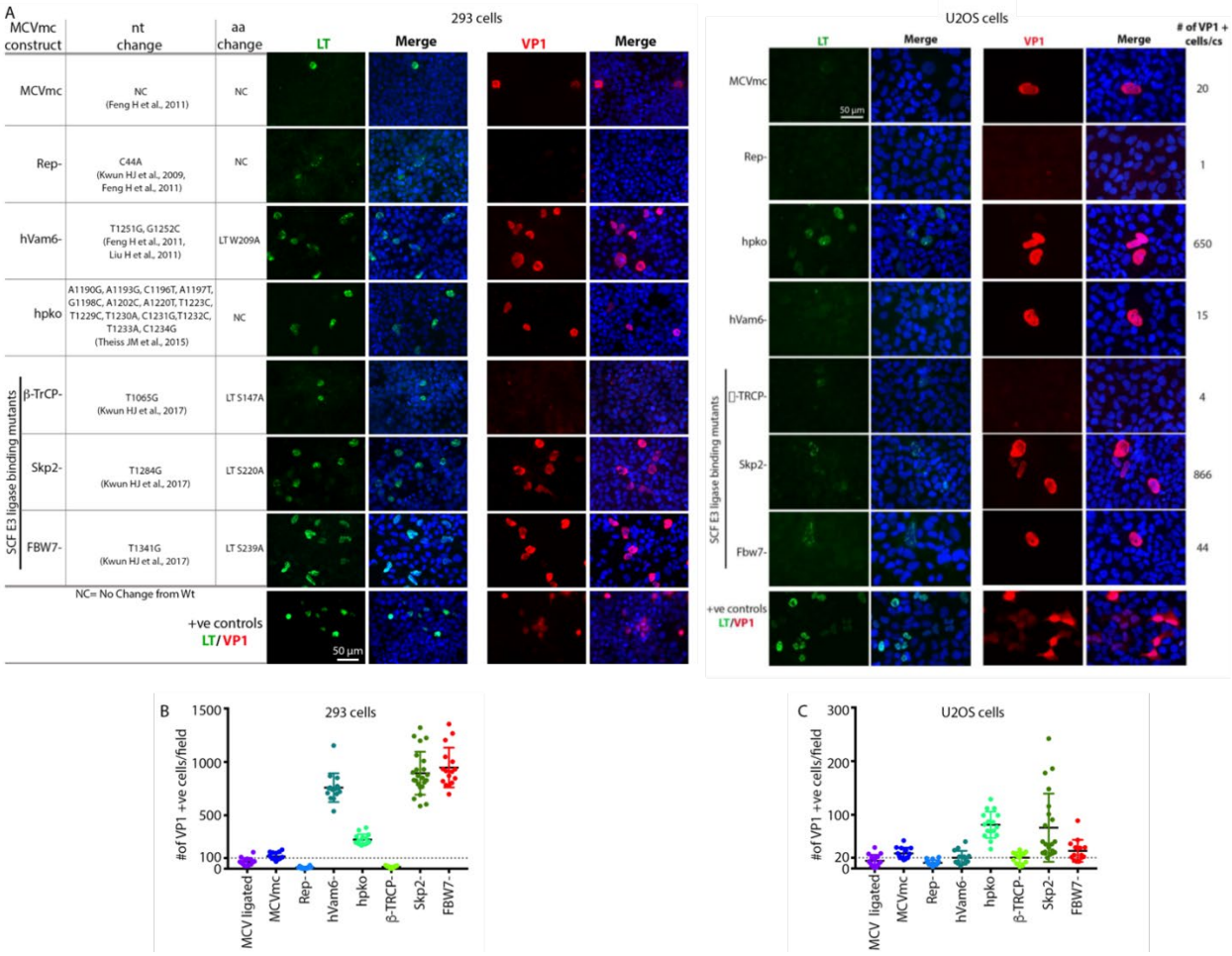


Figure 25 MCVmc is amenable to viral gene mutational analysis.

(A) The left panel shows a table of tested MCV mutants; nucleotide (nt) and amino acid (aa) changes (based on HF strain: GenBank #JF813003) are noted for each mutant [76, 138, 146, 147, 194]. The right 2 panels show immunofluorescence of LT-AF488 (pseudo color green) or VP1-AF488 (pseudo color red) and DAPI (blue) in 293 and U2OS cells transfected with MCVmc or mutants 5-day post-transfection. LT- or VP1-expression construct transfected 293 cells were used as positive controls. Images were originally acquired at 40× magnification. (B and C) Quantification of the number of VP1 positive cells per 4× magnification field in 293 or U2OS cells transfected with MCVmc or mutants 5-day post-transfection. The number of VP1 positive cells was quantified using a Cytation 5 cell imaging multi-mode reader.

Transfection of various mutated MCVmcs (**Table 17**) into either 293 or U2OS cells allowed both qualitative (**Figure 25A**) and quantitative (**Figure 25B and C**) assessment of replication permissivity as compared to wild-type (wt) MCVmcs. LT and VP1 positive cells displayed nuclei that are 2-5-fold larger than uninfected cells with condensed LT nuclear puncta consistent with viral replication centers, particularly in U2OS cells (**Figure 25A**). T4 DNA ligase-treated and MCVmc wt genomes showed similar low-level replication (measured by VP1 expression) that was, nevertheless, significantly greater than the MCVmc-Rep^{-} mutant. Mutation of the LT β-TrCP phosphodegron site (S147A) required for replication also did not show VP1 expression, indicative of replication loss, as previously reported [147]. Elimination of the Fbw7 interaction site markedly increased MCV replication in 293 cells but not U2OS cells (**Figure 25B and C**), a pattern similar to that seen for the Vam6p binding site mutant while the Skp2 phosphodegron mutant increased MCV replication in both cell types. MCVmc carrying mutations in the miR-M1 hairpin markedly enhanced MCV replication independent of the cell type.

4.3.5 Generation of Fluorescent MCV Reporter Viruses (MCVmc.VP1-mS and MCVmc.VP1-P2A-mS)

To directly visualize the viral replication process in real time, we fused an mScarlet fluorescent protein coding sequence 3' to the MCV VP1 protein-coding sequence (MCVmc.VP1-mS) to generate a VP1-mScarlet fusion protein (**Figure 26A**). In a second reporter virus (MCVmc.VP1-P2A-mS), the porcine teschovirus-1 2A peptide sequence (P2A: GSGATNFSLLKQAGDVEENPGP) that initiates ribosome skipping during translation [441] was engineered between VP1 and mScarlet proteins to minimize functional consequences of the 232

amino acid fusion tag. Nevertheless, MCVmc.VP1-P2A-mS still carries 21 extra non-MCV amino acids on the C-terminus of VP1.

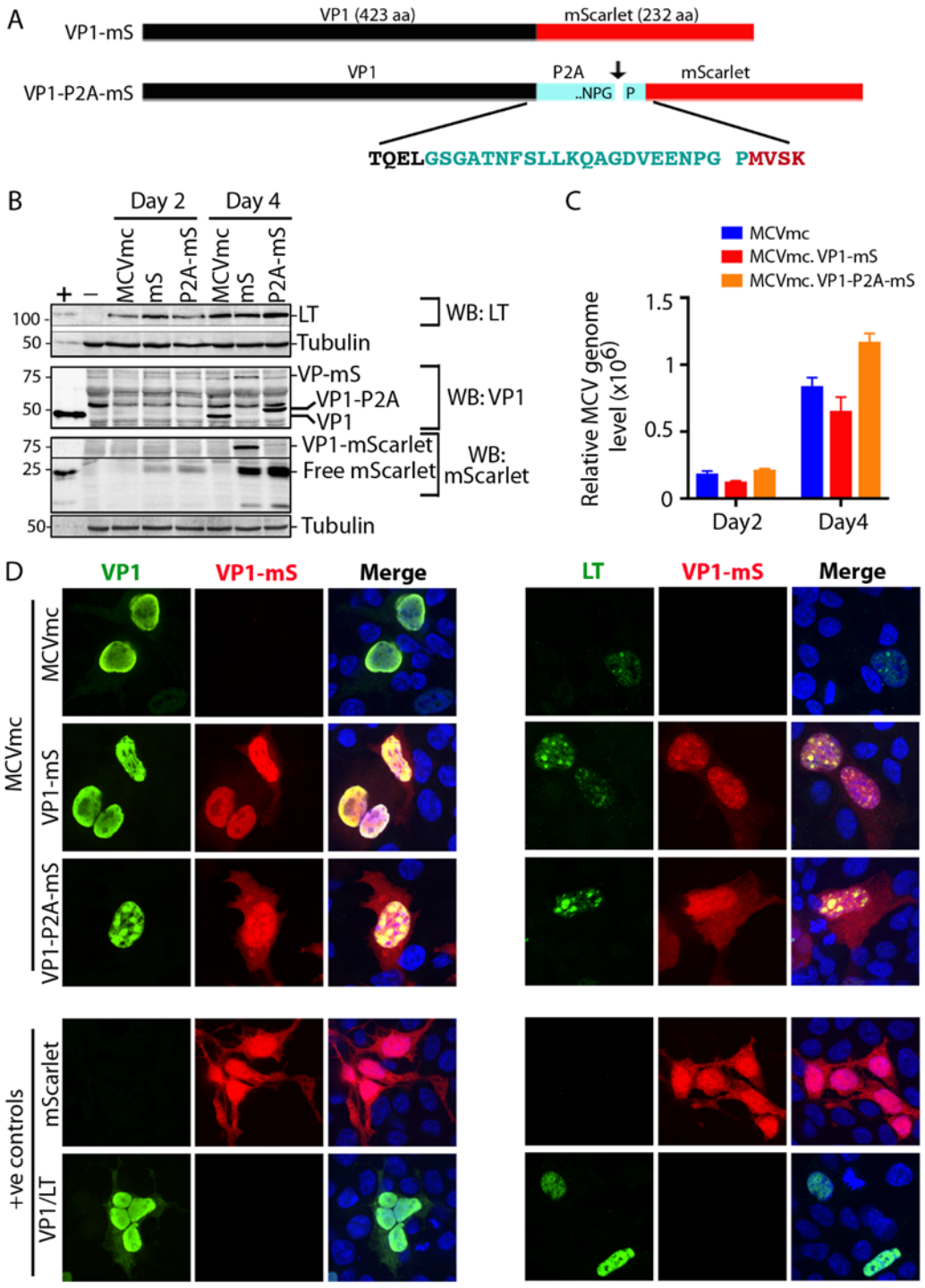


Figure 26 Generation of an mScarlet reporter MCVmc.

(A) Schematic of VP1 fluorescent fusion constructs representing VP1 (black) containing mScarlet (red) tag in the presence (cyan) or absence of a P2A linker. The arrow indicates the P2A ribosome-skipping site and the whole P2A peptide sequence (cyan) is shown with four flanking amino acids from VP1 (black) and mScarlet (red). (B) Immunoblot of MCV-encoded proteins (LT and VP1) and mScarlet in 293 cells transfected with MCVmc, MCVmc.VP1-mS and MCVmc.VP1-P2A-mS after 2- and 4-day post-transfection. Un-transfected 293 cells were used as a negative control (−), while 293 cells transfected with LT or VP1 expression construct were used as a positive control (+). α-tubulin was used as an endogenous protein-loading control. Protein molecular weight markers are shown on the left. (C) Quantification of DpnI-resistant replicated MCV DNA in 293 cells transfected with MCVmc, MCVmc.VP1-mS or MCVmc.VP1-P2A-mS 2- and 4-day post-transfection qPCR results. The $\Delta\Delta C_T$ method was used to calculate relative MCV DNA levels; GAPDH was used as the endogenous loading control, while MCVmc was used as the experimental control. Error bars indicate ± SD of three independent replicates. (D) Confocal images of mScarlet expression (pseudo color red) and LT-AF488 (pseudo color green) or VP1-AF488 (pseudo color Green) immunofluorescence in MCVmc, MCVmc.VP1-mS and MCVmc.VP1-P2A-mS-transfected 293 cells 5-day post-transfection. LT, VP1 or mScarlet expression construct-transfected 293 cells were used as positive controls. Images were originally acquired at ×40 magnification.

293 cells were transfected with these MCVmc reporter genomes and analyzed by immunoblotting and qPCR. Both MCVmc.VP1-mS and MCVmc.VP1-P2A-mS express LT at comparable levels to MCVmc at days 2- and 4-day post-transfection (**Figure 26B**). As expected, VP1 (~50 kDa) shifted to higher molecular masses for MCVmc.VP1-mS and MCVmc.VP1-P2A-mS as a result of the residual fusion peptide/protein. DpnI-resistant DNA detection (**Figure 26C**) showed comparable increases in MCV genome copy numbers for MCVmc and the reporter viruses, although the mScarlet fusion protein virus (MCVmc.VP1-mS) had consistently reduced replication. Using a primary mouse anti-VP1 followed by a AF488-conjugated secondary anti-mouse antibody, dual color IF of transfected cells showed complete nuclear co-localization of mScarlet (red) with VP1 (green) for the MCVmc.VP1-mS fusion virus (**Figure 26D, left panel**).

In contrast, cells transfected with the MCVmc.VP1-P2A-mS virus showed mScarlet staining also in the cytoplasm, consistent with a cleaved fluorescent tag. Comparison of LT and mScarlet expression revealed discrete localization of LT protein at putative nuclear replication puncta (**Figure 26D, right panel**).

4.3.6 Replication Kinetics for MCV Reporter Viruses.

Since sT is involved in MCV replication, MCVmc, MCVmc.VP1-mS or MCVmc.VP1-P2A-mS were transfected into wild type 293 cells instead of 293 TRE-sT cells and monitored by flow cytometry and immunoblotting to analyze MCV replication kinetics (**Figure 27A and B**). mScarlet expression from both fusion viruses was evident as early as 2-day post-transfection and the number of mScarlet-positive cells increased through 10-day post-transfection. Consistent with data from **Figure 26**, MCVmc.VP-P2A-mS showed a growth advantage as measured by the number of positive cells compared to MCVmc.VP1-mS. All three viruses showed the highest LT protein expression at days 3-6 which declined at later time points (days 8-10). VP1, in contrast, peaked at days 5-6 (**Figure 27B**).

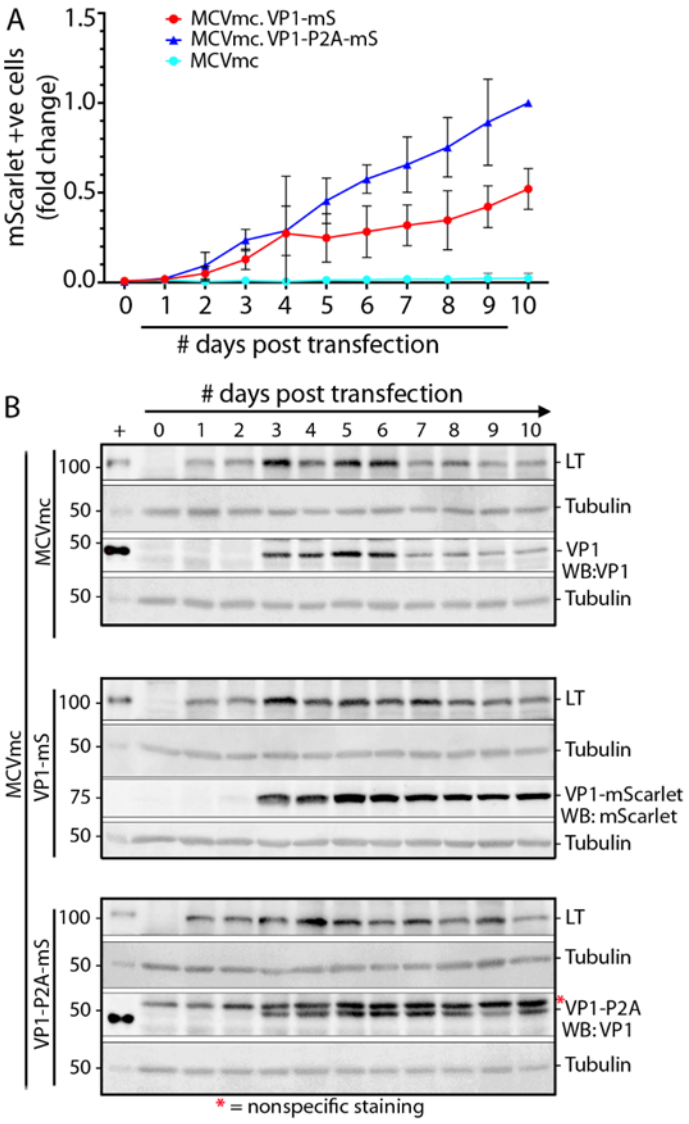


Figure 27 MCV replication kinetics analysis using mScarlet reporter.

