Functional Characterization of the Non-Coding Control Region of Human Polyomaviruses

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Elvis Tasih Ajuh

aus Kamerun

Basel, 2017

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von

Prof. Dr. Marcel Tanner
Prof. Dr. med. Hans H. Hirsch
Prof. Dr. Volker Thiel

Basel, den 19.09.2017

Prof. Dr. Martin Spiess
Dekan

TABLE OF CONTENTS

1 ABREVIATION

$5HT_{2A}R$	5-hydroxy-tryptamine-2A serotonin receptor
BKPyV	BK polyomavirus
BKPyVAN	BKPyV-associated nephropathy
BKPyVHC	BKPyV-associated hemorrhagic cystitis
BP	Base pairs
BRE	B recognition element
CMV	Cytomegalovirus
CNS	Central nervous system
CPE	Core promoter element
CSF	Cerebrospinal fluid
DNA	Deoxyribonucleic acid
DPE	Downstream promoter element
Ds-DNA	Double-stranded deoxyribonucleic acid
ER	Endoplasmic reticulum
ERAD	Endoplasmic reticulum-associated degredation
EVGR	Early viral gene region
HIV-1	Human immunodeficiency virus-1
HPyV10	Human polyomavirus 10
HPyV12	Human polyomavirus 12
HPyV6	Human polyomavirus 6
HPyV7	Human polyomavirus 7
HPyV9	Human polyomavirus 9
HPyVs	Human polyomaviruses
Inr	Initiation element
JCPyV	JC polyomavirus
LTag	Large tumor antigen
LVGR	Late viral gene region
MCPyV	Merkel cell polyomavirus

mRNA	Messenger ribonucleic acid
miRNA	Micro RNA
NCCR	Non-coding control region
NK	Cells natural killer cells
NLS	Nuclear localization signal
PBMCs	Peripheral blood mononuclear cells
PML	Progressive multifocal leukoencephalopathy
pRB	Retinoblastoma protein
PyVAN	Polyomavirus associated nephropathy
PyVHC	Polyomavirus associated hemorrhagic cystitis
PyVs	polyomaviruses
RNA	Ribonucleic acid
RCA	Rolling circle amplification
RPA	Replication protein A
rr-NCCR	Rearranged non-coding control region
sTag	Small T antigen
STLPyV	Saint Louis polyomavirus
SV40	Simian virus 40
TSS	Transcription start site
TSPyV	*Trichodysplasia spinulosa* polyomavirus
VLP(s)	Virus-like particle(s)
Vp1	Major capsid protein 1
Vp2	Minor capsid protein 2
Vp3	Minor capsid protein 3
ww	Wild type or Archetype

2 ABSTRACT

With the advent of advanced molecular biology techniques, 13 human polyomaviruses (HPyVs) have been identified from different human body compartments. Generally, infection with HPyVs is harmless in immunocompetent people. However, some of these viruses are known to cause severe morbidities in the immunosuppressed. Five HPyVs, BK polyomavirus (BKPyV), JC polyomavirus (JCPyV), Merkel cell polyomavirus (MCPyV), Trichodysplasia spinolosa polyomavirus (TSPyV) and HPyV7 are known to cause diseases. Although other HPyVs have been detected in diseased tissues or cancers, evidence of their involvement is lacking. Furthermore, less is known about the cell tropism, replication in cell culture, and gene regulation of HPyVs. The non-coding control region (NCCR) of HPyVs functions as a bi-directional promoter/enhancer system, coordinating the respective steps of the viral replication cycle. Furthermore, it also determines viral persistence, host cell specificity, replication and virulence. Efficient replication of BKPyV and JCPyV have been demonstrated in cell cultures, but less is known with regards to the novel HPyVs. NCCR activity which can be used as an indicator of HPyV cell tropism and replication in cell culture is lacking for the newly discovered HPyVs. Despite similarities in the genome organization of HPyVs, the NCCRs of the HPyVs are different with respect to large T antigen binding sites, transcriptional factor binding sites (TFBS) and length. We hypothesized that the HPyV-NCCRs will display different activity in the same or different host cells. A bi-directional reporter vector recapitulating the HPyV genome with a red fluorescence protein (dsRed2) and a green fluorescence protein (EGFP) as markers of EVGR and LVGR, respectively, was designed. The reporter was used to analyze the NCCR activity of the 13 HPyVs in different cell lines originating from kidney, skin, lungs, cervix, brain and colon cancers. Our result demonstrated that the bi-directional HPyV-NCCR activity substantially differ in the same and different host cells. Indicating that the HPyV-NCCRs differentially sense and interpret the host cells' transcription factors and different host cells' transcription factors modulate the basal HPyV-NCCRs expression. As previously reported for BKPyV and JCPyV, rearranged (rr)-NCCR variants of newly discovered HPyVs associated with diseases showed higher EVGR expression compared to their respective archetype, indicating the NCCR a major pathogenicity

determinant. Analyzing the HPyV-NCCRs' activity in cell lines expressing T antigens (Tags) displayed activation of EVGR expression, suggesting that the bi-directional reporter recapitulates an essential NCCR response reported for the viral replication cycle. Furthermore, it indicates that cell lines expressing Tags of respective HPyVs can be used for the propagation of respective HPyVs. This data serves as a basis for understanding host and viral factors that may permit the identification of suitable cell culture systems and secondary host cell tropism for HPyVs. However, actual viral replication studies are needed to confirm replication of HPyV genomes in some of the cell lines tested. Furthermore, the role of the LTag-binding motif with regards to the regulation of EVGR and LVGR has been well-characterized in the monkey polyomavirus, simian virus 40, although very little is known with respect to the role of the LTag-binding motif for HPyVs. Given that the LTag recognition sites influence EVGR and LVGR expression and the LTag-binding sites organization of SV40 is similar to that of some HPyVs, such as BKPyV, we hypothesized that the role of the LTag-binding motif in the regulation of EVGR and LVGR expression in BKPyV is potentially similar to that of SV40. The EVGR and LVGR expression of BKPyV archetype (ww)-NCCR was compared with those of BKPyV ww-NCCR defective in one or more LTag-binding motifs in HEK293, HEK293T and HEK293TT cell lines, expressing none, small and large amounts of SV40 LTag, respectively. Our results indicated that abundant LTag decreased EVGR-expression by probably interacting with LTag binding *site A*, suggesting abundant LTag expression may down regulate EVGR expression, similar to SV40's-EVGR autoregulation. Contrarily, the LVGR expression is proportionally increased in the presence of increasing amounts of LTag expression (HEK293T and HEK293TT cells), by LTag potentially interacting with binding *sites A* and *D.* These results may open avenues for designing new therapeutic strategies targeting this reduction of EVGR expression in the presence of abundant LTag expression. Actual viral replication studies are needed to verify whether these mutants are replicative competent or not.

Concluding, a basic approach was used to elucidate factors that may allow the replication of HPyVs in cell cultures, that will greatly enhance our understanding of the basic biology of these viruses. Furthermore, putative host cell tropism of HPyVs was suggested. The effect of LTag-binding sites with respect to BKPyV's EVGR and LVGR expression, which could be extrapolated to other HPyVs was shown.

3 INTRODUCTION

3.1 Polyomaviridae

Polyomaviruses (PyVs) belonging to the family *Polyomaviridae* infect a wide range of vertebrates from birds to mammals. However, PyVs are limited in the species they can productively infect (1). The *Polyomaviridae* Study Group of the International Committee on Taxonomy of Viruses has recently revised the family *Polyomaviridae* by a sequence- and host-based rationale (1). Based on the observed distance between large T antigen (LTag) coding sequences, at least 73 species of the family *Polyomaviridae* were classified into four genera (Alpha-, Beta-, Gamma- and Delta-polyomavirus) (2). The Alphapolyomavirus and Betapolyomavirus genera contain polyomaviruses that infect a variety of mammals. The gammapolyomavirus genus contains avian polyomaviruses, while the genus Deltapolyomavirus contains only four human polyomaviruses (HPyVs). The observed distance between LTag-coding sequence of HPyVs is shown **(Fig. 1)**. However, the closest neighbor of a HPyV may be from a PyV from animals. At least 73 PyVs have been identified so far, of which 13 are HPyVs detected in different human-body compartments **(Fig. 2)** (2). Specifically BK polyomavirus (BKPyV) (3), JC polyomavirus (JCPyV) (4), Karolinska institute polyomavirus (KIPyV) (5), Washington university polyomavirus (WUPyV) (6), Merkel cell polyomavirus (MCPyV) (7), human polyomavirus 6 (HPyV6) (8), human polyomavirus 7 (HPyV7) (8), *Trichodysplasia spinulosa* polyomavirus (TSPyV) (9), human polyomavirus 9 (HPyV9) (10), human polyomavirus 10 (HPyV10) (11), St. Louis polyomavirus (STLPyV) (12) human polyomavirus 12 (HPyV12) (13), New Jersey polyomavirus (NJPyV) (14). While this work was in progress, the International Agency for Research on Cancer (IARC) in Lyon described a new PyV (LIPyV), which now needs confirmation as a HPyV (15). HPyVs generally cause asymptomatic infections in early childhood, which can then reactivate upon immunosuppression causing severe diseases. Of all these HPyVs, only BKPyV, JCPyV, MCPyV, TSPyV and HPyV7 are known to cause diseases (Fig. 2). The seroprevalence of HPyVs is up to 99% in the general adult population (Table 1) (16-18), suggesting a low-level infection of these viruses in the general population.

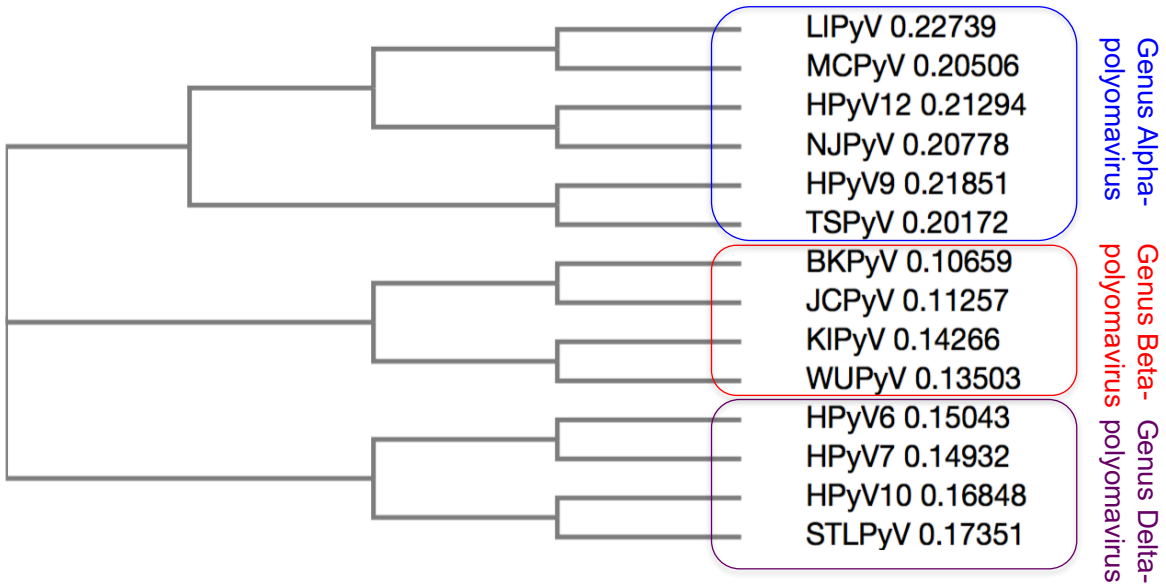


Fig. 1: A neighbor joining phylogenetic tree without distance corrections based on observed distance between HPyVs large T antigen coding sequences. The phylogram was constructed after multiple sequence alignment of the large T antigen sequences of the HPyVs using clustal omega. Next to each branch the evolutionary distance is depicted. Based on the tree, HPyVs could be grouped under 3 genera (Alpha-, Beta- and Delta-polyomavirus). GenBank accession numbers, from which the large T antigen sequences are based are found in table 6.