(A) Numbers of mScarlet-positive cells in 293 cells transfected with MCVmc, MCVmc.VP1-mS, or MCVmc.VP1-P2A-mS from day 1 through day 10 post-transfection were quantified by flow cytometry. Data were normalized to MCVmc.VP1-P2A-mS day 10 post-transfection. Error bars indicate ± SD of three independent experiments. (B) Immunoblot of MCV-encoded proteins (LT and VP1) or mScarlet in 293 cells transfected with MCVmc, MCVmc.VP1-mS, and MCVmc.VP1-P2A-mS from 1 to 10 days post-transfection. LT or VP1 expression construct-transfected 293 cells were used as a positive control (+); while α-tubulin was used as an endogenous protein-loading control. Protein molecular weight markers are shown on the left. Results are representative of three independent experiments. *Indicates non-specific staining.

4.3.7 Single-Round Transmission of MCV mScarlet Reporter Viruses

MCVmc.VP1-mS and MCVmc.VP1-P2A-mS DNAs were transfected into 293 cells, harvested after 10 days, lysed and purified on a 27-60% discontinuous OptiPrep gradient. Collected fractions were treated with Benzonase or left untreated before MCV DNA quantification by qPCR (**Figure 28A**). All preparations showed a peak of Benzonase-sensitive DNA at fractions 3-4, consistent with unencapsidated DNA in protein aggregates [442]. MCVmc.VP1 banding was present in fractions 7-10 that co-migrated with the Benzonase-protected MCV genome, consistent with full encapsidation. Neither MCVmc.VP1-mS nor MCVmc.VP1-P2A-mS generated Benzonase-protected MCV DNA. For MCVmc.VP1-P2A-mS, VP1 was present in dense fractions, consistent with multimerization without generation of complete capsid.

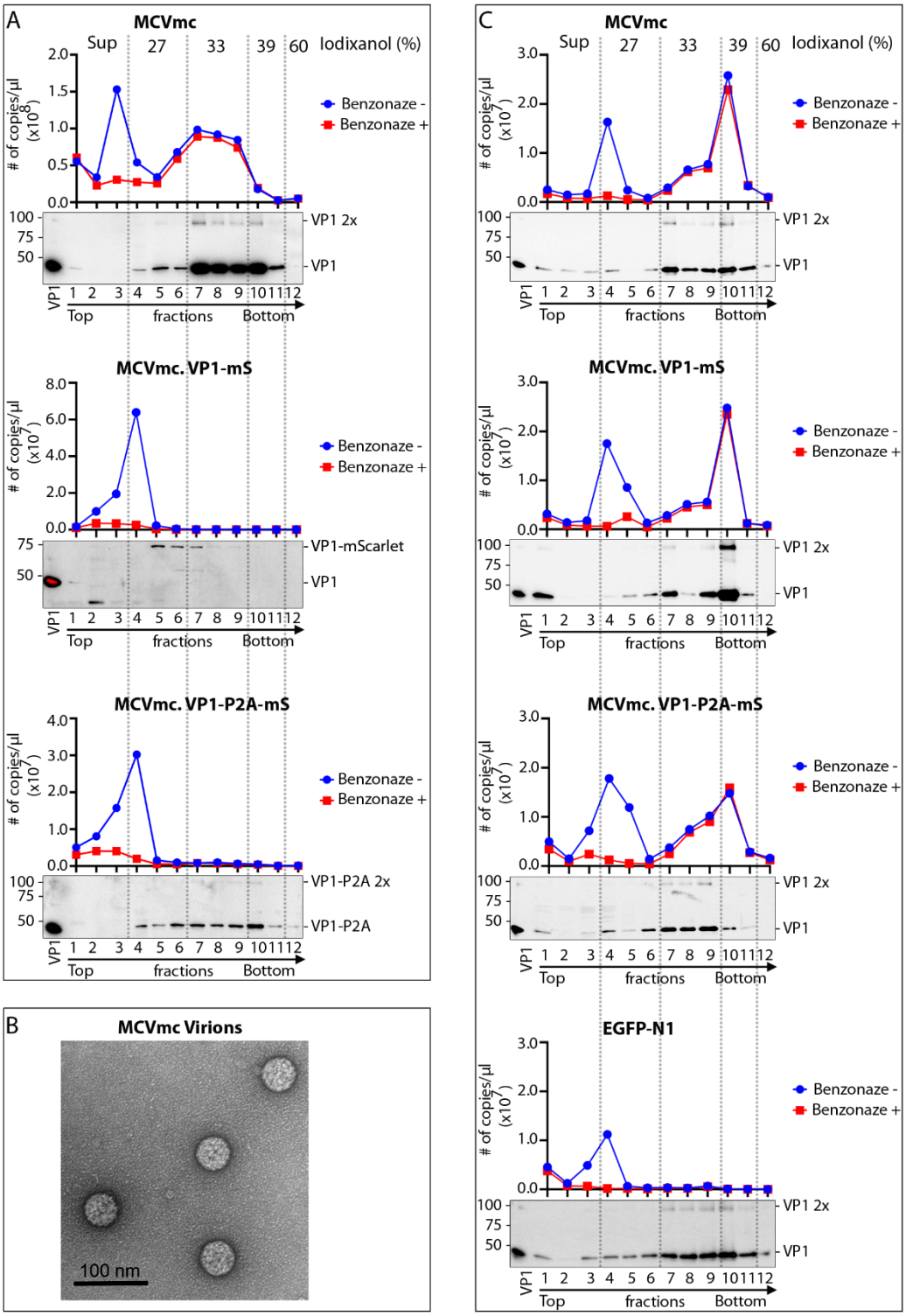


Figure 28 Virus production and infection using MCV mScarlet reporter virus.

(A) Western blot and qPCR analysis of fractions from MCVmc, MCVmc.VP1-mS, and MCVmc.VP1-P2A-mS virions purified over an Opti-Prep gradient (Iodixanol) (concentrations noted on top). In the qPCR

quantification, the blue curve shows total MCV DNA copy number, while the red curve shows Benzonase-protected MCV DNA copy numbers. Results represent two independent experiments. (B) Image of MCVmc virions, from panel (A), by negative staining electron microscopy. (C) Western blot and qPCR analysis in Opti-Prep fractions from heterologous VP1/VP2 packaged MCVmc, MCVmc.VP1-mS, and MCVmc.VP1-P2A-mS. Results represent one-time experiments. Iodixanol concentration for each fraction is noted on top. qPCR quantification: blue curves show total MCV DNA copy number, while red curves show the number of Benzonase-resistant MCV DNA copies.

To generate encapsidated MCV mScarlet reporter viruses, heterologous MCV genome was co-transfected together with expression plasmids for the native VP1 and VP2 proteins at a 3:1 ratio and tested for packaging efficiency in parallel with MCVmc.

Complementation of VP1 and VP2 proteins in trans rescued encapsidation of both MCVmc.VP1-mS and MCVmc.VP1-P2A-mS genomes, as evidenced by the appearance of MCV genomic DNA in the VP1-enriched heavy fraction corresponding to encapsidated genomes (**Figure 28C**). Infection using 5×10^8 genome copies (MOI 2×10^3) in 293 TRE-sTco cells generated a small number of mScarlet-expressing cells 8-10 days after infection, representing a single round of infection using the pseudovirus-packaged fluorescent-encoded MCV genomes (MCVmc.VP1-mS: 57 cells/well; MCVmc.VP1-P2A-mS: 62 cells/well; and EGFP-N1 (control): 2063 cells/well) (**Figure 29A and B**).

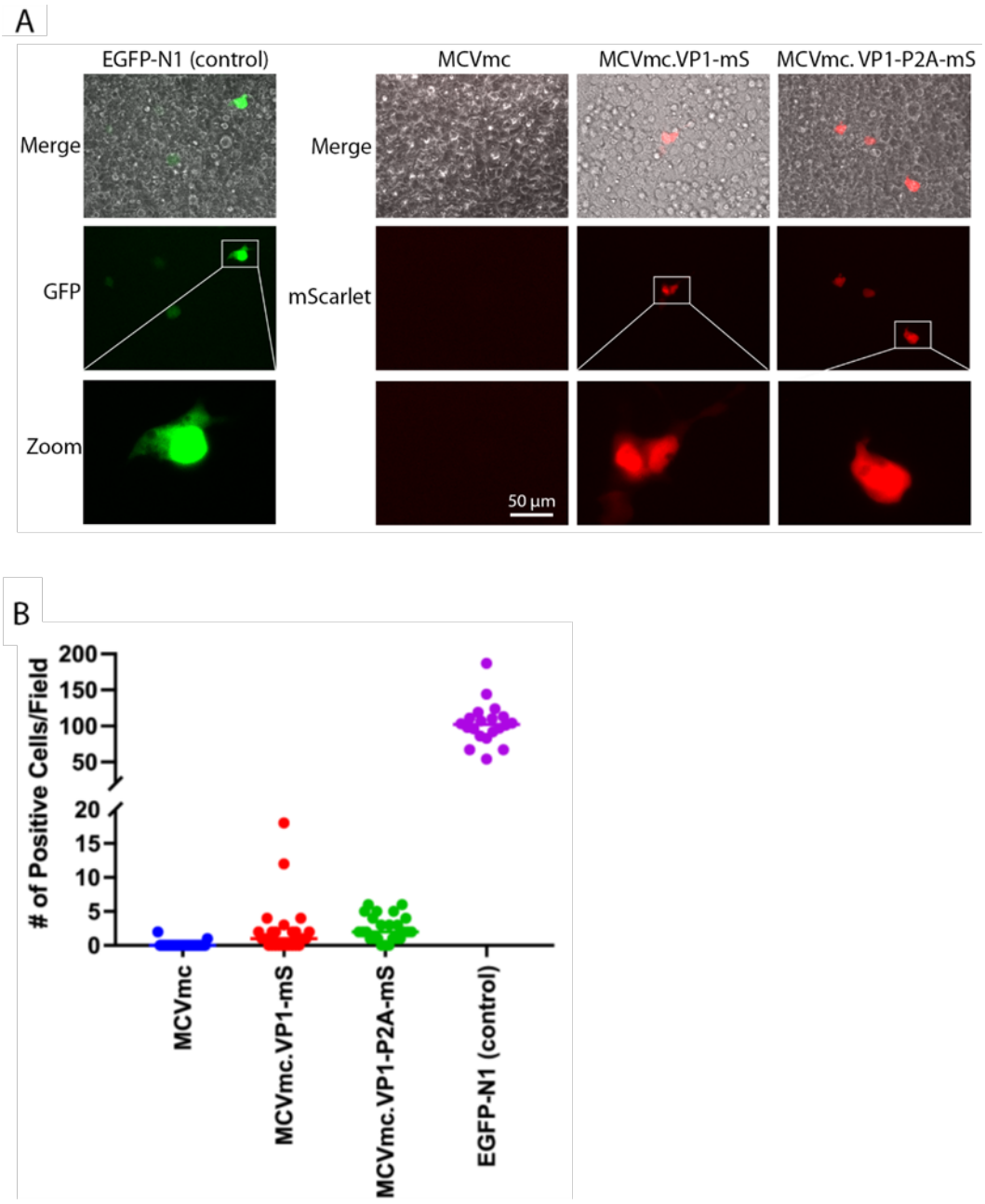


Figure 29 Single round infection by pseudovirus packaged MCV reporter.

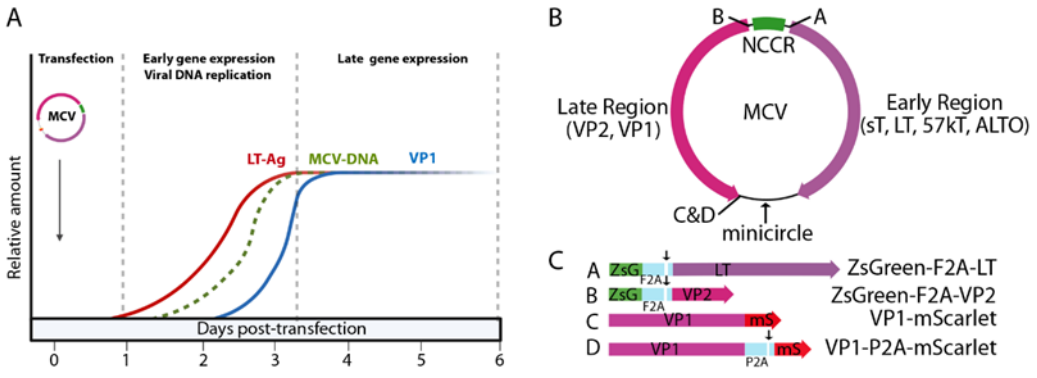
(A) Images represent 293 TRE-sTco cells infected with exogenous VP1/VP2-packaged pEGFP-N1 (pseudo color green), MCVmc, MCVmc.VP1-mS (pseudo color red) or MCVmc.VP1-P2A-mS (red) reporter. Images were originally acquired at 40× magnification. (B) The numbers of mScarlet and EGFP positive cells, from panel (A), were counted using Cytation 5 cell imaging multi-mode reader. Each dot represents the numbers of mScarlet or EGFP positive 293 cells per field 10-day post-infection.

4.4 Discussion

To gain a better understanding of MCV biology, we generated an MCVmc that enabled the production of a covalently closed circular genome free of bacterial sequences and amenable to genetic manipulation (**Figure 23**). Use of MCVmc improves experimental reproducibility due to the presence of only a single copy of the viral genome per minicircle in contrast to *in vitro* ligated MCV genomes (**Figure 23A**) that can exist as concatenated forms produced by *in vitro* ligation. In addition to the ease of minicircle production, we find that MCVmc replication and gene expression is similar to *in vitro* ligated MCV genomes, underscoring the validity of this system for studying the MCV lifecycle.

Minicircle technology has been used for stable gene expression and circularization of small DNA virus genomes, such as HBV and HPV free from bacterial sequences [403, 404, 440]. Limitations in the system described here derive from the fact that MCV is the smallest genome generated using this recombination approach. The constraints on placing recombination sites and fluorescent protein insertions into the MCV genome without eliminating viral viability required multiple trial and error cloning attempts (**Figure 30**). The recombination site leaves an extra 39 bp sequence and careful selection of its location was required to retain virus functionality. As shown in **Figure 30**, introduction of exogenous sequences into MCV genes can alter protein expression and genome packaging. While the recombinant MCV minicircle is a valuable new technology for defining MCV biology, careful comparisons are required to assess potential changes to the virus and its lifecycle caused by exogenous sequences. We do not find evidence that the recombination site we selected between the C-termini of the VP1 and LT genes and the "scar" of extraneous sequences affects virus replication or gene expression, but this remains a potential caveat to this approach. We find that introduction of the MCV350 NCCR mutation (Rep^{-}) [146] eliminates late

protein VP1 expression, consistent with newly replicated viral DNA being required for VP1 protein expression. The β-TrCP phosphorylation site at LT aa 149 is similarly required for replication and late gene, but not early gene, expression (**Figure 25A**). Conversely, introduction of mutations that abolished restriction factors for MCV replication, including LT-binding sites for Vam6p [194], interaction residues to SCF E3-ligases Fbw7 and Skp2 [147] or the MCV miR-M1 hairpin [76] markedly increase VP1 expression. Of note, some mutants displayed cell-type-dependent effects. Whereas Fbw7 and hVam6p mutants show these host cell proteins to be replication restriction factors in 293 cells, they are neutral in U2OS cells (**Figure 25**). Different cell lines are characterized by different expression programs and profiles, which can greatly affect many biological processes. Similar to MCCs, 293 cells are of neuroendocrine origin, expressing neurofilaments and cytokeratin [443, 444], whereas U2OS are of osteosarcoma origin. It is possible that variation in the expression of specific host cell proteins between these cell lines may contribute to the cell-type-specific effects we observed.



D

	Constructs	Fluorescent Tag	Insertion Site	Protein Expression		Virus Replication	Genome Pakaging
				LT	VP1		
A	MCVmc. ZsG-F2A-LT	ZsGreen-F2A	N-terminus of LT	+	-	+	-
B	MCVmc. ZsG-F2A-VP2	ZsGreen-F2A	N-terminus of VP2	+	-	+	-
C	MCVmc. VP1-mS	mScarlet	C-terminus of VP1	+	+	+	-
D	MCVmc. VP1-P2A-mS	P2A-mScarlet	C-terminus of VP1	+	+	+	-

Figure 30 Kinetics of MCVmc reporter viruses.

(A) A model of MCV replication kinetics, early/late gene expression as well as viral DNA replication is depicted. LT-Ag and VP1 expression kinetics are based on the Western blots in Figure 27. MCV genome

replication kinetics is hypothetical. (B) A map of MCVmcs with sites of fluorescent protein cassettes and minicircle vector insertion is indicated. (C) A detailed schematic of fluorescently tagged MCV viral proteins that can be expressed from panel (A). Arrows in (B) indicate ribosome-skipping sites for F2A and P2A peptide sequences. (D) A table of tested MCVmc fluorescent reporter genomes. +: detected in transfected cells; −: not detected in transfected cells.