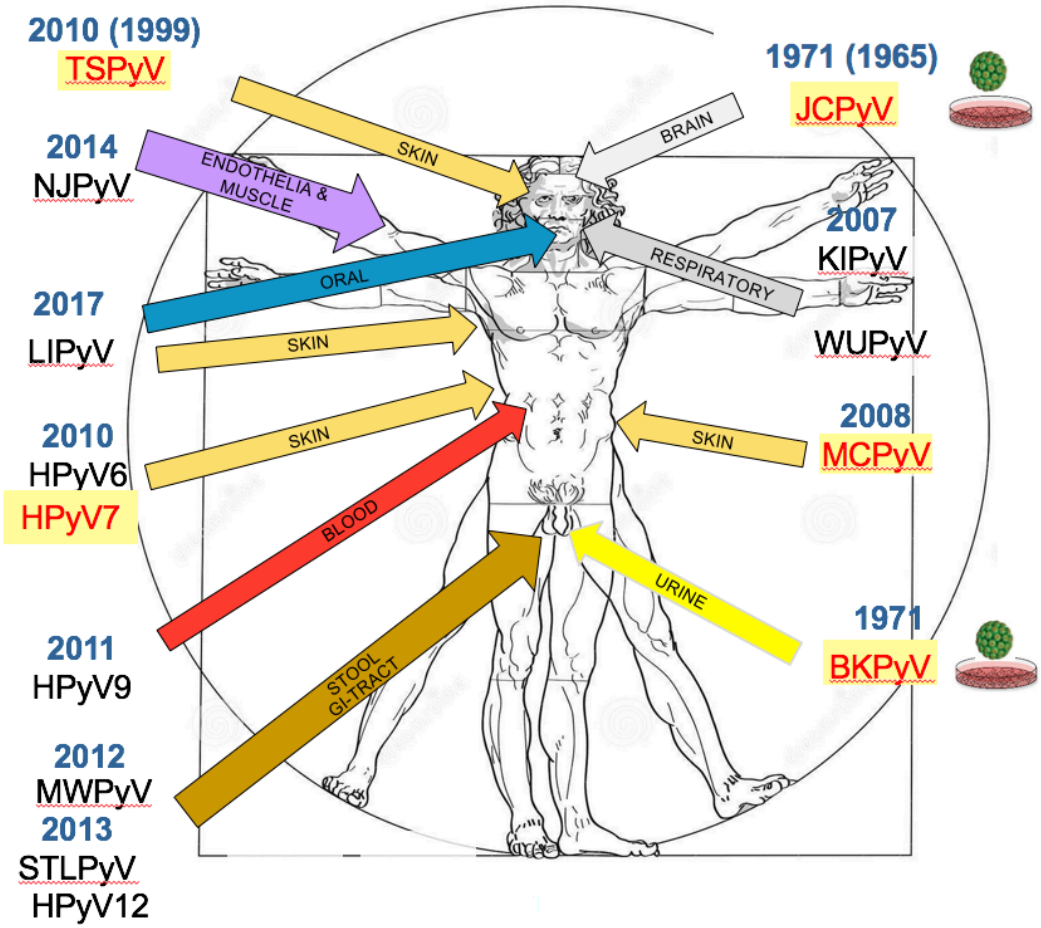


Fig. 2: The 13 HPyVs identified from different compartment of the human body including LIPyV which was discovered recently but needs confirmation as a HPyV (indicated with the arrows). The year of identification of each HPyV is depicted. Only 5 HPyVs highlighted in yellow can cause diseases. BKPyV and JCPyV are the two HPyVs that can be efficiently propagated in cell culture.

3.1.1 Structure and genome architecture of polyomavirus

Polyomaviruses are non-enveloped viruses with a supercoiled double-stranded circular DNA genome of about 5000 base pairs (bp). The genome is associated with cellular histones to form a minichromosome. It is encapsidated in an icosahedral-shaped viral capsid of approximately 45 nano meter (nm) in diameter (19, 20). The capsid is made up of 72 pentamers of the major capsid protein, viral protein 1 (Vp1), which is capable of self-assembly into viral-like particles (VLPs) (19, 21). Each Vp1 pentamer is associated with a minor capsid protein, Vp2 or Vp3 **(Fig. 3A)**. The genome of polyomaviruses is divided into 3 main regions, the non-coding control

region (NCCR) and two transcriptional regions; early viral gene region (EVGR) and the late viral gene region (LVGR) **(Fig. 3B)**. They are termed EVGR and LVGR due to the time of transcription during the course of the viral replication cycle. The EVGR and LVGR are transcribed in opposite strands and directions by the host cell RNA polymerase II as a single pre-messenger RNA.

3.1.1.1 Non-coding control region

The NCCR functions as a bi-directional promoter, harboring the origin of viral replication (ori), early and late promoters, enhancers regions, transcriptional start sites and a multitude of transcription factor-binding sites (TFBS). The NCCR can be referred to as the ‘’brain’’ of polyomaviruses. It is defined as the region between the start codon of the LVGR and the start codon of the EVGR. In BKPyV and JCPyV, the NCCR region proximal to the LVGR contains enhancer elements, LVGR transcriptional initiation sites and usually undergoes rearrangements including deletions, duplications and point mutations (22), whereas the region proximal to the EVGR harbors the origin of viral genome replication, EVGR transcriptional initiation sites and mostly not rearranged (23). This region also contains palindromes and dyad symmetry elements (23). Rearrangements of NCCR region proximal to the LVGR region result in different strains of the same species, which confers tissues specific expression of EVGR and LVGR expression. Rearranged BKPyV and JCPyV NCCR variants are common in BKPyV- and JCPyV-associated diseases (24-26). Nonetheless, BKPyV-induced hemorrhagic cystitis and nephropathy was not linked to any particular NCCR architecture (27). Furthermore, polyomaviruses containing similar NCCR rearrangement have been reported in different patient groups and healthy individuals (18, 28). The exact role of rearranged NCCR variants in pathogenesis and the mechanisms of generating the rearrangements are poorly understood. Rearranged BKPyV and JCPyV NCCR variants with diverse duplications and deletions from patients, have been reported to increase EVGR expression, replication capacity and cytopathology in vitro (25, 26). This suggests that BKPyV and JCPyV variants with rearranged NCCR can be linked to increase EVGR expression, replication capacity and disease (22, 25), which has not been studied for the novel HPyVs. A multitude of TFBS within the HPyV NCCR control the cell specific EVGR and LVGR expression of HPyVs. The roles of some of these TFBS in the

control of BKPyV and JCPyV gene expression are known (29-34), whereas little or no information is known with respect to the newly discovered HPyVs. Nonetheless, since some of the TFBS are conserved among HPyVs, the knowledge gained from BKPyV and JCPyV can be extrapolated to the other HPyVs.

3.1.1.2 Early viral gene region

The EVGR pre-mRNA is spliced into different variants encoding for large T antigen (LTag), small T antigen (sTag) **(Fig. 4)** and other T antigens variants, which are produced by alternative splicing. Some of the HPyVs encode additional early proteins. BKPyV encodes 17kT, JCPyV; T'135, T'136 and T'165, MCPyV; 57KT and ALTO, TSPyV; tinyT, 21kT, MTag and ALTO HPyV9; 145T, STLPyV may encode 229T, HPyV12; 84T and NJPyV may encode 299T (18).

3.1.1.3 Late viral gene region

The LVGR encodes the viral structural proteins, Vp1, Vp2, Vp3 and a non-structural late protein, agno, which are produced by alternative translational start sites. Only simian virus 40 (SV40), BKPyV and JCPyV encode the late non-structural protein, agno. The capsid of PyVs are composed of 72 pentamers of the major capsid protein Vp1, which are strengthened by disulphide and calcium bridges, and conserved throughout the polyomavirus family (20). These Vp1 pentamers form barrel-like structures, and within the hollows of Vp1, are contained minor capsid proteins, Vp2 and Vp3.