Addressing practical implementation, introduction of a fluorescently tagged viral protein facilitates direct visualization of viral gene expression and the determination of MCV replication kinetics (**Figure 27 and 30A**). Extensive manipulation shows that the site of fluorescent reporter placement is critical. While fusion of mScarlet (mS) to the VP1 C-terminus was tolerated for full genome replication and late gene expression, introduction of fluorescent tags (mS or ZsGreen, ZsG) on either side of the NCCR, in the sense (5' of LT coding sequence) or antisense (5' to VP2 coding sequence) orientation eliminated late gene expression (**Figure 30B-D**). This could be due to interference with proper late gene leader-to-leader splicing [119, 120, 407] or disruption of transcriptional regulatory elements in the NCCR. Ribosome-skipping sequences (P2A or foot-and-mouth disease virus ribosome skipping sequence, F2A: GSGVKQTNLFDLLKLAGDVESNPGP) cloned between VP1, VP2 or LT and fluorescent reporter peptides may minimize interference to replication when compared to their bulky fusion counterparts (**Figure 26C**). However, because the fluorescent reporter tags are cleaved from their corresponding, co-translated viral protein by ribosome skipping, actual fluorescent signals cannot serve as surrogates for localization purposes. Hence, the increased intensity of the cytoplasmic signal from MCVmc.VP1-P2A-mS compared with the predominantly nuclear signal in MCVmc.VP1-mS is likely to reflect cytoplasmic accumulation of cleaved reporter protein rather than VP1 (**Figure 26D**).

Development of the VP1 reporter virus allows measurement of replication kinetics after genome transfection by flow cytometry (**Figure 27A**), which can also be applied to real-time live-

cell microscopy. However, neither the VP1-mS nor VP1-P2A-mS minicircle virus was able to produce Benzonase-protected encapsidated virus genomes and no viral particles were visualized in heavy gradient fractions by electron microscopy. This is likely due to extra amino acids at the VP1 C-terminus (the P2A site leaves a C-terminal 21 amino acid "tag" on VP1) (**Figure 26**). Previous structure studies of MCV viral particles show that alterations at the VP1 C-terminus are functionally crucial. Deletions, including a 37 amino acid sequence unique to MCV, disrupted the ability of VP1 to form VLP [255]. Our studies clearly show that addition of even 21 aa onto the VP1 C-terminus disrupts virion assembly.

We used a protocol similar to that reported by Liu et al. [176], to infect a panel of primary cells. Infection with MCV-ligated- [176, 445] and MCVmc-produced virions that can in turn infect primary HFF-1 cells, BJ and immortalized BJ-hTert cells, while infection of human mesenchymal stem cells is minimal (**Figure 24**). Use of single-round infection with MCVmc.VP1-mS and MCVmc.VP1-P2A-mS reporter viruses (**Figure 29**), packaged into VLP and gradient-purified, can allow the ready manipulation of factors regulating MCV entry, uncoating and genome replication [102, 424, 425].

MCVmc provides an easy and efficient tool to assay MCV replication and its interaction with the host. Our studies confirm the roles of restriction factors in controlling MCV replication, including the Fbw7 interaction site in the LT protein. Fluorescent gene expression coupled with late gene expression provides an excellent tool to visualize MCV replication kinetics in real time and for use in high-throughput screening targeting MCV replication.

5.0 Conclusions and Perspectives

5.1 Summary and Discussion

The advent of new technologies consistently propels biological research forward, facilitating breakthroughs in numerous fields. For instance, as described in Chapter 1.1, advanced technologies have been instrumental in the discovery of novel human tumor viruses, revealing their role in cancer development and broadening our understanding of oncogenic processes. Moreover, cutting-edge tools have paved the way for the characterization of viral-encoded circular RNAs (circRNAs). Despite being relatively novel entities in the realm of molecular biology, circRNAs are now recognized for their potential as regulators of gene expression, playing a crucial role in various physiological and pathological processes. Thanks to technological advancements, we are beginning to unravel the complexities of these circular genetic materials and their influence on the viral life cycle, potentially leading to new strategies for antiviral therapies.

Previously, our laboratory made the important discovery of circular RNAs (circRNAs) encoded by Kaposi sarcoma-associated herpesvirus (KSHV) and Epstein-Barr virus (EBV) [293]. To further elucidate the characteristics of KSHV circRNAs, we scrutinized the expression profiles of circ-vIRF4, circPANs, and circK7.3s across a range of primary effusion lymphoma (PEL) cell lines. Our findings demonstrated that while KSHV circRNAs are ubiquitously expressed, their expression patterns vary across different PEL cell lines. Intriguingly, we discovered that the levels of circPANs and circK7.3s were augmented during KSHV lytic replication, a pattern mirroring their linear counterparts. In contrast, the levels of circ-vIRF4 were not enhanced during lytic replication. This finding implies that the biogenesis of circ-vIRF4 may be regulated via a distinct

mechanism. We also investigated whether KSHV circRNA molecules are incorporated into KSHV virions and found that this incorporation depends on their distribution in the cytoplasm [351, 367]. This trait allows viral circRNAs to function as immediate early genes during the viral infection. Furthermore, we assessed the stability of circ-vIRF4 and found that it has a half-life that is twice as long as its linear transcript. This intriguing finding suggests that the biogenesis of RNAs may be differentially regulated through circularization, opening new avenues for understanding RNA stability and functionality in the context of viral infections.

Following the identification of circRNAs encoded by KSHV and EBV, we further discovered an array of circRNAs encoded by Merkel cell polyomavirus (MCV) and Rat polyomavirus 2 (RatPyV2). This discovery was achieved through RNase R^+ RNA sequencing and RNase R treatment followed by polyadenylation and poly(A)$^+$ RNA Depletion (RPAD). Among the identified circRNAs, we particularly characterized the most abundant MCV-encoded circRNA, circMCV-T. While some viral circRNAs have been reported to produce protein products via m^6A-mediated cap-independent translation [301, 409], our investigation into circMCV-T, through polysome fractionation and Western blot analysis, failed to identify any associated protein products. This observation implies that circMCV-T does not express protein products, however, if circMCV-T is capable to translate protein products remains controversial [151]. When we examined the stability of circMCV-T, we found that it displayed an unexpected sensitivity to RNase R treatment, diverging from the typical stability characteristics of other circRNAs [333, 347]. Given these peculiarities, we sought to determine whether an MCV-encoded miRNA, miR-M1, targeted circMCV-T, analogous to its interaction with early linear transcripts [72, 76]. Our results, derived from co-expression experiments of miR-M1 and a hairpin knockout mutant (hpko), confirmed that miR-M1 indeed downregulated circMCV-T. This discovery suggests that

circMCV-T could serve as a decoy for miR-M1. Further analysis aimed to discern whether circMCV-T could stabilize early transcripts by sequestering or degrading miR-M1. However, the coexpression of circMCV-T led only to a modest increase in T-Ag (sT and LT) transcripts. Interestingly, a hpko mutant version of circMCV-T, which lacked miR-M1 expression and binding sites, demonstrated an enhanced ability to increase MCV's early gene transcription. This outcome suggested an additional circMCV-T-mediated mechanism to stabilize linear T-Ag transcripts or increase the transcription of T-Ag, independent of sequestering miR-M1 [151]. Despite these findings, neither the wild-type nor the hpko mutant circMCV-T significantly increased MCV replication when co-expressed with the entire MCV genome which was produced using the recently adopted minicircle technology. Furthermore, we propose that competitive splicing may provide an intriguing explanation for circMCV-T's function. Given that circMCV-T utilizes the same splice sites as LT and 57kT, the production of circMCV-T could potentially reduce the output of these linear transcripts. This phenomenon could potentially function as a mechanism that shifts the equilibrium of T-Ag expression more towards sT. Consequently, circMCV-T molecules, which absorb miR-M1, may also function as additional factors that protect sT transcripts. Thus, circMCV-T could potentially serve as a crucial component within a regulatory system that intensifies lytic viral replication through the upregulation of sT. This occurs despite the presence of miR-M1, an entity whose expression is conventionally associated with the suppression of lytic replication, thus fostering the establishment of MCV latency. Overall, the results of our research indicate that viral-encoded circRNAs add a new dimension to the regulation of MCV replication. They contribute to an intricate equilibrium involving the regulation of three classes of MCV transcripts: circMCV-T, miR-M1, and linear T-Ag transcripts. This delicate regulatory interplay ultimately fine-tunes the process of MCV replication.

Building upon the capabilities of minicircle technology, we have successfully produced the first reporter MCV virus [184]. Initially, we compared the MCV minicircle (MCVmc) with the ligated MCV genome. This comparison confirmed that the additional 39 bps non-MCV sequence present in MCVmc does not impede MCVmc replication or gene expression. Furthermore, we introduced mutations in the viral genome previously found to be of functional consequence. Through this, we were able to corroborate the roles of multiple host cell restriction factors in viral replication. However, when we attempted to integrate a fluorescent reporter into the MCV genome, we discovered that the placement of this reporter was crucial for its functioning. Notably, the attachment of the mScarlet (mS) tag to the C-terminus of the VP1 protein was found to be compatible with full genome replication and late gene expression. Conversely, the incorporation of fluorescent tags such as mS, ZsGreen, or mNeonGreen at either terminus of the non-coding control region (NCCR) resulted in the abrogation of VP1 expression. This outcome could potentially be attributed to the disruption of late gene leader-to-leader splicing or a perturbation of the transcriptional regulation elements [120]. Moreover, any tag inserted into the C-terminus of LT abolished LT expression. A plausible explanation for this could be the disruption caused to the polyadenylation signal of the early region by the insertion [76]. This finding underscores the necessity for meticulous planning in the placement of tags within the MCV genome to circumvent artificial effects. Despite these complexities, reporter MCVmcs provided us with the tools to analyze MCV replication kinetics through flow cytometry. Nevertheless, it is important to note that the C-terminus of VP1 plays a crucial role in its function [255]. As observed, even a residual string of 21 amino acids at the C-terminus of VP1 can disrupt the encapsidation of MCV. Consequently, we resorted to single-round infections with the reporter MCVmc genome packaged into Virus-Like Particles (VLPs). This approach enables MCV entry, uncoating, and genome

replication. In conclusion, this study demonstrates the value of MCVmc and fluorescent reporter MCV as innovative tools to delve deeper into the biology of MCV. Through these methodologies, we are better equipped to understand and elucidate the complex mechanisms of MCV replication and infection.

5.2 Future Directions

We have identified the existence of multiple MCV-encoded circRNAs, though their precise biological functions remain to be elucidated. Through the implementation of a laccase circRNA expression vector, it becomes feasible to express these circRNAs within mammalian cells. Our approach involves the use of a biotinylated probe, specifically designed to target the splice junction, allowing the pull-down of viral circRNAs via streptavidin beads. This technique will enable the co-precipitation of cellular miRNAs and proteins that interact with the MCV circRNAs. Upon extraction of the co-precipitated miRNAs and proteins, methodologies such as small RNA sequencing and mass spectrometry can be deployed to identify these specific miRNA and protein interactors, respectively. The subsequent validation of these interactions can be achieved through circRNA pull-down, followed by techniques such as reverse transcription quantitative PCR (RT-qPCR) and Western blotting. This comprehensive analysis will provide a foundational framework to further explore and understand the intricate roles that these interactions may play in host cell reprogramming and MCV replication processes.

By employing minicircle technology in conjunction with MCV reporter viruses, we have enabled the observation of MCV VP1 expression, an indicator of efficient viral replication. This has been achieved through the use of live cell imaging and flow cytometry. Despite these

advancements, we have yet to introduce a tag for early genes, largely due to the need to preserve the integrity of MCV transcription and splicing processes. As a result, our research is ongoing, with a focus on identifying an alternative method for tagging these early genes.

The application of reporter MCVmc allows us to address critical questions concerning the MCV life cycle. A key aspect we are interested in is whether MCV replication and gene expression are linked to the cell cycle. If so, which phase of the cell cycle is most conducive to MCV replication and transcription? Furthermore, we are curious if MCV can manipulate cell cycle progression to optimize its replication. If this is the case, what mechanism enables MCV to regulate cell cycle progression? In order to shed light on these questions, we can employ MCVmc reporter transfection and infection into host cells expressing Fluorescence Ubiquitination Cell Cycle Indicators (FUCCI). This will allow us to ascertain whether VP1 expression occurs during specific cell phases and whether there are any discernable differences in cell cycle length between MCV positive and negative cells.

The MCVmc reporters also provide us with the means to study the longevity of MCV's persistence and integration, which emulates the early stage of Merkel Cell Carcinoma (MCC) development, within host cells at a single-cell level. This holds particular significance given that MCV does not express a tethering protein, leaving the mechanism of MCV persistence largely unexplored.

In summary, the use of MCVmc reporters has facilitated a novel approach to studying viral life cycles and the mechanisms behind tumorigenesis. By offering unprecedented insight into these processes, we hope to enhance our understanding of MCV's role in MCC development and potentially contribute to the development of novel therapeutics and prevention strategies.

Bibliography

1. Payne, L.N. and V. Nair, *The long view: 40 years of avian leukosis research.* Avian Pathol, 2012. **41**(1): p. 11-9.
2. Rous, P., *A Sarcoma of the Fowl Transmissible by an Agent Separable from the Tumor Cells.* J Exp Med, 1911. **13**(4): p. 397-411.
3. Shope, R.E. and E.W. Hurst, *Infectious Papillomatosis of Rabbits : With a Note on the Histopathology.* J Exp Med, 1933. **58**(5): p. 607-24.
4. Gross, L., *A filterable agent, recovered from Ak leukemic extracts, causing salivary gland carcinomas in C3H mice.* Proc Soc Exp Biol Med, 1953. **83**(2): p. 414-21.
5. Stewart, S.E. *Leukemia in mice produced by a filterable agent present in AKR leukemic tissues with notes on a sarcoma produced by the same agent*. in *Anatomical Record*. 1953. WILEY-LISS DIV JOHN WILEY & SONS INC, 605 THIRD AVE, NEW YORK, NY 10158-0012.
6. Sweet, B.H. and M.R. Hilleman, *The vacuolating virus, S.V. 40.* Proc Soc Exp Biol Med, 1960. **105**: p. 420-7.
7. Poulin, D.L. and J.A. Decaprio, *Is There a Role for SV40 in Human Cancer?* Journal of Clinical Oncology, 2006. **24**(26): p. 4356-4365.
8. Epstein, M.A., B.G. Achong, and Y.M. Barr, *Virus Particles in Cultured Lymphoblasts from Burkitt's Lymphoma.* Lancet, 1964. **1**(7335): p. 702-3.
9. Henle, W., et al., *Herpes-type virus and chromosome marker in normal leukocytes after growth with irradiated Burkitt cells.* Science, 1967. **157**(3792): p. 1064-5.
10. Saha, A., et al., *Tumor viruses and cancer biology: Modulating signaling pathways for therapeutic intervention.* Cancer Biol Ther, 2010. **10**(10): p. 961-78.
11. Farrell, P.J., *Epstein-Barr Virus and Cancer.* Annu Rev Pathol, 2019. **14**: p. 29-53.
12. Patel, P.D., et al., *The Association of Epstein-Barr Virus With Cancer.* Cureus, 2022. **14**(6): p. e26314.
13. Qiu, J. and D.A. Thorley-Lawson, *EBV microRNA BART 18-5p targets MAP3K2 to facilitate persistence in vivo by inhibiting viral replication in B cells.* Proc Natl Acad Sci U S A, 2014. **111**(30): p. 11157-62.
14. Hooykaas, M.J.G., et al., *EBV MicroRNA BART16 Suppresses Type I IFN Signaling.* J Immunol, 2017. **198**(10): p. 4062-4073.
15. Cai, L., et al., *EBV-miR-BART7-3p Imposes Stemness in Nasopharyngeal Carcinoma Cells by Suppressing SMAD7.* Front Genet, 2019. **10**: p. 939.
16. Liu, Q., M. Shuai, and Y. Xia, *Knockdown of EBV-encoded circRNA circRPMS1 suppresses nasopharyngeal carcinoma cell proliferation and metastasis through sponging multiple miRNAs.* Cancer Manag Res, 2019. **11**: p. 8023-8031.
17. Gong, L.P., et al., *Epstein-Barr virus-derived circular RNA LMP2A induces stemness in EBV-associated gastric cancer.* EMBO Rep, 2020. **21**(10): p. e49689.
18. Blumberg, B.S., *A "New" Antigen in Leukemia Sera.* JAMA: The Journal of the American Medical Association, 1965. **191**(7): p. 541.
19. Blumberg, B.S., et al., *The relation of infection with the hepatitis B agent to primary hepatic carcinoma.* Am J Pathol, 1975. **81**(3): p. 669-82.