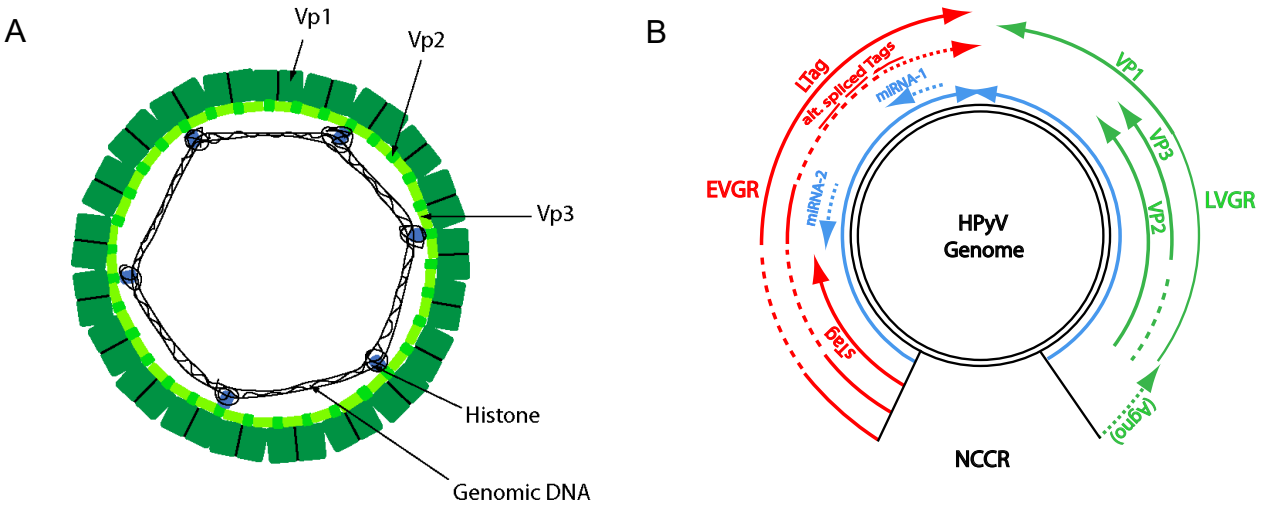


Fig. 3: (A) Schematic representation of the HPyV viral particle. (B) Representation of general features of the HPyV genome.

3.1.2 Polyomavirus life cycle

3.1.2.1 Host cell receptors, viral entry and un-coating

The detailed mechanisms of polyomavirus entry, trafficking through the cytoplasm and nuclear entry is not well understood. However, much of the understanding of the polyomavirus life cycle **(Fig. 5)** have been deduced from investigating SV40, BKPyV and JCPyV infection mechanisms (35). Vp1 is known to interact with the host cell receptors, as for instance mutation of the Vp1 calcium binding residues block viral entry into the host cell, resulting to disruption of virion assembly (36). The monkey polyomavirus, SV40 binds the monosialotetrahexosylganglioside (GM1) receptor and enters the cell via caveolar-mediated endocytosis. Additionally, SV40 can also independently enter the cell via caveolin or clathrin mediated endocytosis (37). On the other hand, BKPyV uses the disialoganglioside GT1b and trisialogangliosides GD1 as receptors to enter the cells by caveolar-mediated endocytosis, while JCPyV binds to serotonin 5-hydroxytryptamine 2A (5-HT2A) and enters the cell via clathrin-mediated endocytosis (38, 39). After internalization, the viral particle is trafficked through the microtubule network to the endoplasmic reticulum (ER) (40-42). ER transiting is unique to PyVs, no other DNA virus known to replicate in the nucleus uses this route. Interactions between the PyVs capsid proteins and the ER proteins have been well studied in SV40 (43). Partial un-coating of the viral particle takes place in the ER. The disulfide-isomerase unlinks the disulfide bonds between the Vp1 pentamers in the capsid (40, 44). The minor capsid proteins Vp2 and Vp3 become accessible with the help of chaperone proteins, which is important for subsequent steps in the viral life cycle. Once the viral particle is partially disassembled, it is then retrograde-transported from the ER to the cytoplasm via the ER associated degradation (ERAD) pathway. The viral particle is then guided into the nucleus via the nuclear pore with the help of the nuclear localization signal (NLS) present in Vp2 and Vp3 (45).

3.1.2.2 Early viral gene expression

The early and late promoters within the NCCR control the bidirectional EVGR and LVGR transcription. The EVGR and LVGR transcription progresses outward from the NCCR via opposite strands of the genome. Once the virus gets into the nucleus, EVGR transcription is initiated by cellular RNA polymerase II and controlled by cis-acting sequences within the NCCR, such as TATA box, GC-rich sequences and enhancers. A single pre-EVGR mRNA is produced which undergoes alternative splicing into sTag, LTag and other T antigens spliced variants. The T antigens' coding regions are overlapping at the N-terminal sequence, but differ at the C-terminal sequence (46).

3.1.2.2.1 Large T antigen: the multifunctional protein

The LTag of HPyVs is a multifunctional protein that interacts with numerous cellular and viral factors to initiate viral gene expression and viral genome replication (1). The structure and functions of the LTag have been largely described in the SV40 model, which is also relevant for BKPyV and JCPyV with amino acids homology of approximately 81% between SV40 and BKPyV, and 79% between SV40 and JCPyV (47). The large T antigen contained five regions that are conserved among HPyVs, that is, DnaJ region, LXCXE motif, nuclear localization signal, origin-binding domain (OBD) and ATPase/helicase domain. LTag binds to pentanucleotide sequences, GRGGC or its compliment GCCYC at the ori via the OBD region (48). The NCCR of all HPyVs contains at least one of these pentanucleotide sequences (49). Large T antigen binds to the ori within the NCCRs, unwinds the duplex DNA viral genome by using the ATPase/helicase activity (48). The ATPase gives the energy required for the helicase activity, while a zinc-binding domain proximal to the helicase domain is needed for the formation of LTag hexamer at the ori (48). In addition, LTag drives the cell into S-phase of the cell cycle resulting in the activation of DNA damage response (DDR) in order to promote replication of viral genome. The DnaJ domain cooperates with a proximally located motif called LXCXE to bind the retinoblastoma tumor suppressor proteins, pRB/p130/p107 (48). This binding disrupts the interactions between E2F (E2 transcription factor) and pRB/p130/p107. Dissociating pRB from E2F permits the formation of E2F-E2F dimers, which interacts with its target sequence and drives S-phase progression and DNA synthesis (48). Normally, improper entry of the host cell into S-phase would signal a p53-mediated apoptosis.