20. Beasley, R.P., et al., *Hepatocellular carcinoma and hepatitis B virus. A prospective study of 22 707 men in Taiwan.* Lancet, 1981. **2**(8256): p. 1129-33.
21. Buynak, E.B., et al., *Vaccine against human hepatitis B.* JAMA, 1976. **235**(26): p. 2832-4.
22. Fields, B.N., D.M. Knipe, and P.M. Howley, *Fields virology*. 5th ed. 2007, Philadelphia: Lippincott Williams & Wilkins. xvii, (various paging) : ill (chiefly col.).
23. Choo, Q.-L., et al., *Isolation of a cDNA cLone Derived from a Blood-Borne Non-A, Non-B Viral Hepatitis Genome.* Science, 1989. **244**(4902): p. 359-362.
24. Poiesz, B.J., et al., *Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma.* Proc Natl Acad Sci U S A, 1980. **77**(12): p. 7415-9.
25. Hinuma, Y., et al., *Adult T-cell leukemia: antigen in an ATL cell line and detection of antibodies to the antigen in human sera.* Proc Natl Acad Sci U S A, 1981. **78**(10): p. 6476-80.
26. Matsuoka, M. and K.T. Jeang, *Human T-cell leukaemia virus type 1 (HTLV-1) infectivity and cellular transformation.* Nat Rev Cancer, 2007. **7**(4): p. 270-80.
27. Dürst, M., et al., *A papillomavirus DNA from a cervical carcinoma and its prevalence in cancer biopsy samples from different geographic regions.* Proceedings of the National Academy of Sciences, 1983. **80**(12): p. 3812-3815.
28. Boshart, M., et al., *A new type of papillomavirus DNA, its presence in genital cancer biopsies and in cell lines derived from cervical cancer.* The EMBO Journal, 1984. **3**(5): p. 1151-1157.
29. Frazer, I.H., D.R. Lowy, and J.T. Schiller, *Prevention of cancer through immunization: Prospects and challenges for the 21st century.* European Journal of Immunology, 2007. **37**(S1): p. S148-S155.
30. Arbyn, M., et al., *EUROGIN 2011 roadmap on prevention and treatment of HPV-related disease.* Int J Cancer, 2012. **131**(9): p. 1969-82.
31. Lisitsyn, N., N. Lisitsyn, and M. Wigler, *Cloning the differences between two complex genomes.* Science, 1993. **259**(5097): p. 946-51.
32. Chang, Y., et al., *Identification of Herpesvirus-Like DNA Sequences in AIDS-Sssociated Kaposi's Sarcoma.* Science, 1994. **266**(5192): p. 1865-1869.
33. Rady, P.L., et al., *Herpesvirus-like DNA sequences in classic Kaposi's sarcomas.* J Med Virol, 1995. **47**(2): p. 179-83.
34. Schalling, M., et al., *A role for a new herpes virus (KSHV) in different forms of Kaposi's sarcoma.* Nat Med, 1995. **1**(7): p. 707-8.
35. Lennette, E.T., D.J. Blackbourn, and J.A. Levy, *Antibodies to human herpesvirus type 8 in the general population and in Kaposi's sarcoma patients.* Lancet, 1996. **348**(9031): p. 858-61.
36. Aluigi, M.G., et al., *KSHV sequences in biopsies and cultured spindle cells of epidemic, iatrogenic and Mediterranean forms of Kaposi's sarcoma.* Res Virol, 1996. **147**(5): p. 267-75.
37. Dictor, M., et al., *Human herpesvirus 8 (Kaposi's sarcoma-associated herpesvirus) DNA in Kaposi's sarcoma lesions, AIDS Kaposi's sarcoma cell lines, endothelial Kaposi's sarcoma simulators, and the skin of immunosuppressed patients.* Am J Pathol, 1996. **148**(6): p. 2009-16.

38. Feng, H., et al., *Clonal integration of a polyomavirus in human Merkel cell carcinoma.* Science, 2008. **319**(5866): p. 1096-100.
39. Girardi, A.J., et al., *Development of tumors in hamsters inoculated in the neonatal period with vacuolating virus, SV-40.* Proc Soc Exp Biol Med, 1962. **109**: p. 649-60.
40. Gardner, S.D., et al., *New human papovavirus (B.K.) isolated from urine after renal transplantation.* Lancet, 1971. **1**(7712): p. 1253-7.
41. Padgett, B.L., et al., *Cultivation of papova-like virus from human brain with progressive multifocal leucoencephalopathy.* Lancet, 1971. **1**(7712): p. 1257-60.
42. Egli, A., et al., *Prevalence of polyomavirus BK and JC infection and replication in 400 healthy blood donors.* J Infect Dis, 2009. **199**(6): p. 837-46.
43. Antonsson, A., et al., *Prevalence and stability of antibodies to the BK and JC polyomaviruses: a long-term longitudinal study of Australians.* J Gen Virol, 2010. **91**(Pt 7): p. 1849-53.
44. Laine, H.K., et al., *Seroprevalence of polyomaviruses BK and JC in Finnish women and their spouses followed-up for three years.* Sci Rep, 2023. **13**(1): p. 879.
45. Stolt, A., et al., *Seroepidemiology of the human polyomaviruses.* J Gen Virol, 2003. **84**(Pt 6): p. 1499-1504.
46. Kean, J.M., et al., *Seroepidemiology of human polyomaviruses.* PLoS Pathog, 2009. **5**(3): p. e1000363.
47. Gossai, A., et al., *Seroepidemiology of Human Polyomaviruses in a US Population.* Am J Epidemiol, 2016. **183**(1): p. 61-9.
48. Allander, T., et al., *Identification of a third human polyomavirus.* J Virol, 2007. **81**(8): p. 4130-6.
49. Kamminga, S., et al., *Seroprevalence of fourteen human polyomaviruses determined in blood donors.* PLoS One, 2018. **13**(10): p. e0206273.
50. Gaynor, A.M., et al., *Identification of a novel polyomavirus from patients with acute respiratory tract infections.* PLoS Pathog, 2007. **3**(5): p. e64.
51. Jeles, K., M. Katona, and E. Csoma, *Seroprevalence of Four Polyomaviruses Linked to Dermatological Diseases: New Findings and a Comprehensive Analysis.* Viruses, 2022. **14**(10).
52. Schowalter, R.M., et al., *Merkel Cell Polyomavirus and Two Previously Unknown Polyomaviruses Are Chronically Shed from Human Skin.* Cell Host & Microbe, 2010. **7**(6): p. 509-515.
53. Sroller, V., et al., *Seroprevalence rates of HPyV6, HPyV7, TSPyV, HPyV9, MWPyV and KIPyV polyomaviruses among the healthy blood donors.* J Med Virol, 2016. **88**(7): p. 1254-61.
54. Van Der Meijden, E., et al., *Discovery of a New Human Polyomavirus Associated with Trichodysplasia Spinulosa in an Immunocompromized Patient.* PLoS Pathogens, 2010. **6**(7): p. e1001024.
55. Scuda, N., et al., *A Novel Human Polyomavirus Closely Related to the African Green Monkey-Derived Lymphotropic Polyomavirus.* Journal of Virology, 2011. **85**(9): p. 4586-4590.
56. Siebrasse, E.A., et al., *Identification of MW Polyomavirus, a Novel Polyomavirus in Human Stool.* Journal of Virology, 2012. **86**(19): p. 10321-10326.

57. Lim, E.S., et al., *Discovery of STL polyomavirus, a polyomavirus of ancestral recombinant origin that encodes a unique T antigen by alternative splicing.* Virology, 2013. **436**(2): p. 295-303.
58. Lim, E.S., et al., *Common exposure to STL polyomavirus during childhood.* Emerg Infect Dis, 2014. **20**(9): p. 1559-61.
59. Korup, S., et al., *Identification of a Novel Human Polyomavirus in Organs of the Gastrointestinal Tract.* PLoS ONE, 2013. **8**(3): p. e58021.
60. Gaboriaud, P., et al., *Age-specific seroprevalence of human polyomavirus 12 and Saint Louis and New Jersey polyomaviruses.* Emerg Microbes Infect, 2018. **7**(1): p. 22.
61. Mishra, N., et al., *Identification of a Novel Polyomavirus in a Pancreatic Transplant Recipient With Retinal Blindness and Vasculitic Myopathy.* Journal of Infectious Diseases, 2014. **210**(10): p. 1595-1599.
62. Zhou, X., et al., *Characterization of the self-assembly of New Jersey polyomavirus VP1 into virus-like particles and the virus seroprevalence in Japan.* Sci Rep, 2019. **9**(1): p. 13085.
63. Gheit, T., et al., *Isolation and characterization of a novel putative human polyomavirus.* Virology, 2017. **506**: p. 45-54.
64. Kamminga, S., et al., *Serology Identifies LIPyV as a Feline Rather than a Human Polyomavirus.* Viruses, 2023. **15**(7).
65. DeCaprio, J.A. and R.L. Garcea, *A cornucopia of human polyomaviruses.* Nat Rev Microbiol, 2013. **11**(4): p. 264-76.
66. Shuda, M., et al., *T antigen mutations are a human tumor-specific signature for Merkel cell polyomavirus.* Proc Natl Acad Sci U S A, 2008. **105**(42): p. 16272-7.
67. Carter, J.J., et al., *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, 2013. **110**(31): p. 12744-9.
68. Zerrahn, J., et al., *Independent expression of the transforming amino-terminal domain of SV40 large I antigen from an alternatively spliced third SV40 early mRNA.* EMBO J, 1993. **12**(12): p. 4739-46.
69. Abend, J.R., et al., *A truncated T antigen expressed from an alternatively spliced BK virus early mRNA.* J Gen Virol, 2009. **90**(Pt 5): p. 1238-1245.
70. Sullivan, C.S., et al., *SV40-encoded microRNAs regulate viral gene expression and reduce susceptibility to cytotoxic T cells.* Nature, 2005. **435**(7042): p. 682-6.
71. Seo, G.J., et al., *Evolutionarily conserved function of a viral microRNA.* J Virol, 2008. **82**(20): p. 9823-8.
72. Seo, G.J., C.J. Chen, and C.S. Sullivan, *Merkel cell polyomavirus encodes a microRNA with the ability to autoregulate viral gene expression.* Virology, 2009. **383**(2): p. 183-7.
73. Akhbari, P., et al., *MCV-miR-M1 Targets the Host-Cell Immune Response Resulting in the Attenuation of Neutrophil Chemotaxis.* J Invest Dermatol, 2018. **138**(11): p. 2343-2354.
74. Sullivan, C.S., et al., *Murine Polyomavirus encodes a microRNA that cleaves early RNA transcripts but is not essential for experimental infection.* Virology, 2009. **387**(1): p. 157-67.
75. Bauman, Y., et al., *An identical miRNA of the human JC and BK polyoma viruses targets the stress-induced ligand ULBP3 to escape immune elimination.* Cell Host Microbe, 2011. **9**(2): p. 93-102.

76. Theiss, J.M., et al., *A Comprehensive Analysis of Replicating Merkel Cell Polyomavirus Genomes Delineates the Viral Transcription Program and Suggests a Role for mcv-miR-M1 in Episomal Persistence.* PLoS Pathog, 2015. **11**(7): p. e1004974.
77. An, P., M.T. Saenz Robles, and J.M. Pipas, *Large T antigens of polyomaviruses: amazing molecular machines.* Annu Rev Microbiol, 2012. **66**: p. 213-36.
78. Sullivan, C.S., P. Cantalupo, and J.M. Pipas, *The molecular chaperone activity of simian virus 40 large T antigen is required to disrupt Rb-E2F family complexes by an ATP-dependent mechanism.* Mol Cell Biol, 2000. **20**(17): p. 6233-43.
79. Garimella, R., et al., *Hsc70 contacts helix III of the J domain from polyomavirus T antigens: addressing a dilemma in the chaperone hypothesis of how they release E2F from pRb.* Biochemistry, 2006. **45**(22): p. 6917-29.
80. Diaz, J., et al., *Phosphorylation of large T antigen regulates merkel cell polyomavirus replication.* Cancers (Basel), 2014. **6**(3): p. 1464-86.
81. Valle, M., et al., *Large T-antigen double hexamers imaged at the simian virus 40 origin of replication.* Mol Cell Biol, 2000. **20**(1): p. 34-41.
82. Tsurimoto, T., M.P. Fairman, and B. Stillman, *Simian virus 40 DNA replication in vitro: identification of multiple stages of initiation.* Mol Cell Biol, 1989. **9**(9): p. 3839-49.
83. Bullock, P.A., *The initiation of simian virus 40 DNA replication in vitro.* Crit Rev Biochem Mol Biol, 1997. **32**(6): p. 503-68.
84. Trowbridge, P.W., R. Roy, and D.T. Simmons, *Human topoisomerase I promotes initiation of simian virus 40 DNA replication in vitro.* Mol Cell Biol, 1999. **19**(3): p. 1686-94.
85. Waga, S. and B. Stillman, *Anatomy of a DNA replication fork revealed by reconstitution of SV40 DNA replication in vitro.* Nature, 1994. **369**(6477): p. 207-12.
86. Bruckner, A., et al., *The mouse DNA polymerase alpha-primase subunit p48 mediates species-specific replication of polyomavirus DNA in vitro.* Mol Cell Biol, 1995. **15**(3): p. 1716-24.
87. Pallas, D.C., et al., *Polyoma small and middle T antigens and SV40 small t antigen form stable complexes with protein phosphatase 2A.* Cell, 1990. **60**(1): p. 167-76.
88. Pallas, D.C., et al., *The third subunit of protein phosphatase 2A (PP2A), a 55-kilodalton protein which is apparently substituted for by T antigens in complexes with the 36- and 63-kilodalton PP2A subunits, bears little resemblance to T antigens.* J Virol, 1992. **66**(2): p. 886-93.
89. Kwun, H.J., et al., *Restricted protein phosphatase 2A targeting by Merkel cell polyomavirus small T antigen.* J Virol, 2015. **89**(8): p. 4191-200.
90. Freund, R., et al., *Polyoma virus middle T is essential for virus replication and persistence as well as for tumor induction in mice.* Virology, 1992. **191**(2): p. 716-23.
91. Fluck, M.M. and B.S. Schaffhausen, *Lessons in signaling and tumorigenesis from polyomavirus middle T antigen.* Microbiol Mol Biol Rev, 2009. **73**(3): p. 542-63, Table of Contents.
92. Gjoerup, O. and Y. Chang, *Update on human polyomaviruses and cancer.* Adv Cancer Res, 2010. **106**: p. 1-51.
93. Salunke, D.M., D.L. Caspar, and R.L. Garcea, *Polymorphism in the assembly of polyomavirus capsid protein VP1.* Biophys J, 1989. **56**(5): p. 887-900.
94. Chen, X.S., T. Stehle, and S.C. Harrison, *Interaction of polyomavirus internal protein VP2 with the major capsid protein VP1 and implications for participation of VP2 in viral entry.* EMBO J, 1998. **17**(12): p. 3233-40.

95. Liddington, R.C., et al., *Structure of simian virus 40 at 3.8-A resolution.* Nature, 1991. **354**(6351): p. 278-84.
96. Gerits, N. and U. Moens, *Agnoprotein of mammalian polyomaviruses.* Virology, 2012. **432**(2): p. 316-26.
97. Saribas, A.S., et al., *Emerging From the Unknown: Structural and Functional Features of Agnoprotein of Polyomaviruses.* J Cell Physiol, 2016. **231**(10): p. 2115-27.
98. Smith, A.E., H. Lilie, and A. Helenius, *Ganglioside-dependent cell attachment and endocytosis of murine polyomavirus-like particles.* FEBS Lett, 2003. **555**(2): p. 199-203.
99. Tsai, B., et al., *Gangliosides are receptors for murine polyoma virus and SV40.* EMBO J, 2003. **22**(17): p. 4346-55.
100. Dugan, A.S., S. Eash, and W.J. Atwood, *An N-linked glycoprotein with alpha(2,3)-linked sialic acid is a receptor for BK virus.* J Virol, 2005. **79**(22): p. 14442-5.
101. Erickson, K.D., R.L. Garcea, and B. Tsai, *Ganglioside GT1b is a putative host cell receptor for the Merkel cell polyomavirus.* J Virol, 2009. **83**(19): p. 10275-9.
102. Schowalter, R.M. and C.B. Buck, *The Merkel cell polyomavirus minor capsid protein.* PLoS Pathog, 2013. **9**(8): p. e1003558.
103. Neu, U., et al., *Structures of Merkel cell polyomavirus VP1 complexes define a sialic acid binding site required for infection.* PLoS Pathog, 2012. **8**(7): p. e1002738.
104. Pho, M.T., A. Ashok, and W.J. Atwood, *JC virus enters human glial cells by clathrin-dependent receptor-mediated endocytosis.* J Virol, 2000. **74**(5): p. 2288-92.
105. Assetta, B., et al., *5-HT2 receptors facilitate JC polyomavirus entry.* J Virol, 2013. **87**(24): p. 13490-8.
106. Anderson, H.A., Y. Chen, and L.C. Norkin, *Bound simian virus 40 translocates to caveolin-enriched membrane domains, and its entry is inhibited by drugs that selectively disrupt caveolae.* Mol Biol Cell, 1996. **7**(11): p. 1825-34.
107. Richterova, Z., et al., *Caveolae are involved in the trafficking of mouse polyomavirus virions and artificial VP1 pseudocapsids toward cell nuclei.* J Virol, 2001. **75**(22): p. 10880-91.
108. Damm, E.M., et al., *Clathrin- and caveolin-1-independent endocytosis: entry of simian virus 40 into cells devoid of caveolae.* J Cell Biol, 2005. **168**(3): p. 477-88.
109. Eash, S., W. Querbes, and W.J. Atwood, *Infection of vero cells by BK virus is dependent on caveolae.* J Virol, 2004. **78**(21): p. 11583-90.
110. Ashok, A. and W.J. Atwood, *Contrasting roles of endosomal pH and the cytoskeleton in infection of human glial cells by JC virus and simian virus 40.* J Virol, 2003. **77**(2): p. 1347-56.
111. Becker, M., et al., *Infectious Entry of Merkel Cell Polyomavirus.* J Virol, 2019. **93**(6).
112. Jiang, M., et al., *Early events during BK virus entry and disassembly.* J Virol, 2009. **83**(3): p. 1350-8.
113. Magnuson, B., et al., *ERp29 triggers a conformational change in polyomavirus to stimulate membrane binding.* Mol Cell, 2005. **20**(2): p. 289-300.
114. Gilbert, J., et al., *Downregulation of protein disulfide isomerase inhibits infection by the mouse polyomavirus.* J Virol, 2006. **80**(21): p. 10868-70.
115. Schelhaas, M., et al., *Simian Virus 40 depends on ER protein folding and quality control factors for entry into host cells.* Cell, 2007. **131**(3): p. 516-29.
116. Walczak, C.P. and B. Tsai, *A PDI family network acts distinctly and coordinately with ERp29 to facilitate polyomavirus infection.* J Virol, 2011. **85**(5): p. 2386-96.

117. Geiger, R., et al., *BAP31 and BiP are essential for dislocation of SV40 from the endoplasmic reticulum to the cytosol.* Nat Cell Biol, 2011. **13**(11): p. 1305-14.
118. Inoue, T. and B. Tsai, *A large and intact viral particle penetrates the endoplasmic reticulum membrane to reach the cytosol.* PLoS Pathog, 2011. **7**(5): p. e1002037.
119. Hyde-DeRuyscher, R. and G.G. Carmichael, *Polyomavirus early-late switch is not regulated at the level of transcription initiation and is associated with changes in RNA processing.* Proc Natl Acad Sci U S A, 1988. **85**(23): p. 8993-7.
120. Hyde-DeRuyscher, R.P. and G.G. Carmichael, *Polyomavirus late pre-mRNA processing: DNA replication-associated changes in leader exon multiplicity suggest a role for leader-to-leader splicing in the early-late switch.* J Virol, 1990. **64**(12): p. 5823-32.
121. Clayson, E.T., L.V. Brando, and R.W. Compans, *Release of simian virus 40 virions from epithelial cells is polarized and occurs without cell lysis.* J Virol, 1989. **63**(5): p. 2278-88.
122. Suzuki, T., et al., *The human polyoma JC virus agnoprotein acts as a viroporin.* PLoS Pathog, 2010. **6**(3): p. e1000801.
123. Raghava, S., et al., *The SV40 late protein VP4 is a viroporin that forms pores to disrupt membranes for viral release.* PLoS Pathog, 2011. **7**(6): p. e1002116.
124. Raghava, S., et al., *SV40 late protein VP4 forms toroidal pores to disrupt membranes for viral release.* Biochemistry, 2013. **52**(22): p. 3939-48.
125. Evans, G.L., et al., *Anion homeostasis is important for non-lytic release of BK polyomavirus from infected cells.* Open Biol, 2015. **5**(8).
126. Hughes, M.P., et al., *Merkel Cell Carcinoma: Epidemiology, Target, and Therapy.* Curr Dermatol Rep, 2014. **3**(1): p. 46-53.
127. Busam, K.J., et al., *Merkel cell polyomavirus expression in merkel cell carcinomas and its absence in combined tumors and pulmonary neuroendocrine carcinomas.* Am J Surg Pathol, 2009. **33**(9): p. 1378-85.
128. Paulson, K.G., et al., *Merkel cell carcinoma: Current US incidence and projected increases based on changing demographics.* J Am Acad Dermatol, 2018. **78**(3): p. 457-463 e2.
129. Swann, M.H. and J. Yoon, *Merkel cell carcinoma.* Semin Oncol, 2007. **34**(1): p. 51-6.
130. Heath, M., et al., *Clinical characteristics of Merkel cell carcinoma at diagnosis in 195 patients: the AEIOU features.* J Am Acad Dermatol, 2008. **58**(3): p. 375-81.
131. Albores-Saavedra, J., et al., *Merkel cell carcinoma demographics, morphology, and survival based on 3870 cases: a population based study.* J Cutan Pathol, 2010. **37**(1): p. 20-7.
132. Uchi, H., *Merkel Cell Carcinoma: An Update and Immunotherapy.* Front Oncol, 2018. **8**: p. 48.
133. Ma, J.E. and J.D. Brewer, *Merkel cell carcinoma in immunosuppressed patients.* Cancers (Basel), 2014. **6**(3): p. 1328-50.
134. Hemminki, K., et al., *Kaposi sarcoma and Merkel cell carcinoma after autoimmune disease.* Int J Cancer, 2012. **131**(3): p. E326-8.
135. Silling, S., et al., *Epidemiology of Merkel Cell Polyomavirus Infection and Merkel Cell Carcinoma.* Cancers (Basel), 2022. **14**(24).
136. Katano, H., et al., *Detection of Merkel cell polyomavirus in Merkel cell carcinoma and Kaposi's sarcoma.* J Med Virol, 2009. **81**(11): p. 1951-8.
137. Paulson, K.G., et al., *Antibodies to merkel cell polyomavirus T antigen oncoproteins reflect tumor burden in merkel cell carcinoma patients.* Cancer Res, 2010. **70**(21): p. 8388-97.

138. Feng, H., et al., *Cellular and viral factors regulating Merkel cell polyomavirus replication.* PLoS One, 2011. **6**(7): p. e22468.
139. Foulongne, V., et al., *Human skin microbiota: high diversity of DNA viruses identified on the human skin by high throughput sequencing.* PLoS One, 2012. **7**(6): p. e38499.
140. Martel-Jantin, C., et al., *Genetic variability and integration of Merkel cell polyomavirus in Merkel cell carcinoma.* Virology, 2012. **426**(2): p. 134-42.
141. Martel-Jantin, C., et al., *Molecular epidemiology of merkel cell polyomavirus: evidence for geographically related variant genotypes.* J Clin Microbiol, 2014. **52**(5): p. 1687-90.
142. Nguyen, K.D., et al., *Human polyomavirus 6 and 7 are associated with pruritic and dyskeratotic dermatoses.* J Am Acad Dermatol, 2017. **76**(5): p. 932-940 e3.
143. Torres, C., et al., *Phylodynamics of Merkel-cell polyomavirus and human polyomavirus 6: A long-term history with humans.* Mol Phylogenet Evol, 2018. **126**: p. 210-220.
144. Jin, H.T., et al., *The frequency of Merkel cell polyomavirus in whole blood from immunocompetent and immunosuppressed patients with kidney disease and healthy donors.* Microb Pathog, 2019. **131**: p. 75-80.
145. Passerini, S., et al., *Detection of Merkel Cell Polyomavirus (MCPyV) DNA and Transcripts in Merkel Cell Carcinoma (MCC).* Pathogens, 2023. **12**(7).
146. Kwun, H.J., et al., *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, 2009. **83**(23): p. 12118-28.
147. Kwun, H.J., Y. Chang, and P.S. Moore, *Protein-mediated viral latency is a novel mechanism for Merkel cell polyomavirus persistence.* Proc Natl Acad Sci U S A, 2017. **114**(20): p. E4040-E4047.
148. Pastrana, D.V., et al., *Quantitation of human seroresponsiveness to Merkel cell polyomavirus.* PLoS Pathog, 2009. **5**(9): p. e1000578.
149. Neumann, F., et al., *Replication, gene expression and particle production by a consensus Merkel Cell Polyomavirus (MCPyV) genome.* PLoS One, 2011. **6**(12): p. e29112.
150. Abere, B., et al., *Merkel Cell Polyomavirus Encodes Circular RNAs (circRNAs) Enabling a Dynamic circRNA/microRNA/mRNA Regulatory Network.* mBio, 2020. **11**(6).
151. Yang, R., et al., *Characterization of ALTO-encoding circular RNAs expressed by Merkel cell polyomavirus and trichodysplasia spinulosa polyomavirus.* PLoS Pathog, 2021. **17**(5): p. e1009582.
152. Sastre-Garau, X., et al., *Merkel cell carcinoma of the skin: pathological and molecular evidence for a causative role of MCV in oncogenesis.* J Pathol, 2009. **218**(1): p. 48-56.
153. Shuda, M., et al., *Human Merkel cell polyomavirus infection I. MCV T antigen expression in Merkel cell carcinoma, lymphoid tissues and lymphoid tumors.* Int J Cancer, 2009. **125**(6): p. 1243-9.
154. Laude, H.C., et al., *Distinct merkel cell polyomavirus molecular features in tumour and non tumour specimens from patients with merkel cell carcinoma.* PLoS Pathog, 2010. **6**(8): p. e1001076.
155. Houben, R., et al., *Merkel cell polyomavirus-infected Merkel cell carcinoma cells require expression of viral T antigens.* J Virol, 2010. **84**(14): p. 7064-72.
156. Shuda, M., Y. Chang, and P.S. Moore, *Merkel cell polyomavirus-positive Merkel cell carcinoma requires viral small T-antigen for cell proliferation.* J Invest Dermatol, 2014. **134**(5): p. 1479-1481.

157. Houben, R., et al., *Characterization of functional domains in the Merkel cell polyoma virus Large T antigen.* Int J Cancer, 2015. **136**(5): p. E290-300.
158. Hesbacher, S., et al., *RB1 is the crucial target of the Merkel cell polyomavirus Large T antigen in Merkel cell carcinoma cells.* Oncotarget, 2016. **7**(22): p. 32956-68.
159. Kwun, H.J., et al., *Merkel cell polyomavirus small T antigen controls viral replication and oncoprotein expression by targeting the cellular ubiquitin ligase SCFFbw7.* Cell Host Microbe, 2013. **14**(2): p. 125-35.
160. Shuda, M., et al., *Merkel Cell Polyomavirus Small T Antigen Induces Cancer and Embryonic Merkel Cell Proliferation in a Transgenic Mouse Model.* PLoS One, 2015. **10**(11): p. e0142329.
161. Verhaegen, M.E., et al., *Merkel cell polyomavirus small T antigen is oncogenic in transgenic mice.* J Invest Dermatol, 2015. **135**(5): p. 1415-1424.
162. Cheng, J., et al., *Merkel cell polyomavirus recruits MYCL to the EP400 complex to promote oncogenesis.* PLoS Pathog, 2017. **13**(10): p. e1006668.
163. Tolstov, Y.L., et al., *Human Merkel cell polyomavirus infection II. MCV is a common human infection that can be detected by conformational capsid epitope immunoassays.* Int J Cancer, 2009. **125**(6): p. 1250-6.
164. Moll, I., C. Kuhn, and R. Moll, *Cytokeratin 20 is a general marker of cutaneous Merkel cells while certain neuronal proteins are absent.* J Invest Dermatol, 1995. **104**(6): p. 910-5.
165. Cheuk, W., et al., *Immunostaining for thyroid transcription factor 1 and cytokeratin 20 aids the distinction of small cell carcinoma from Merkel cell carcinoma, but not pulmonary from extrapulmonary small cell carcinomas.* Arch Pathol Lab Med, 2001. **125**(2): p. 228-31.
166. Halata, Z., M. Grim, and K.I. Bauman, *Friedrich Sigmund Merkel and his "Merkel cell", morphology, development, and physiology: review and new results.* Anat Rec A Discov Mol Cell Evol Biol, 2003. **271**(1): p. 225-39.
167. Koljonen, V., et al., *Neuroendocrine differentiation in primary Merkel cell carcinoma--possible prognostic significance.* Anticancer Res, 2005. **25**(2A): p. 853-8.
168. Van Keymeulen, A., et al., *Epidermal progenitors give rise to Merkel cells during embryonic development and adult homeostasis.* J Cell Biol, 2009. **187**(1): p. 91-100.
169. Godlewski, J., et al., *Plasticity of neuropeptidergic neoplasm cells in the primary and metastatic Merkel cell carcinoma.* Folia Histochem Cytobiol, 2013. **51**(2): p. 168-73.
170. Kervarrec, T., et al., *Diagnostic accuracy of a panel of immunohistochemical and molecular markers to distinguish Merkel cell carcinoma from other neuroendocrine carcinomas.* Mod Pathol, 2019. **32**(4): p. 499-510.
171. Nguyen, B.D. and A.E. McCullough, *Imaging of Merkel cell carcinoma.* Radiographics, 2002. **22**(2): p. 367-76.
172. Su, L.D., et al., *CD117 (KIT receptor) expression in Merkel cell carcinoma.* Am J Dermatopathol, 2002. **24**(4): p. 289-93.
173. Deichmann, M., et al., *Adhesion molecules CD171 (L1CAM) and CD24 are expressed by primary neuroendocrine carcinomas of the skin (Merkel cell carcinomas).* J Cutan Pathol, 2003. **30**(6): p. 363-8.
174. Strong, S., et al., *KIT receptor (CD117) expression in Merkel cell carcinoma.* Br J Dermatol, 2004. **150**(2): p. 384-5.

175. Li, J., et al., *Merkel cell polyomavirus large T antigen disrupts host genomic integrity and inhibits cellular proliferation.* J Virol, 2013. **87**(16): p. 9173-88.
176. Liu, W., et al., *Identifying the Target Cells and Mechanisms of Merkel Cell Polyomavirus Infection.* Cell Host Microbe, 2016. **19**(6): p. 775-87.
177. Peretti, A., et al., *Analysis of human beta-papillomavirus and Merkel cell polyomavirus infection in skin lesions and eyebrow hair bulbs from a cohort of patients with chronic lymphocytic leukaemia.* Br J Dermatol, 2014. **171**(6): p. 1525-8.
178. Hampras, S.S., et al., *Merkel cell polyomavirus (MCV) T-antigen seroreactivity, MCV DNA in eyebrow hairs, and squamous cell carcinoma.* Infect Agent Cancer, 2015. **10**: p. 35.
179. Bellaud, G., et al., *Prevalence of human polyomavirus DNA in eyebrow hairs plucked from patients with psoriasis treated with TNF inhibitors.* J Eur Acad Dermatol Venereol, 2015. **29**(5): p. 1019-21.
180. Spurgeon, M.E. and P.F. Lambert, *Merkel cell polyomavirus: a newly discovered human virus with oncogenic potential.* Virology, 2013. **435**(1): p. 118-30.
181. Harrison, C.J., et al., *Asymmetric assembly of Merkel cell polyomavirus large T-antigen origin binding domains at the viral origin.* J Mol Biol, 2011. **409**(4): p. 529-42.
182. Liu, W., et al., *Merkel Cell Polyomavirus Infection and Detection.* J Vis Exp, 2019(144).
183. Abere, B., et al., *Replication Kinetics for a Reporter Merkel Cell Polyomavirus.* Viruses, 2022. **14**(3).
184. Kay, M.A., C.Y. He, and Z.Y. Chen, *A robust system for production of minicircle DNA vectors.* Nat Biotechnol, 2010. **28**(12): p. 1287-9.
185. Alves, C.P.A., D.M.F. Prazeres, and G.A. Monteiro, *Minicircle Biopharmaceuticals–An Overview of Purification Strategies.* Frontiers in Chemical Engineering, 2021. **2**.
186. Darquet, A.M., et al., *A new DNA vehicle for nonviral gene delivery: supercoiled minicircle.* Gene Ther, 1997. **4**(12): p. 1341-9.
187. Bigger, B.W., et al., *An araC-controlled bacterial cre expression system to produce DNA minicircle vectors for nuclear and mitochondrial gene therapy.* J Biol Chem, 2001. **276**(25): p. 23018-27.
188. Nehlsen, K., S. Broll, and J. Bode, *Replicating minicircles: Generation of nonviral episomes for the efficient modification of dividing cells.* Gene Ther Mol Biol, 2006. **10**: p. 233-244.
189. Jechlinger, W., et al., *Minicircle DNA immobilized in bacterial ghosts: in vivo production of safe non-viral DNA delivery vehicles.* J Mol Microbiol Biotechnol, 2004. **8**(4): p. 222-31.
190. Mayrhofer, P., et al., *Minicircle-DNA production by site specific recombination and protein-DNA interaction chromatography.* J Gene Med, 2008. **10**(11): p. 1253-69.
191. Simcikova, M., et al., *On the dual effect of glucose during production of pBAD/AraC-based minicircles.* Vaccine, 2014. **32**(24): p. 2843-6.
192. Chen, Z.Y., et al., *Minicircle DNA vectors devoid of bacterial DNA result in persistent and high-level transgene expression in vivo.* Mol Ther, 2003. **8**(3): p. 495-500.
193. Topalis, D., G. Andrei, and R. Snoeck, *The large tumor antigen: a "Swiss Army knife" protein possessing the functions required for the polyomavirus life cycle.* Antiviral Res, 2013. **97**(2): p. 122-36.
194. Liu, X., et al., *Merkel cell polyomavirus large T antigen disrupts lysosome clustering by translocating human Vam6p from the cytoplasm to the nucleus.* J Biol Chem, 2011. **286**(19): p. 17079-90.