However, to circumvent the p53-mediated apoptosis, LTag binds p53 via the P53-binding domain within the helicase domain of LTag and allows progressive viral replication (1, 50). This p53 protein is a critical tumor suppressor in humans and its involvement in tumorigenesis has been widely demonstrated in the SV40 model (48, 51).

Likewise, sTag also has tumorigenic potentials. It promotes cells growth and viral replication by interacting with protein phosphatase 2A (PP2A), and disrupts its phosphatase ability (52). It has been reported to stimulate intercellular kinases to also promote cellular growth pathways (53).

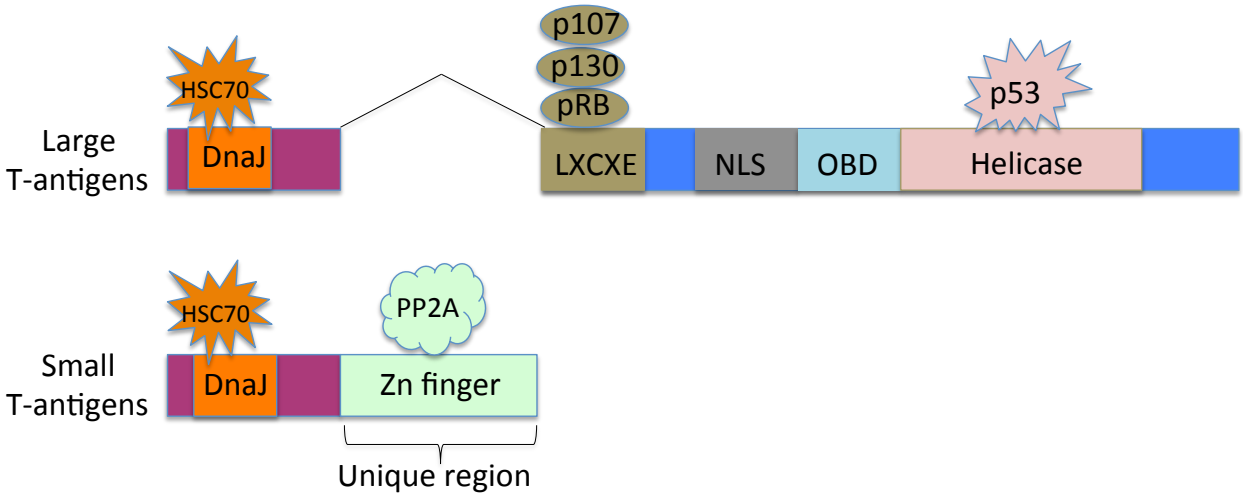


Fig. 4: Schematic representation of the large and small T antigens.

3.1.2.2.2 Viral encoded microRNA

During the viral life cycle, small noncoding RNAs (miRNAs) post-translationally regulate gene expression by directed-mRNAs cleavage or repression of translation. SV40, BKPyV, JCPyV, and MCPyV encode miRNA that target and deplete early viral transcripts, such as LTag mRNA (54-56). SV40 miRNA mutant infected cells were observed to be more sensitive to T antigen specific cytotoxic T-cell lysis in vitro, suggesting the importance of miRNA in viral evasion of the host immune response. However, similar amounts of infectious virions were produced with the SV40 miRNA mutant compared to the wild-type (55). Moreover, the stress-induced ligand, UL16 binding protein 3 (ULBP3), which is important in the activation of the natural killer T-cell receptor (NKG2D), has been reported to be targeted by BKPyV and JCPyV

miRNA (57). The NKG2D mediated killing of infected cells was reduced due to the downregulation of ULBP3 by the viral miRNA. This indicated that PyVs could use this mechanism to latently infect the host without being eliminated by the host immune system (57). Interestingly, archetype (ww) BKPyV encoded miRNA has been reported to be involved in limiting BKPyV ww replication in renal tubular epithelia cells (58). Nevertheless, regulation of the miRNA could be accomplished through the balance of regulatory elements within the NCCRs controlling EVGR and miRNA expression prior to genome replication (58). The BKPyV and JCPyV miRNA is encoded on the late strand with complementarity to the 3′ coding end of the T antigen mRNA (Fig. 2B) (54). Since the miRNA is located on the late strand, one could suggest that the miRNA is expressed from the late promoter, although it may also be expressed from within or outside the NCCR. The miRNA of BKPyV and JCPyV are identical; this indicates that in the case of co-infection, the miRNA from each virus can downregulates each other's T antigens expression. Since PyVs use miRNA to evade the host immune system, therapeutically targeting miRNA could be used for treating PyVs-associated diseases. Therefore, investigating the presence and function of miRNA of the other HPyVs is warranted.

3.1.2.3 Viral genome replication

PyVs depend on the host cell machineries to replicate its genome. LTag coordinates viral genome replication by recruiting multiple host factors required for viral genome replication. LTag monomers bind to the viral ori and assemble into two hexamers. Once the binding to ori occurs, LTag uses its helicase activity to unwind the double stranded DNA (48) (as explained above). Next, replication protein A (RPA) is recruited to bind the single stranded unwound DNA, to prevent it from annealing and forming secondary structures.

3.1.2.4 Late viral gene expression

Concomitantly to viral genome replication, late viral proteins are expressed. LTag begins suppression of EVGR expression at this point, but the exact mechanism is not well understood. It is suggested that, high LTag amounts binding the ori, particularly to LTag binding *site I* in SV40 (59), results in the downregulation (auto-regulation) of EVGR expression and promotion of LVGR expression. Additionally, miRNA could

also be involved as explained above. LVGR transcription produces at least 3 structural proteins, Vp1, Vp2 and Vp3. They are encoded in an overlapping manner; Vp3 is completely encoded in the 3' end of Vp2. SV40 has been reported to encode Vp4 (60). Furthermore, unlike the other HPyVs the LVGR of SV40, BKPyV and JCPyV encodes a small late non-structural protein called agno (6, 7, 61, 62).

3.1.2.5 Encapsidation and viral release

In baculovirus expression systems, Vp1 can self-assemble into the capsid protein without the presence of Vp2 and Vp3, but Vp1 cannot assemble into the capsid protein in the cytoplasm, which might be due to low calcium concentrations in the cytoplasm. This theory was supported by the higher calcium concentration in the nucleus, where virions were observed by confocal and electron microscopy (19, 21, 63). The LVGR capsid proteins made in the cytoplasm are transported to the nucleus with the help of the nuclear localization signal sequence within the proteins (NLS). Once in the nucleus they are assembled into virions. Vp1, the major capsid protein associates with minor capsid proteins Vp2 or Vp3 (not encoded by MCPyV), which extend internally anchoring the minichromosome. Viral release of the fully assembled infectious particles is poorly understood, but thought to occur via a number of processes including host cell lysis, agnoprotein increasing membrane permeability, Vp3 lytic properties, SV40 Vp4-mediated viroporin release (60, 64, 65). Vp4 was reported to be encoded by SV40 and function as a viroporin to form pores for viral release (60). However, recently, Henriksen et al. demonstrated that Vp4 is not important for viral release, since there wasn't any difference between wild-type and Vp4 mutant release (66).

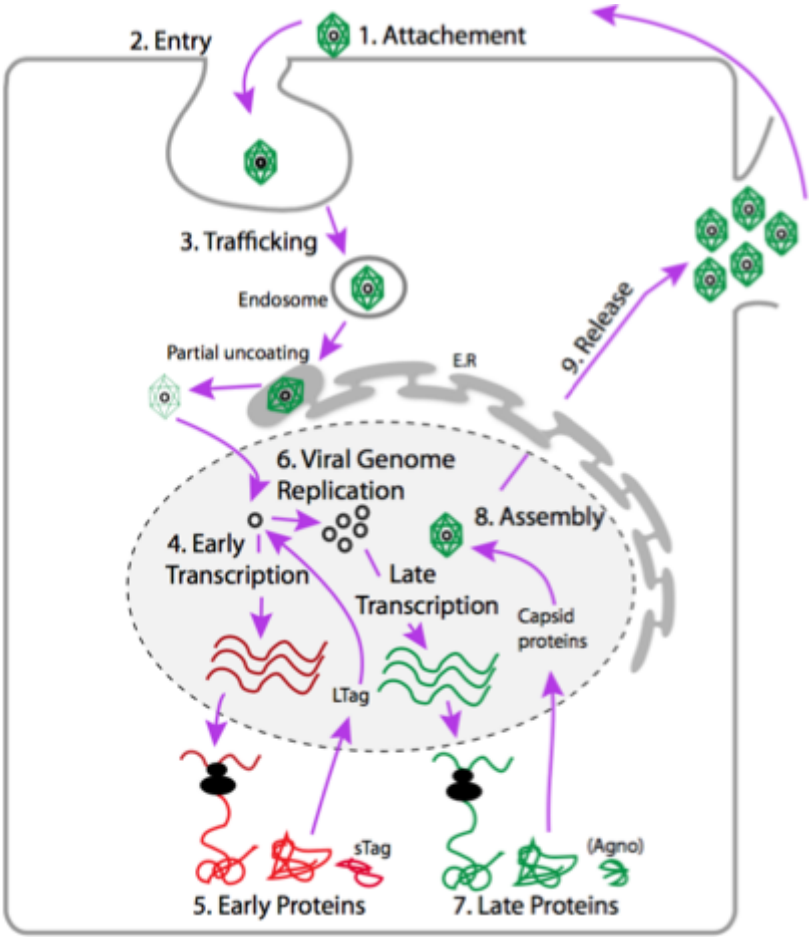


Fig. 5: Schematic representation of polyomavirus life cycle. **1.** Attachment of host-cell receptors. **2.** Entry by endocytosis. **3.** Transport of viral particle via endosomes to ER and later to the nucleus via the nuclear pore. **4 and 5.** EVGR transcription and translation. **6.** Nuclear translocation of LTag and viral genome replication. **7.** LVGR transcription and translation. **8.** Newly-synthesized viral genomes encapsidation. **9.** Infectious viral particles release.

3.1.3 **The emerging family of polyomaviruses**

3.1.3.1 Murine polyomavirus

The mouse polyomavirus (MPyV) was one of the first PyV to be isolated by Ludwig Gross in 1953 while studying murine leukaemia virus (67). Heat inactivation of an inoculum, which usually protects against infection with Murine leukaemia virus, still resulted in leukaemia and tumors in the parotids glands of newborn mice. Thus, he concluded that the tumor was due to another infectious agent. Gross named this

infectious agent after further experimentation “polyoma” due to its ability to cause multiple tumors in various tissues in newborn mice. Since the MPyV discovery, viral oncogenesis has been intensively studied (46).

3.1.3.2 Simian Virus 40

SV40 was discovered in 1960 as a contaminant of a batch of polio vaccines generated in primary kidney cells from Rhesus monkeys (68). Between 1955 and 1963 approximately 100 million adults and children received the polio vaccine and the same methods for generating the vaccine were used throughout the world. A vacoulating agent, later named Simian virus (SV40) was identified during safety testing in some cultures causing cytopathic effect (1, 68). Furthermore, investigations of culture systems infected with SV40 revealed that this virus had the ability to cause tumors in animal models (69). SV40 has been linked to non-Hodgkin lymphoma, adult and pediatric brain tumors, mesothelioma, and osteosarcoma (70). Epidemiological studies show that populations exposed to the SV40-contaminated polio vaccine failed to show increased risk of tumor development (71-73). However, the understanding of cell cycle and oncogenic pathways that permit the discovery of tumor suppressors pRB and p53 was greatly improved due to SV40 research (1, 46, 74). For more than 50 years the question of whether SV40 infection is an agent of human cancer has remained highly controversial (75).

3.1.3.3 Human Polyomaviruses

3.1.3.3.1 BK and JC polyomavirus

BKPyV and JCPyV were the two HPyVs discovered in 1971. They were named after the initials of the two patients from whom the viruses were isolated (3, 4, 76). BKPyV was isolated from the urine of a kidney transplant recipient (3), while JCPyV was isolated from the brain tissue of a progressive multifocal leukoencephalopathy (PML) patient (3, 4, 76). Inoculating human fetal brain tissue with isolated JCPyV agent revealed that JCPyV was the causative agent of PML, a fatal demyelinating disease in the central nervous system (CNS) (4). Infection of BKPyV and JCPyV occurs during early childhood. Seroprevalence of BKPyV and JCPyV in the general adult population ranges from 50-98% (77). The mechanism of transmission of BKPyV and

JCPyV is poorly understood, but thought to be via the respiratory or fecal-oral route (78-80). After primary infection of BKPyV or JCPyV, the viruses establish a persistent infection, possibly in kidney, neuronal, hematopoietic or lymphoid tissues (81-83). BKPyV has been detected in various organs/compartments of the human body; peripheral blood mononuclear cells (PBMCs), brain cells, urogenital cells, lymphocytes, cervix, rectum, sperm, skin and numerous cancers such as, neuroblastoma, osteosarcoma, prostate, cervix and Kaposi's sarcoma (84-87). JCPyV has been detected in PBMCs, kidney, lungs, bone marrow, gastrointestinal tract, tonsils, and lymphoid organs and in various cancers such as, B-cell lymphoma, lung, colon, urothelial, prostate, esophageal and cancers of the CNS (88-91). Reactivation of JCPyV and BKPyV can occur upon immune suppression. The incidence of PML has increased greatly with the pandemic of acquired immunodeficiency syndrome (AIDS) in the 1980s (90). BKPyV reactivation occurs mostly in renal transplant recipients resulting in nephropathy, which is associated with the risk of organ rejection (90, 92-94). Reactivation of BKPyV in bone marrow transplant (BMT) recipients could lead to hemorrhagic cystitis (95, 96). Both viruses are known to have four main serological subtypes, however, the clinical importance of these subtypes is poorly understood (97)