195. Gao, J., et al., *Merkel cell polyomavirus T-antigens regulate DICER1 mRNA stability and translation through HSC70.* iScience, 2021. **24**(11): p. 103264.
196. Kumar, S., et al., *Merkel cell polyomavirus oncoproteins induce microRNAs that suppress multiple autophagy genes.* Int J Cancer, 2020. **146**(6): p. 1652-1666.
197. Kitagawa, K. and M. Kitagawa, *The SCF-type E3 Ubiquitin Ligases as Cancer Targets.* Curr Cancer Drug Targets, 2016. **16**(2): p. 119-29.
198. Nwogu, N., L.E. Ortiz, and H.J. Kwun, *Merkel Cell Polyomavirus Large T Antigen Unique Domain Regulates Its Own Protein Stability and Cell Growth.* Viruses, 2020. **12**(9).
199. Ortiz, L.E., A.M. Pham, and H.J. Kwun, *Identification of the Merkel Cell Polyomavirus Large Tumor Antigen Ubiquitin Conjugation Residue.* Int J Mol Sci, 2021. **22**(13).
200. Caplan, S., et al., *Human Vam6p promotes lysosome clustering and fusion in vivo.* J Cell Biol, 2001. **154**(1): p. 109-22.
201. Pham, A.M., et al., *Merkel Cell Polyomavirus Large T Antigen Induces Cellular Senescence for Host Growth Arrest and Viral Genome Persistence through Its Unique Domain.* Cells, 2023. **12**(3).
202. Borchert, S., et al., *High-affinity Rb binding, p53 inhibition, subcellular localization, and transformation by wild-type or tumor-derived shortened Merkel cell polyomavirus large T antigens.* J Virol, 2014. **88**(6): p. 3144-60.
203. Adams, P.D. and W.G. Kaelin, Jr., *Transcriptional control by E2F.* Semin Cancer Biol, 1995. **6**(2): p. 99-108.
204. Chellappan, S., et al., *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, 1992. **89**(10): p. 4549-53.
205. DeCaprio, J.A., *The role of the J domain of SV40 large T in cellular transformation.* Biologicals, 1999. **27**(1): p. 23-8.
206. Ali, S.H. and J.A. DeCaprio, *Cellular transformation by SV40 large T antigen: interaction with host proteins.* Semin Cancer Biol, 2001. **11**(1): p. 15-23.
207. Sullivan, C.S., A.E. Baker, and J.M. Pipas, *Simian virus 40 infection disrupts p130-E2F and p107-E2F complexes but does not perturb pRb-E2F complexes.* Virology, 2004. **320**(2): p. 218-28.
208. Arora, R., et al., *Survivin is a therapeutic target in Merkel cell carcinoma.* Sci Transl Med, 2012. **4**(133): p. 133ra56.
209. Houben, R., et al., *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, 2012. **130**(4): p. 847-56.
210. Houben, R., et al., *Merkel Cell Polyomavirus Large T Antigen is Dispensable in G2 and M-Phase to Promote Proliferation of Merkel Cell Carcinoma Cells.* Viruses, 2020. **12**(10).
211. Thanguturi, S., et al., *Investigation of the RB1-SOX2 axis constitutes a tool for viral status determination and diagnosis in Merkel cell carcinoma.* Virchows Arch, 2022. **480**(6): p. 1239-1254.
212. Spurgeon, M.E., et al., *Merkel cell polyomavirus large T antigen binding to pRb promotes skin hyperplasia and tumor development.* PLoS Pathog, 2022. **18**(5): p. e1010551.
213. Dresang, L.R., et al., *Response of Merkel cell polyomavirus-positive merkel cell carcinoma xenografts to a survivin inhibitor.* PLoS One, 2013. **8**(11): p. e80543.

214. Nakamura, T., et al., *Nuclear localization of Merkel cell polyomavirus large T antigen in Merkel cell carcinoma.* Virology, 2010. **398**(2): p. 273-9.
215. Tegtmeyer, P., et al., *Topography of simian virus 40 A protein-DNA complexes: arrangement of protein bound to the origin of replication.* J Virol, 1983. **46**(1): p. 151-61.
216. DeLucia, A.L., et al., *Topography of simian virus 40 A protein-DNA complexes: arrangement of pentanucleotide interaction sites at the origin of replication.* J Virol, 1983. **46**(1): p. 143-50.
217. Chang, C.F., H. Tada, and K. Khalili, *The role of a pentanucleotide repeat sequence, AGGGAAGGGA, in the regulation of JC virus DNA replication.* Gene, 1994. **148**(2): p. 309-14.
218. Deyerle, K.L., F.G. Sajjadi, and S. Subramani, *Analysis of origin of DNA replication of human papovavirus BK.* J Virol, 1989. **63**(1): p. 356-65.
219. Li, L., et al., *Sequences flanking the pentanucleotide T-antigen binding sites in the polyomavirus core origin help determine selectivity of DNA replication.* J Virol, 1995. **69**(12): p. 7570-8.
220. Rose, P.E. and B.S. Schaffhausen, *Zinc-binding and protein-protein interactions mediated by the polyomavirus large T antigen zinc finger.* J Virol, 1995. **69**(5): p. 2842-9.
221. Meinke, G., et al., *The crystal structure of the SV40 T-antigen origin binding domain in complex with DNA.* PLoS Biol, 2007. **5**(2): p. e23.
222. Boichuk, S., et al., *Multiple DNA damage signaling and repair pathways deregulated by simian virus 40 large T antigen.* J Virol, 2010. **84**(16): p. 8007-20.
223. Jiang, M., et al., *Roles of ATM and ATR-mediated DNA damage responses during lytic BK polyomavirus infection.* PLoS Pathog, 2012. **8**(8): p. e1002898.
224. Tsang, S.H., et al., *Host DNA damage response factors localize to merkel cell polyomavirus DNA replication sites to support efficient viral DNA replication.* J Virol, 2014. **88**(6): p. 3285-97.
225. Verhalen, B., et al., *Viral DNA replication-dependent DNA damage response activation during BK polyomavirus infection.* J Virol, 2015. **89**(9): p. 5032-9.
226. White, M.K., et al., *The DNA damage response promotes polyomavirus JC infection by nucleus to cytoplasm NF- kappaB activation.* Virol J, 2017. **14**(1): p. 31.
227. Erickson, K.D. and R.L. Garcea, *Viral replication centers and the DNA damage response in JC virus-infected cells.* Virology, 2019. **528**: p. 198-206.
228. Justice, J.L., J.M. Needham, and S.R. Thompson, *BK Polyomavirus Activates the DNA Damage Response To Prolong S Phase.* J Virol, 2019. **93**(14).
229. Justice, J.L., et al., *BK Polyomavirus Requires the Mismatch Repair Pathway for DNA Damage Response Activation.* J Virol, 2022. **96**(8): p. e0202821.
230. Siebels, S., et al., *Merkel Cell Polyomavirus DNA Replication Induces Senescence in Human Dermal Fibroblasts in a Kap1/Trim28-Dependent Manner.* mBio, 2020. **11**(2).
231. Marechal, A. and L. Zou, *DNA damage sensing by the ATM and ATR kinases.* Cold Spring Harb Perspect Biol, 2013. **5**(9).
232. Cheng, J., et al., *Merkel cell polyomavirus large T antigen has growth-promoting and inhibitory activities.* J Virol, 2013. **87**(11): p. 6118-26.
233. Li, J., et al., *Phosphorylation of Merkel cell polyomavirus large tumor antigen at serine 816 by ATM kinase induces apoptosis in host cells.* J Biol Chem, 2015. **290**(3): p. 1874-84.

234. Donati, B., E. Lorenzini, and A. Ciarrocchi, *BRD4 and Cancer: going beyond transcriptional regulation.* Mol Cancer, 2018. **17**(1): p. 164.
235. Wang, X., et al., *Bromodomain protein Brd4 plays a key role in Merkel cell polyomavirus DNA replication.* PLoS Pathog, 2012. **8**(11): p. e1003021.
236. Arora, R., A. Vats, and V. Chimankar, *MCV Truncated Large T antigen interacts with BRD4 in tumors.* Matters (Zur), 2019. **2019**.
237. Pozhidaeva, A. and I. Bezsonova, *USP7: Structure, substrate specificity, and inhibition.* DNA Repair (Amst), 2019. **76**: p. 30-39.
238. Czech-Sioli, M., et al., *The Ubiquitin-Specific Protease Usp7, a Novel Merkel Cell Polyomavirus Large T-Antigen Interaction Partner, Modulates Viral DNA Replication.* J Virol, 2020. **94**(5).
239. Rozenblatt-Rosen, O., et al., *Interpreting cancer genomes using systematic host network perturbations by tumour virus proteins.* Nature, 2012. **487**(7408): p. 491-5.
240. Shuda, M., et al., *Human Merkel cell polyomavirus small T antigen is an oncoprotein targeting the 4E-BP1 translation regulator.* J Clin Invest, 2011. **121**(9): p. 3623-34.
241. Welcker, M. and B.E. Clurman, *FBW7 ubiquitin ligase: a tumour suppressor at the crossroads of cell division, growth and differentiation.* Nat Rev Cancer, 2008. **8**(2): p. 83-93.
242. Shuda, M., et al., *CDK1 substitutes for mTOR kinase to activate mitotic cap-dependent protein translation.* Proc Natl Acad Sci U S A, 2015. **112**(19): p. 5875-82.
243. Velasquez, C., et al., *Mitotic protein kinase CDK1 phosphorylation of mRNA translation regulator 4E-BP1 Ser83 may contribute to cell transformation.* Proc Natl Acad Sci U S A, 2016. **113**(30): p. 8466-71.
244. Park, D.E., et al., *Dual inhibition of MDM2 and MDM4 in virus-positive Merkel cell carcinoma enhances the p53 response.* Proc Natl Acad Sci U S A, 2019. **116**(3): p. 1027-1032.
245. Bista, M., M. Petrovich, and A.R. Fersht, *MDMX contains an autoinhibitory sequence element.* Proc Natl Acad Sci U S A, 2013. **110**(44): p. 17814-9.
246. Chen, L., et al., *Autoinhibition of MDMX by intramolecular p53 mimicry.* Proc Natl Acad Sci U S A, 2015. **112**(15): p. 4624-9.
247. Bates, S., et al., *p14ARF links the tumour suppressors RB and p53.* Nature, 1998. **395**(6698): p. 124-5.
248. Allen, M.A., et al., *Global analysis of p53-regulated transcription identifies its direct targets and unexpected regulatory mechanisms.* Elife, 2014. **3**: p. e02200.
249. Nomura, K., et al., *Structural analysis of MDM2 RING separates degradation from regulation of p53 transcription activity.* Nat Struct Mol Biol, 2017. **24**(7): p. 578-587.
250. Rapchak, K., et al., *Merkel cell polyomavirus small T antigen is a viral transcription activator that is essential for viral genome maintenance.* PLoS Pathog, 2022. **18**(12): p. e1011039.
251. Wertz, I.E. and V.M. Dixit, *Signaling to NF-kappaB: regulation by ubiquitination.* Cold Spring Harb Perspect Biol, 2010. **2**(3): p. a003350.
252. Ramos, F., et al., *Role of protein phosphatases PP1, PP2A, PP4 and Cdc14 in the DNA damage response.* Cell Stress, 2019. **3**(3): p. 70-85.
253. Tsang, S.H., et al., *The Oncogenic Small Tumor Antigen of Merkel Cell Polyomavirus Is an Iron-Sulfur Cluster Protein That Enhances Viral DNA Replication.* J Virol, 2016. **90**(3): p. 1544-56.

254. Gupta, P., et al., *Merkel Cell Polyomavirus Downregulates N-myc Downstream-Regulated Gene 1, Leading to Cellular Proliferation and Migration.* J Virol, 2020. **94**(3).
255. Bayer, N.J., et al., *Structure of Merkel Cell Polyomavirus Capsid and Interaction with Its Glycosaminoglycan Attachment Receptor.* J Virol, 2020. **94**(20).
256. Lee, Y., et al., *MicroRNA maturation: stepwise processing and subcellular localization.* EMBO J, 2002. **21**(17): p. 4663-70.
257. Cai, X., C.H. Hagedorn, and B.R. Cullen, *Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs.* RNA, 2004. **10**(12): p. 1957-66.
258. Lee, Y., et al., *The nuclear RNase III Drosha initiates microRNA processing.* Nature, 2003. **425**(6956): p. 415-9.
259. Ketting, R.F., et al., *Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans.* Genes Dev, 2001. **15**(20): p. 2654-9.
260. Hutvagner, G., et al., *A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA.* Science, 2001. **293**(5531): p. 834-8.
261. Grishok, A., et al., *Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing.* Cell, 2001. **106**(1): p. 23-34.
262. Cullen, B.R., *Transcription and processing of human microRNA precursors.* Mol Cell, 2004. **16**(6): p. 861-5.
263. Hammond, S.M., et al., *An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells.* Nature, 2000. **404**(6775): p. 293-6.
264. Lewis, B.P., et al., *Prediction of mammalian microRNA targets.* Cell, 2003. **115**(7): p. 787-98.
265. Brennecke, J., et al., *Principles of microRNA-target recognition.* PLoS Biol, 2005. **3**(3): p. e85.
266. Lewis, B.P., C.B. Burge, and D.P. Bartel, *Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets.* Cell, 2005. **120**(1): p. 15-20.
267. Bartel, D.P., *MicroRNAs: target recognition and regulatory functions.* Cell, 2009. **136**(2): p. 215-33.
268. Skalsky, R.L., et al., *Kaposi's sarcoma-associated herpesvirus encodes an ortholog of miR-155.* J Virol, 2007. **81**(23): p. 12836-45.
269. Abend, J.R., T. Uldrick, and J.M. Ziegelbauer, *Regulation of tumor necrosis factor-like weak inducer of apoptosis receptor protein (TWEAKR) expression by Kaposi's sarcoma-associated herpesvirus microRNA prevents TWEAK-induced apoptosis and inflammatory cytokine expression.* J Virol, 2010. **84**(23): p. 12139-51.
270. Suffert, G., et al., *Kaposi's sarcoma herpesvirus microRNAs target caspase 3 and regulate apoptosis.* PLoS Pathog, 2011. **7**(12): p. e1002405.
271. Liu, X., C. Happel, and J.M. Ziegelbauer, *Kaposi's Sarcoma-Associated Herpesvirus MicroRNAs Target GADD45B To Protect Infected Cells from Cell Cycle Arrest and Apoptosis.* J Virol, 2017. **91**(3).
272. Choy, E.Y., et al., *An Epstein-Barr virus-encoded microRNA targets PUMA to promote host cell survival.* J Exp Med, 2008. **205**(11): p. 2551-60.
273. Marquitz, A.R., et al., *The Epstein-Barr Virus BART microRNAs target the pro-apoptotic protein Bim.* Virology, 2011. **412**(2): p. 392-400.

274. Shinozaki-Ushiku, A., et al., *Profiling of Virus-Encoded MicroRNAs in Epstein-Barr Virus-Associated Gastric Carcinoma and Their Roles in Gastric Carcinogenesis.* J Virol, 2015. **89**(10): p. 5581-91.
275. Kim, H., H. Choi, and S.K. Lee, *Epstein-Barr Virus MicroRNA miR-BART20-5p Suppresses Lytic Induction by Inhibiting BAD-Mediated caspase-3-Dependent Apoptosis.* J Virol, 2016. **90**(3): p. 1359-68.
276. Liang, D., et al., *A human herpesvirus miRNA attenuates interferon signaling and contributes to maintenance of viral latency by targeting IKKepsilon.* Cell Res, 2011. **21**(5): p. 793-806.
277. Abend, J.R., et al., *Kaposi's sarcoma-associated herpesvirus microRNAs target IRAK1 and MYD88, two components of the toll-like receptor/interleukin-1R signaling cascade, to reduce inflammatory-cytokine expression.* J Virol, 2012. **86**(21): p. 11663-74.
278. Ramalingam, D. and J.M. Ziegelbauer, *Viral microRNAs Target a Gene Network, Inhibit STAT Activation, and Suppress Interferon Responses.* Sci Rep, 2017. **7**: p. 40813.
279. Lu, Y., et al., *Epstein-Barr Virus miR-BART6-3p Inhibits the RIG-I Pathway.* J Innate Immun, 2017. **9**(6): p. 574-586.
280. Bouvet, M., et al., *Multiple Viral microRNAs Regulate Interferon Release and Signaling Early during Infection with Epstein-Barr Virus.* mBio, 2021. **12**(2).
281. Nachmani, D., et al., *Diverse herpesvirus microRNAs target the stress-induced immune ligand MICB to escape recognition by natural killer cells.* Cell Host Microbe, 2009. **5**(4): p. 376-85.
282. Bellare, P. and D. Ganem, *Regulation of KSHV lytic switch protein expression by a virus-encoded microRNA: an evolutionary adaptation that fine-tunes lytic reactivation.* Cell Host Microbe, 2009. **6**(6): p. 570-5.
283. Lin, X., et al., *miR-K12-7-5p encoded by Kaposi's sarcoma-associated herpesvirus stabilizes the latent state by targeting viral ORF50/RTA.* PLoS One, 2011. **6**(1): p. e16224.
284. Feederle, R., et al., *The members of an Epstein-Barr virus microRNA cluster cooperate to transform B lymphocytes.* J Virol, 2011. **85**(19): p. 9801-10.
285. Moody, R., et al., *KSHV microRNAs mediate cellular transformation and tumorigenesis by redundantly targeting cell growth and survival pathways.* PLoS Pathog, 2013. **9**(12): p. e1003857.
286. Bernhardt, K., et al., *A Viral microRNA Cluster Regulates the Expression of PTEN, p27 and of a bcl-2 Homolog.* PLoS Pathog, 2016. **12**(1): p. e1005405.
287. Cai, L., et al., *Epstein-Barr virus-encoded microRNA BART1 induces tumour metastasis by regulating PTEN-dependent pathways in nasopharyngeal carcinoma.* Nat Commun, 2015. **6**: p. 7353.
288. Chen, Y., et al., *Epstein-Barr virus microRNAs regulate B cell receptor signal transduction and lytic reactivation.* PLoS Pathog, 2019. **15**(1): p. e1007535.
289. Li, W., et al., *A KSHV microRNA enhances viral latency and induces angiogenesis by targeting GRK2 to activate the CXCR2/AKT pathway.* Oncotarget, 2016. **7**(22): p. 32286-305.
290. Skalsky, R.L., et al., *The viral and cellular microRNA targetome in lymphoblastoid cell lines.* PLoS Pathog, 2012. **8**(1): p. e1002484.
291. Lei, X., et al., *A Kaposi's sarcoma-associated herpesvirus microRNA and its variants target the transforming growth factor beta pathway to promote cell survival.* J Virol, 2012. **86**(21): p. 11698-711.

292. Yang, L., J.E. Wilusz, and L.L. Chen, *Biogenesis and Regulatory Roles of Circular RNAs.* Annu Rev Cell Dev Biol, 2022. **38**: p. 263-289.
293. Toptan, T., et al., *Circular DNA tumor viruses make circular RNAs.* Proc Natl Acad Sci U S A, 2018. **115**(37): p. E8737-E8745.
294. Abere, B., et al., *Kaposi's Sarcoma-Associated Herpesvirus-Encoded circRNAs Are Expressed in Infected Tumor Tissues and Are Incorporated into Virions.* mBio, 2020. **11**(1).
295. Tagawa, T., et al., *Discovery of Kaposi's sarcoma herpesvirus-encoded circular RNAs and a human antiviral circular RNA.* Proc Natl Acad Sci U S A, 2018. **115**(50): p. 12805-12810.
296. Ungerleider, N., et al., *The Epstein Barr virus circRNAome.* PLoS Pathog, 2018. **14**(8): p. e1007206.
297. Huang, J.T., et al., *Identification of virus-encoded circular RNA.* Virology, 2019. **529**: p. 144-151.
298. Zhang, J.Y., et al., *ebv-circRPMS1 promotes the progression of EBV-associated gastric carcinoma via Sam68-dependent activation of METTL3.* Cancer Lett, 2022. **535**: p. 215646.
299. Du, Y., et al., *Hypoxia-induced ebv-circLMP2A promotes angiogenesis in EBV-associated gastric carcinoma through the KHSRP/VHL/HIF1alpha/VEGFA pathway.* Cancer Lett, 2022. **526**: p. 259-272.
300. Tan, K.E., et al., *Identification and characterization of a novel Epstein-Barr Virus-encoded circular RNA from LMP-2 Gene.* Sci Rep, 2021. **11**(1): p. 14392.
301. Zhao, J., et al., *Transforming activity of an oncoprotein-encoding circular RNA from human papillomavirus.* Nat Commun, 2019. **10**(1): p. 2300.
302. Zhu, M., et al., *Hepatocellular carcinoma progression mediated by hepatitis B virus-encoded circRNA HBV_circ_1 through interaction with CDK1.* Mol Ther Nucleic Acids, 2021. **25**: p. 668-682.
303. Chang, Y., et al., *Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma.* Science, 1994. **266**(5192): p. 1865-9.
304. Cesarman, E., et al., *Kaposi's sarcoma-associated herpesvirus-like DNA sequences in AIDS-related body-cavity-based lymphomas.* N Engl J Med, 1995. **332**(18): p. 1186-91.
305. Soulier, J., et al., *Kaposi's sarcoma-associated herpesvirus-like DNA sequences in multicentric Castleman's disease.* Blood, 1995. **86**(4): p. 1276-80.
306. Sarid, R., et al., *Transcription mapping of the Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8) genome in a body cavity-based lymphoma cell line (BC-1).* J Virol, 1998. **72**(2): p. 1005-12.
307. Dupin, N., et al., *Distribution of human herpesvirus-8 latently infected cells in Kaposi's sarcoma, multicentric Castleman's disease, and primary effusion lymphoma.* Proc Natl Acad Sci U S A, 1999. **96**(8): p. 4546-51.
308. Staskus, K.A., et al., *Kaposi's sarcoma-associated herpesvirus gene expression in endothelial (spindle) tumor cells.* J Virol, 1997. **71**(1): p. 715-9.
309. Orenstein, J.M., et al., *Visualization of human herpesvirus type 8 in Kaposi's sarcoma by light and transmission electron microscopy.* AIDS, 1997. **11**(5): p. F35-45.
310. Parravicini, C., et al., *Differential viral protein expression in Kaposi's sarcoma-associated herpesvirus-infected diseases: Kaposi's sarcoma, primary effusion lymphoma, and multicentric Castleman's disease.* Am J Pathol, 2000. **156**(3): p. 743-9.

311. Staskus, K.A., et al., *Cellular tropism and viral interleukin-6 expression distinguish human herpesvirus 8 involvement in Kaposi's sarcoma, primary effusion lymphoma, and multicentric Castleman's disease.* J Virol, 1999. **73**(5): p. 4181-7.
312. Katano, H., et al., *Expression and localization of human herpesvirus 8-encoded proteins in primary effusion lymphoma, Kaposi's sarcoma, and multicentric Castleman's disease.* Virology, 2000. **269**(2): p. 335-44.
313. Hosseinipour, M.C., et al., *Viral profiling identifies multiple subtypes of Kaposi's sarcoma.* MBio, 2014. **5**(5): p. e01633-14.
314. Lin, Y.T., et al., *Small RNA profiling reveals antisense transcription throughout the KSHV genome and novel small RNAs.* RNA, 2010. **16**(8): p. 1540-58.
315. Cai, X., et al., *Kaposi's sarcoma-associated herpesvirus expresses an array of viral microRNAs in latently infected cells.* Proc Natl Acad Sci U S A, 2005. **102**(15): p. 5570-5.
316. Samols, M.A., et al., *Cloning and identification of a microRNA cluster within the latency-associated region of Kaposi's sarcoma-associated herpesvirus.* J Virol, 2005. **79**(14): p. 9301-5.
317. Sullivan, C.S., *High conservation of Kaposi sarcoma--associated herpesvirus microRNAs implies important function.* J Infect Dis, 2007. **195**(5): p. 618-20.
318. Marshall, V., et al., *Conservation of virally encoded microRNAs in Kaposi sarcoma--associated herpesvirus in primary effusion lymphoma cell lines and in patients with Kaposi sarcoma or multicentric Castleman disease.* J Infect Dis, 2007. **195**(5): p. 645-59.
319. Hansen, A., et al., *KSHV-encoded miRNAs target MAF to induce endothelial cell reprogramming.* Genes Dev, 2010. **24**(2): p. 195-205.
320. Lei, X., et al., *Regulation of herpesvirus lifecycle by viral microRNAs.* Virulence, 2010. **1**(5): p. 433-5.
321. Lei, X., et al., *Regulation of NF-kappaB inhibitor IkappaBalpha and viral replication by a KSHV microRNA.* Nat Cell Biol, 2010. **12**(2): p. 193-9.
322. Gottwein, E., et al., *Viral microRNA targetome of KSHV-infected primary effusion lymphoma cell lines.* Cell Host Microbe, 2011. **10**(5): p. 515-26.
323. Qin, Z., et al., *Upregulation of xCT by KSHV-encoded microRNAs facilitates KSHV dissemination and persistence in an environment of oxidative stress.* PLoS Pathog, 2010. **6**(1): p. e1000742.
324. Arias, C., et al., *KSHV 2.0: a comprehensive annotation of the Kaposi's sarcoma-associated herpesvirus genome using next-generation sequencing reveals novel genomic and functional features.* PLoS Pathog, 2014. **10**(1): p. e1003847.
325. Chandriani, S., Y. Xu, and D. Ganem, *The lytic transcriptome of Kaposi's sarcoma-associated herpesvirus reveals extensive transcription of noncoding regions, including regions antisense to important genes.* J Virol, 2010. **84**(16): p. 7934-42.
326. Sun, R., et al., *Polyadenylylated nuclear RNA encoded by Kaposi sarcoma-associated herpesvirus.* Proc Natl Acad Sci U S A, 1996. **93**(21): p. 11883-8.
327. Dresang, L.R., et al., *Coupled transcriptome and proteome analysis of human lymphotropic tumor viruses: insights on the detection and discovery of viral genes.* BMC Genomics, 2011. **12**: p. 625.
328. Zhong, W., et al., *Restricted expression of Kaposi sarcoma-associated herpesvirus (human herpesvirus 8) genes in Kaposi sarcoma.* Proc Natl Acad Sci U S A, 1996. **93**(13): p. 6641-6.

329. Taylor, J.L., et al., *Transcriptional analysis of latent and inducible Kaposi's sarcoma-associated herpesvirus transcripts in the K4 to K7 region.* J Virol, 2005. **79**(24): p. 15099-106.
330. Majerciak, V., et al., *A viral genome landscape of RNA polyadenylation from KSHV latent to lytic infection.* PLoS Pathog, 2013. **9**(11): p. e1003749.
331. Wang, Y., et al., *Kaposi's sarcoma-associated herpesvirus ori-Lyt-dependent DNA replication: cis-acting requirements for replication and ori-Lyt-associated RNA transcription.* J Virol, 2004. **78**(16): p. 8615-29.
332. Schifano, J.M., et al., *Expression of the Antisense-to-Latency Transcript Long Noncoding RNA in Kaposi's Sarcoma-Associated Herpesvirus.* J Virol, 2017. **91**(4).
333. Jeck, W.R., et al., *Circular RNAs are abundant, conserved, and associated with ALU repeats.* RNA, 2013. **19**(2): p. 141-57.
334. Holdt, L.M., A. Kohlmaier, and D. Teupser, *Molecular roles and function of circular RNAs in eukaryotic cells.* Cell Mol Life Sci, 2018. **75**(6): p. 1071-1098.
335. Hansen, T.B., et al., *Natural RNA circles function as efficient microRNA sponges.* Nature, 2013. **495**(7441): p. 384-8.
336. Memczak, S., et al., *Circular RNAs are a large class of animal RNAs with regulatory potency.* Nature, 2013. **495**(7441): p. 333-8.
337. Ashwal-Fluss, R., et al., *circRNA biogenesis competes with pre-mRNA splicing.* Mol Cell, 2014. **56**(1): p. 55-66.
338. Abdelmohsen, K., et al., *Identification of HuR target circular RNAs uncovers suppression of PABPN1 translation by CircPABPN1.* RNA Biol, 2017. **14**(3): p. 361-369.
339. Du, W.W., et al., *Foxo3 circular RNA retards cell cycle progression via forming ternary complexes with p21 and CDK2.* Nucleic Acids Res, 2016. **44**(6): p. 2846-58.
340. Li, Z., et al., *Exon-intron circular RNAs regulate transcription in the nucleus.* Nat Struct Mol Biol, 2015. **22**(3): p. 256-64.
341. Zhang, Y., et al., *Circular intronic long noncoding RNAs.* Mol Cell, 2013. **51**(6): p. 792-806.
342. Legnini, I., et al., *Circ-ZNF609 Is a Circular RNA that Can Be Translated and Functions in Myogenesis.* Mol Cell, 2017. **66**(1): p. 22-37 e9.
343. Pamudurti, N.R., et al., *Translation of CircRNAs.* Mol Cell, 2017. **66**(1): p. 9-21 e7.
344. Li, X., et al., *Coordinated circRNA Biogenesis and Function with NF90/NF110 in Viral Infection.* Mol Cell, 2017. **67**(2): p. 214-227 e7.
345. Liu, C.X., et al., *Structure and Degradation of Circular RNAs Regulate PKR Activation in Innate Immunity.* Cell, 2019. **177**(4): p. 865-880 e21.
346. Xia, P., et al., *A Circular RNA Protects Dormant Hematopoietic Stem Cells from DNA Sensor cGAS-Mediated Exhaustion.* Immunity, 2018. **48**(4): p. 688-701 e7.
347. Enuka, Y., et al., *Circular RNAs are long-lived and display only minimal early alterations in response to a growth factor.* Nucleic Acids Res, 2016. **44**(3): p. 1370-83.
348. Zhang, Y., et al., *Circular RNAs: emerging cancer biomarkers and targets.* J Exp Clin Cancer Res, 2017. **36**(1): p. 152.
349. Maass, P.G., et al., *A map of human circular RNAs in clinically relevant tissues.* J Mol Med (Berl), 2017. **95**(11): p. 1179-1189.
350. Ungerleider, N.A., et al., *Comparative Analysis of Gammaherpesvirus Circular RNA Repertoires: Conserved and Unique Viral Circular RNAs.* J Virol, 2019. **93**(6).