3.1.3.3.2 KI and WU polyomavirus

Approximately 36 years after the discovery of the first two HPyVs, the interest in human polyomaviruses was rekindled in 2007, when two new HPyVs, KIPyV and WUPyV, were discovered (6, 62). Similar to the initial detection of BKPyV and JCPyV, KIPyV and WUPyV were reported within one month of each other. Using large-scale molecular virus screening, KIPyV was detected in a respiratory secretion at the Karolinska institute (KI) by Allander et al. Within a month, Gaynor et al., identified WUPyV in a respiratory secretion from a child with clinical signs of pneumonia at the Washington University (WU) by using high throughput molecular sequencing techniques (6, 62). Furthermore, these viruses were also detected in respiratory secretions of children with acute respiratory illness (98). Since then, they have also been detected in blood, faces, urine, lung, tonsils, cerebrospinal fluid, spleen, lymphoid tissue, brain of HIV-infected patients with and without PML (99-102). Both viruses are phylogenetically related to each other and belong with BKPyV and

JCPyV to the Genus *Betapolyomavirus* (2, 6). The detection rates of both viruses in respiratory secretions by polymerase chain reaction (PCR) were between 0.4%-9% in studies conducted throughout the world (5, 6, 62, 98, 103-116). Detection rates in immunocompromised populations are estimated to be higher, and many of the reports lack appropriate control population and used small cohorts. The detection of KIPyV and WUPyV with other respiratory pathogens is common, thus making it difficult to draw the etiological link between these two viruses with respiratory illnesses (6, 62, 116, 117). The seroprevalence of KIPyV and WUPyV in healthy adults' ranges between 55%-93.3% and 54%-90%, respectively. This suggests a wide spread of KIPyV and WUPyV infection (77, 118, 119). Like BKPyV and JCPyV, it is hypothesized that KIPyV and WUPyV may persist in their host, reactivate and cause disease in immunosuppressed individuals. So far, no study has shown a significant difference between KIPyV and WUPyV detection rates in immunocompetent or immunocompromised individuals, although higher viral loads of KIPyV have been reported in the immunocompromised host (101, 120-127). Currently, no possible link of KIPyV and WUPyV to cancer has been made. Moreover, a role for both viruses in different cancer types is possible (128-130). Lastly, there have been no definitive association of these viruses with any disease.

3.1.3.3.3 Merkel cell polyomavirus

One year later (2008), Merkel cell polyomavirus (MCPyV) was detected using a digital transcriptome subtraction technique of RNA extracted from a Merkel cell carcinoma (MCC) (7). MCC is rare and aggressive skin cancer predominantly found in elderly and immunocompromised individuals. 50% of individuals with advanced cases live at least 9 months after being diagnosed with MCC, which is greatly reduced in the immunocompromised (7, 131). Compared to SV40, the genome of MCPyV has higher sequence homology to Lymphotrophic polyomavirus (LPyV), a polyomavirus isolated in 1979 from a B-lymphoblastoid African green monkey cell line. DNA of MCPyV has been detected in the peripheral blood of both healthy and immunosuppressed individuals and many other human samples/organs such as: eyebrow hairs, gall bladder, oral samples, lung, tonsils, intestine, appendix, liver, blood, saliva, lymphoid tissue, urine, malignant and non-malignant tonsillar tissues (132-137). Unlike other HPyVs, MCPyV is more closely related to MPyV.

Nonetheless, T antigens of MCPyV have the conserved features of the other HPyVs including; DnaJ binding domain, OBD, helicase/ATPase domain and conserved pRB binding motif (7). The seroprevalence of MCPyV in the healthy adult population ranges between 25%-42% (77). MCPyV was initially found integrated into the human genome (7). Viral integration into the host genome has also been reported for other polyomaviruses that cause tumor in cell culture models, but not before the discovery of MCPyV (7).

3.1.3.3.4 Human polyomavirus 6 and 7

Two new HPyVs were identified on the skin of adults in 2010 (8). HPyV6 and HPyV7 were detected by rolling circle amplification (RCA) from forehead swabs of volunteers (8). Small amounts of HPyV6 and HPyV7 virions are shed from skin of infected individuals. Some initial investigations detected 5/35 people to be positive for HPyV6 and 4/35 positive for HPyV7. These viruses have been detected in cervical specimens of HIV infected and non-infected, malignant and non-malignant tonsillar tissues, respiratory samples, feces and urine of an allogeneic stem cell transplant recipient. LTag of HPyV7 has been found in 23/37 thymomas (62.2%) and 8/20 thymic hyperplasia (40%) tissues. Furthermore, HPyV7 have been detected in PBMCs and skin of single and double lung transplant recipients with pruritic rash (133, 138-140). In addition, both viruses were detected in immunosuppressed patients with pruritic and dyskeratotic dermatoses (141). This suggests that these viruses can cause disease in immunocompromised patients. Seroprevalence studies using VLPs revealed that, HPyV6 and HPyV7 infect 69% and 35% adults, respectively, indicating that exposure to these viruses is common (8). So far definitive association of these viruses to tumorigenesis is unknown (142).

3.1.3.3.5 *Trichodysplasia spinulosa* polyomavirus

In 2010, another human polyomavirus called TSPyV was detected on the skin by RCA (9). TSPyV is the causative agent of *Trichodysplasia spinulosa* (TS) (9). TS is a rare skin disease affecting mostly immunocompromised persons (9, 143). The virus together with HPyV9, HPyV12, MCPyV, NJPyV and LIPyV belong to the genus *Alphapolyomavirus* (Fig. 1) (2). The clinical development of TS occurs as an eruption of erythematous papules with many follicular extensions localized in the limbs and

central face, mostly the nose (144), which was initially reported by Izakovic et al from a 31-year-old kidney transplant recipients undergoing treatment with cyclosporine A (145). Using quantified PCR and electron microscopy (EM) demonstrated that TSPyV was actively infecting TS follicular biopsy (144, 146). It has also been detected in 4% of skin swabs from kidney transplant recipients without clinical TS, indicating that this virus could be found on the skin without causing disease (9). In addition, this virus has been detected in renal allograft of a kidney transplant patient, heart, spleen, lung, colon, bronchus, small intestine, liver, respiratory samples, feces, tissue from a patient with myocarditis and in tonsillar tissues of healthy individuals (138, 147-149). TSPyV has high seroprevalence of about 70% in adult populations (150). This supports the theory that infection with this virus is usually asymptomatic causing TS only in highly immunocompromised individuals (150).