351. Sathish, N., X. Wang, and Y. Yuan, *Tegument Proteins of Kaposi's Sarcoma-Associated Herpesvirus and Related Gamma-Herpesviruses.* Front Microbiol, 2012. **3**: p. 98.
352. Cannon, J.S., et al., *A new primary effusion lymphoma-derived cell line yields a highly infectious Kaposi's sarcoma herpesvirus-containing supernatant.* J Virol, 2000. **74**(21): p. 10187-93.
353. Renne, R., et al., *Lytic growth of Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8) in culture.* Nat Med, 1996. **2**(3): p. 342-6.
354. Arvanitakis, L., et al., *Establishment and characterization of a primary effusion (body cavity-based) lymphoma cell line (BC-3) harboring kaposi's sarcoma-associated herpesvirus (KSHV/HHV-8) in the absence of Epstein-Barr virus.* Blood, 1996. **88**(7): p. 2648-54.
355. Gao, S.J., et al., *KSHV antibodies among Americans, Italians and Ugandans with and without Kaposi's sarcoma.* Nat Med, 1996. **2**(8): p. 925-8.
356. Menezes, J., et al., *Establishment and characterization of an Epstein-Barr virus (EBC)-negative lymphoblastoid B cell line (BJA-B) from an exceptional, EBV-genome-negative African Burkitt's lymphoma.* Biomedicine, 1975. **22**(4): p. 276-84.
357. Morand, P., et al., *Human herpesvirus 8 and Epstein Barr-virus in a cutaneous B-cell lymphoma and a malignant cell line established from the blood of an AIDS patient.* Leuk Lymphoma, 1999. **35**(3-4): p. 379-87.
358. Kati, S., et al., *Generation of high-titre virus stocks using BrK.219, a B-cell line infected stably with recombinant Kaposi's sarcoma-associated herpesvirus.* J Virol Methods, 2015. **217**: p. 79-86.
359. Kati, S., et al., *Activation of the B cell antigen receptor triggers reactivation of latent Kaposi's sarcoma-associated herpesvirus in B cells.* J Virol, 2013. **87**(14): p. 8004-16.
360. Vieira, J. and P.M. O'Hearn, *Use of the red fluorescent protein as a marker of Kaposi's sarcoma-associated herpesvirus lytic gene expression.* Virology, 2004. **325**(2): p. 225-40.
361. Abere, B., et al., *The Kaposi's sarcoma-associated herpesvirus (KSHV) non-structural membrane protein K15 is required for viral lytic replication and may represent a therapeutic target.* PLoS Pathog, 2017. **13**(9): p. e1006639.
362. Abere, B., et al., *Kaposi's Sarcoma-Associated Herpesvirus Nonstructural Membrane Protein pK15 Recruits the Class II Phosphatidylinositol 3-Kinase PI3K-C2alpha To Activate Productive Viral Replication.* J Virol, 2018. **92**(17).
363. Russo, J.J., et al., *Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus (HHV8).* Proc Natl Acad Sci U S A, 1996. **93**(25): p. 14862-7.
364. Nealon, K., et al., *Lytic replication of Kaposi's sarcoma-associated herpesvirus results in the formation of multiple capsid species: isolation and molecular characterization of A, B, and C capsids from a gammaherpesvirus.* J Virol, 2001. **75**(6): p. 2866-78.
365. Lin, X., et al., *MicroRNAs and unusual small RNAs discovered in Kaposi's sarcoma-associated herpesvirus virions.* J Virol, 2012. **86**(23): p. 12717-30.
366. Bechtel, J., A. Grundhoff, and D. Ganem, *RNAs in the virion of Kaposi's sarcoma-associated herpesvirus.* J Virol, 2005. **79**(16): p. 10138-46.
367. Wang, X., et al., *Mono-ubiquitylated ORF45 Mediates Association of KSHV Particles with Internal Lipid Rafts for Viral Assembly and Egress.* PLoS Pathog, 2015. **11**(12): p. e1005332.
368. Rossetto, C.C. and G.S. Pari, *Kaposi's sarcoma-associated herpesvirus noncoding polyadenylated nuclear RNA interacts with virus- and host cell-encoded proteins and*

suppresses expression of genes involved in immune modulation. J Virol, 2011. **85**(24): p. 13290-7.
369. Rossetto, C.C., et al., *Regulation of viral and cellular gene expression by Kaposi's sarcoma-associated herpesvirus polyadenylated nuclear RNA.* J Virol, 2013. **87**(10): p. 5540-53.
370. Rossetto, C.C. and G. Pari, *KSHV PAN RNA associates with demethylases UTX and JMJD3 to activate lytic replication through a physical interaction with the virus genome.* PLoS Pathog, 2012. **8**(5): p. e1002680.
371. Withers, J.B., et al., *Two herpesviral noncoding PAN RNAs are functionally homologous but do not associate with common chromatin loci.* PLoS Pathog, 2018. **14**(11): p. e1007389.
372. Lee, H.R., et al., *Viral interferon regulatory factors.* J Interferon Cytokine Res, 2009. **29**(9): p. 621-7.
373. Jacobs, S.R. and B. Damania, *The viral interferon regulatory factors of KSHV: immunosuppressors or oncogenes?* Front Immunol, 2011. **2**: p. 19.
374. Rivas, C., et al., *Kaposi's sarcoma-associated herpesvirus LANA2 is a B-cell-specific latent viral protein that inhibits p53.* J Virol, 2001. **75**(1): p. 429-38.
375. Kanno, T., et al., *Expression of Kaposi's sarcoma-associated herpesvirus-encoded K10/10.1 protein in tissues and its interaction with poly(A)-binding protein.* Virology, 2006. **352**(1): p. 100-9.
376. Lallemand, F., et al., *Quantitative analysis of human herpesvirus 8 viral load using a real-time PCR assay.* J Clin Microbiol, 2000. **38**(4): p. 1404-8.
377. Drexler, H.G., et al., *Lymphoma cell lines: in vitro models for the study of HHV-8+ primary effusion lymphomas (body cavity-based lymphomas).* Leukemia, 1998. **12**(10): p. 1507-17.
378. Rossetto, C.C. and G.S. Pari, *PAN's Labyrinth: Molecular biology of Kaposi's sarcoma-associated herpesvirus (KSHV) PAN RNA, a multifunctional long noncoding RNA.* Viruses, 2014. **6**(11): p. 4212-26.
379. Borah, S., et al., *A viral nuclear noncoding RNA binds re-localized poly(A) binding protein and is required for late KSHV gene expression.* PLoS Pathog, 2011. **7**(10): p. e1002300.
380. Campbell, M., H.J. Kung, and Y. Izumiya, *Long non-coding RNA and epigenetic gene regulation of KSHV.* Viruses, 2014. **6**(11): p. 4165-77.
381. Campbell, M., et al., *A lytic viral long noncoding RNA modulates the function of a latent protein.* J Virol, 2014. **88**(3): p. 1843-8.
382. Szabo, L. and J. Salzman, *Detecting circular RNAs: bioinformatic and experimental challenges.* Nat Rev Genet, 2016. **17**(11): p. 679-692.
383. Knight, L.M., et al., *Merkel cell polyomavirus small T antigen mediates microtubule destabilization to promote cell motility and migration.* J Virol, 2015. **89**(1): p. 35-47.
384. Berrios, C., et al., *Merkel Cell Polyomavirus Small T Antigen Promotes Pro-Glycolytic Metabolic Perturbations Required for Transformation.* PLoS Pathog, 2016. **12**(11): p. e1006020.
385. Shukla, G.C., J. Singh, and S. Barik, *MicroRNAs: Processing, Maturation, Target Recognition and Regulatory Functions.* Mol Cell Pharmacol, 2011. **3**(3): p. 83-92.
386. Gebert, L.F.R. and I.J. MacRae, *Regulation of microRNA function in animals.* Nat Rev Mol Cell Biol, 2019. **20**(1): p. 21-37.
387. Salzman, J., et al., *Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types.* PLoS One, 2012. **7**(2): p. e30733.

388. Hansen, T.B., et al., *miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA.* EMBO J, 2011. **30**(21): p. 4414-22.
389. Cocquerelle, C., et al., *Mis-splicing yields circular RNA molecules.* FASEB J, 1993. **7**(1): p. 155-60.
390. Szabo, L., et al., *Statistically based splicing detection reveals neural enrichment and tissue-specific induction of circular RNA during human fetal development.* Genome Biol, 2015. **16**: p. 126.
391. Salzman, J., et al., *Cell-type specific features of circular RNA expression.* PLoS Genet, 2013. **9**(9): p. e1003777.
392. Piwecka, M., et al., *Loss of a mammalian circular RNA locus causes miRNA deregulation and affects brain function.* Science, 2017. **357**(6357).
393. Memczak, S., et al., *Identification and Characterization of Circular RNAs As a New Class of Putative Biomarkers in Human Blood.* PLoS One, 2015. **10**(10): p. e0141214.
394. Hansen, T.B., J. Kjems, and C.K. Damgaard, *Circular RNA and miR-7 in cancer.* Cancer Res, 2013. **73**(18): p. 5609-12.
395. Gong, L.P., et al., *Epstein-Barr virus-derived circular RNA LMP2A induces stemness in EBV-associated gastric cancer.* EMBO Rep, 2020: p. e49689.
396. Rigatti, L.H., et al., *Identification and Characterization of Novel Rat Polyomavirus 2 in a Colony of X-SCID Rats by P-PIT assay.* mSphere, 2016. **1**(6).
397. Schrama, D., et al., *Characterization of six Merkel cell polyomavirus-positive Merkel cell carcinoma cell lines: Integration pattern suggest that large T antigen truncating events occur before or during integration.* Int J Cancer, 2019. **145**(4): p. 1020-1032.
398. Rosen, S.T., et al., *Establishment and characterization of a neuroendocrine skin carcinoma cell line.* Lab Invest, 1987. **56**(3): p. 302-12.
399. Van Gele, M., et al., *Combined karyotyping, CGH and M-FISH analysis allows detailed characterization of unidentified chromosomal rearrangements in Merkel cell carcinoma.* Int J Cancer, 2002. **101**(2): p. 137-45.
400. Velasquez, C., et al., *Characterization of a Merkel Cell Polyomavirus-Positive Merkel Cell Carcinoma Cell Line CVG-1.* Front Microbiol, 2018. **9**: p. 713.
401. Guastafierro, A., et al., *Characterization of an early passage Merkel cell polyomavirus-positive Merkel cell carcinoma cell line, MS-1, and its growth in NOD scid gamma mice.* J Virol Methods, 2013. **187**(1): p. 6-14.
402. Kramer, M.C., et al., *Combinatorial control of Drosophila circular RNA expression by intronic repeats, hnRNPs, and SR proteins.* Genes Dev, 2015. **29**(20): p. 2168-82.
403. Henno, L., et al., *Analysis of Human Papillomavirus Genome Replication Using Two- and Three-Dimensional Agarose Gel Electrophoresis.* Curr Protoc Microbiol, 2017. **45**: p. 14B 10 1-14B 10 37.
404. Porter, S.S. and A.A. McBride, *Human Papillomavirus Quasivirus Production and Infection of Primary Human Keratinocytes.* Curr Protoc Microbiol, 2020. **57**(1): p. e101.
405. Gao, Y., J. Zhang, and F. Zhao, *Circular RNA identification based on multiple seed matching.* Brief Bioinform, 2017.
406. Panda, A.C., et al., *High-purity circular RNA isolation method (RPAD) reveals vast collection of intronic circRNAs.* Nucleic Acids Res, 2017. **45**(12): p. e116.
407. Adami, G.R., et al., *Leader-to-leader splicing is required for efficient production and accumulation of polyomavirus late mRNAs.* J Virol, 1989. **63**(1): p. 85-93.

408. Rodig, S.J., et al., *Improved detection suggests all Merkel cell carcinomas harbor Merkel polyomavirus.* J Clin Invest, 2012. **122**(12): p. 4645-53.
409. Yang, Y., et al., *Extensive translation of circular RNAs driven by N(6)-methyladenosine.* Cell Res, 2017. **27**(5): p. 626-641.
410. Kleaveland, B., et al., *A Network of Noncoding Regulatory RNAs Acts in the Mammalian Brain.* Cell, 2018. **174**(2): p. 350-362 e17.
411. Chen, C.J., et al., *Identification of a polyomavirus microRNA highly expressed in tumors.* Virology, 2015. **476**: p. 43-53.
412. Ameres, S.L., et al., *Target RNA-directed trimming and tailing of small silencing RNAs.* Science, 2010. **328**(5985): p. 1534-9.
413. Cazalla, D., T. Yario, and J.A. Steitz, *Down-regulation of a host microRNA by a Herpesvirus saimiri noncoding RNA.* Science, 2010. **328**(5985): p. 1563-6.
414. Ghini, F., et al., *Endogenous transcripts control miRNA levels and activity in mammalian cells by target-directed miRNA degradation.* Nat Commun, 2018. **9**(1): p. 3119.
415. Engels, E.A., et al., *Merkel cell carcinoma and HIV infection.* Lancet, 2002. **359**(9305): p. 497-8.
416. Lemos, B.D., et al., *Pathologic nodal evaluation improves prognostic accuracy in Merkel cell carcinoma: analysis of 5823 cases as the basis of the first consensus staging system.* J Am Acad Dermatol, 2010. **63**(5): p. 751-61.
417. Wong, S.Q., et al., *UV-Associated Mutations Underlie the Etiology of MCV-Negative Merkel Cell Carcinomas.* Cancer Res, 2015. **75**(24): p. 5228-34.
418. Goh, G., et al., *Mutational landscape of MCPyV-positive and MCPyV-negative Merkel cell carcinomas with implications for immunotherapy.* Oncotarget, 2016. **7**(3): p. 3403-15.
419. Starrett, G.J., et al., *Merkel Cell Polyomavirus Exhibits Dominant Control of the Tumor Genome and Transcriptome in Virus-Associated Merkel Cell Carcinoma.* mBio, 2017. **8**(1).
420. Knepper, T.C., et al., *The Genomic Landscape of Merkel Cell Carcinoma and Clinicogenomic Biomarkers of Response to Immune Checkpoint Inhibitor Therapy.* Clin Cancer Res, 2019. **25**(19): p. 5961-5971.
421. Starrett, G.J., et al., *Clinical and molecular characterization of virus-positive and virus-negative Merkel cell carcinoma.* Genome Med, 2020. **12**(1): p. 30.
422. Tolstov, Y.L., et al., *Asymptomatic primary Merkel cell polyomavirus infection among adults.* Emerg Infect Dis, 2011. **17**(8): p. 1371-80.
423. Schowalter, R.M., et al., *Merkel cell polyomavirus and two previously unknown polyomaviruses are chronically shed from human skin.* Cell Host Microbe, 2010. **7**(6): p. 509-15.
424. Schowalter, R.M., D.V. Pastrana, and C.B. Buck, *Glycosaminoglycans and sialylated glycans sequentially facilitate Merkel cell polyomavirus infectious entry.* PLoS Pathog, 2011. **7**(7): p. e1002161.
425. Schowalter, R.M., W.C. Reinhold, and C.B. Buck, *Entry tropism of BK and Merkel cell polyomaviruses in cell culture.* PLoS One, 2012. **7**(7): p. e42181.
426. Scheller, A. and C. Prives, *Simian virus 40 and polyomavirus large tumor antigens have different requirements for high-affinity sequence-specific DNA binding.* J Virol, 1985. **54**(2): p. 532-45.
427. Dobson, S.J., et al., *Merkel cell polyomavirus small tumour antigen activates the p38 MAPK pathway to enhance cellular motility.* Biochem J, 2020. **477**(14): p. 2721-2733.

428. Prives, C., et al., *DNA sequence requirements for replication of polyomavirus DNA in vivo and in vitro.* Mol Cell Biol, 1987. **7**(10): p. 3694-704.
429. Deb, S., et al., *The adenine-thymine domain of the simian virus 40 core origin directs DNA bending and coordinately regulates DNA replication.* Mol Cell Biol, 1986. **6**(12): p. 4578-84.
430. Gannon, J.V. and D.P. Lane, *Interactions between SV40 T antigen and DNA polymerase alpha.* New Biol, 1990. **2**(1): p. 84-92.
431. Onwubiko, N.O., et al., *SV40 T antigen interactions with ssDNA and replication protein A: a regulatory role of T antigen monomers in lagging strand DNA replication.* Nucleic Acids Res, 2020. **48**(7): p. 3657-3677.
432. Myers, R.M., et al., *SV40 gene expression is modulated by the cooperative binding of T antigen to DNA.* Cell, 1981. **25**(2): p. 373-84.
433. Rio, D.C. and R. Tjian, *SV40 T antigen binding site mutations that affect autoregulation.* Cell, 1983. **32**(4): p. 1227-40.
434. Alvarez Orellana, J., et al., *Sirolimus and Other Mechanistic Target of Rapamycin Inhibitors Directly Activate Latent Pathogenic Human Polyomavirus Replication.* J Infect Dis, 2021. **224**(7): p. 1160-1169.
435. Lee, S., et al., *Identification and validation of a novel mature microRNA encoded by the Merkel cell polyomavirus in human Merkel cell carcinomas.* J Clin Virol, 2011. **52**(3): p. 272-5.
436. Nwogu, N., L.E. Ortiz, and H.J. Kwun, *Surface charge of Merkel cell polyomavirus small T antigen determines cell transformation through allosteric FBW7 WD40 domain targeting.* Oncogenesis, 2020. **9**(5): p. 53.
437. Dye, K.N., et al., *Merkel cell polyomavirus Tumor antigens expressed in Merkel cell carcinoma function independently of the ubiquitin ligases Fbw7 and beta-TrCP.* PLoS Pathog, 2019. **15**(1): p. e1007543.
438. Lee, M.S., et al., *Human Mesenchymal Stem Cells of Diverse Origins Support Persistent Infection with Kaposi's Sarcoma-Associated Herpesvirus and Manifest Distinct Angiogenic, Invasive, and Transforming Phenotypes.* mBio, 2016. **7**(1): p. e02109-15.
439. Abere, B., et al., *Merkel Cell Polyomavirus Encodes Circular RNAs (circRNAs) Enabling a Dynamic circRNA/microRNA/mRNA Regulatory Network.* mBio, 2020. **11**(6): p. 1-20.
440. Guo, X., et al., *The recombined cccDNA produced using minicircle technology mimicked HBV genome in structure and function closely.* Sci Rep, 2016. **6**: p. 25552.
441. Kim, J.H., et al., *High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice.* PLoS One, 2011. **6**(4): p. e18556.
442. Revie, D., et al., *Covalent association of protein with replicative form DNA of parvovirus H-1.* Proc Natl Acad Sci U S A, 1979. **76**(11): p. 5539-43.
443. Shaw, G., et al., *Preferential transformation of human neuronal cells by human adenoviruses and the origin of HEK 293 cells.* FASEB J, 2002. **16**(8): p. 869-71.
444. Lin, Y.C., et al., *Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations.* Nat Commun, 2014. **5**: p. 4767.
445. Krump, N.A., et al., *Merkel Cell Polyomavirus Infection Induces an Antiviral Innate Immune Response in Human Dermal Fibroblasts.* J Virol, 2021. **95**(13): p. e0221120.