3.1.3.3.6 Human polyomavirus 9

In 2011, HPyV9 was detected in the serum of a kidney transplant recipient using degenerate primers (10). Whole genome sequencing of HPyV 9 revealed it to be closely related to LPyV and contained all the typical genomic characteristics of a polyomavirus (10). LPyV has been detected in blood in both immunosuppressed and healthy people and has a seroprevalence of 30% (77, 151, 152). A seroprevalence of 40% was determined for HPyV9, although cross reactivity of LPyV and HPyV9 was confirmed (153). This cross reactivity may account for the previous serological evidence of LPyV in human infection (153, 154). HPyV9 has been detected on skin swabs of MCC patients. This is in accordance with the finding that in MCC patient groups, HPyV9 was more common compared to age-matched controls (154). Furthermore, HPyV9 variant was detected in PBMCs of an AIDS patient (155).

3.1.3.3.7 Human polyomavirus 10

HPyV10 was discovered in 2012 in chondyloma specimens from a patient with warts, hypogammaglobinemia, infections, and myelokathexis-congenital disorder of leukocytes (WHIM) syndrome (11, 156). In the same, year a strain of HPyV10 termed Malawi polyomavirus (MWPyV) was detected in stool samples of healthy children from Malawi using shot gun 454 pyrosequencing (11). Another nearly genetically identical strain called Mexican polyomavirus (MXPyV) was detected in stool samples

of Mexican children (157). MWPyV strain was also found in one respiratory sample and stool samples from healthy Mexican children. Furthermore, MWPyV was found in tonsils and adenoids of healthy children, whereas HPyV10 was also detected on skin of HIV-infected and non-infected men suggesting skin tropism for HPyV10 (157, 158).

3.1.3.3.8 St Louis polyomavirus

Like MWPyV, STLPyV was detected in a stool sample from a healthy child in 2013 using 454 pyrosequencing (12). STLPyV is placed in the genus *Deltapolyomavirus* (2). The virus shows high sequence homology to HPyV6, HPyV7 and HPyV10 in the LTag coding sequence (2). STLPyV has been reported to encode a unique protein called 229T, obtained through alternative splicing of the EVGR pre-mRNA (18). Stool specimens collected from children confirmed the presence of STLPyV, however there was no statistical association between STLPyV and gastroenteritis. STLPyV has also been identified from a skin wart of a WHIM syndrome patient (12, 159). Compare to other HPyVs, STLPyV has a high seroprevalence of about 68%-77%, with age-stratified data indicating early childhood acquisition of this virus (160).

3.1.3.3.9 Human polyomavirus 12

HPyV12 was detected in resected human liver tissues and also in rectum, colon and a fecal sample (13). Studies in healthy adults/adolescents and pediatric children revealed a seroprevalence of 23% and 17%, respectively (13), indicating that primary infection of HPyV12 occurs during early childhood. So far, this virus has not been associated with any disease.

3.1.3.3.10 New Jersey polyomavirus

NPyV was detected in endothelial cells of a pancreatic transplant recipient suffering from necrotic plaques on multiple areas of the skin and retinal blindness (14). High viral load and viral particles of NJPyV within the nuclei of vascular endothelia cells were detected by specific real time PCR and EM (14). Reduction of immunosuppression resulted to slow subsiding of the patient's weakness, but the patient's eyesight failed to return. NJPyV viremia was detected until 10 months after the onset of symptoms, even though the patient showed improvement of symptoms. Furthermore, it was not possible to detect NJPyV in sera collected from the general

population or in the patient's pre-transplant serum (14).

3.1.3.3.11 Lyon IARC polyomavirus

The most recently LIPyV was discovered from the skin swabs, oral gargles and an eyebrow hair of cancer-free individuals by using a sensitive degenerative PCR and next-generation sequencing (15). The virus is phylogenetically related to the raccoon PyV identified in neuroglial tumors in free-ranging raccoons. The biological properties of this PyV is yet unknown.

Table 1: HPyVs discovery and potential and proven associated-diseases. N.A: not applicable.

Species/ common Name	Year discovered/ references	Initial biological compartment isolated/detected	another biological compartment detected	Associated-diseases	Seroprevalence / references
BKPyV/ HPyV1	1971 (3)	Urine of kidney transplant recipient	Bladder, blood, bone, brain, CNS, heart lung, skin, spleen, genital, salivary glands	BKPyV associated- hemorrhagic cystitis (BKPyVHC) (96), BKPyV associated-nephropathy BKPyVVAN) (161), progressive multifocal leukoencephalopathy (PML) (162), retinitis (162), pneumonitis (163) meningitis and encephalitis (164), ureteral stenosis (165), prostate cancer (166),	55–90% (77, 167)
JCPyV/ HPyV2	1971 (4)	CSF of a PML patient	Brain, CNS, kidney, lungs, GI tract, blood, tonsils	PML (4), colon cancer (168), multiple sclerosis (MS) (169)	44-90% (77, 167)
KIPyV/ HPyV3	2007 (62)	Respiratory tract secretions	Blood, stool, brain, tonsil, CNS, lung cancer tissue, lymphoid and tonsil tissues.	respiratory disease (18)	55-91% (77, 118, 170)
WUPyV/ HPyV4	2007 (6)	Respiratory tract secretions	Blood, stool, brain, tonsil, CNS, CSF, lymphoid and tonsil tissues.	respiratory disease (171)	69-98% (77, 118, 170)

MCPyV/ HPyV5	2008 (7)	Skin of Merkel cell carcinoma patients	Plucked eyebrow hairs, healthy skin, liver, urine, oral samples, tonsils, gall bladder, lung, saliva, oral samples	80% of Merkel cell carcinoma (7)	58-96% (77, 170, 172-174)
HPyV6	2010 (8)	Human skin swab	Eyebrow hair, stool sample, urine, nasopharyngeal swabs	BRAF inhibitor-induce epithelia proliferations, pruritic and dyskeratotic dermatoses (48, 141)	67–98% (8, 170, 172, 174, 175)
HPyV7	2010 (8)	Human skin swab	Eyebrow hair, stool sample, urine, nasopharyngeal swabs	Pruritic and dyskeratotic dermatoses (48, 141), thymomas (139)	35–86% (8, 170, 172, 174, 175)
TSPyV/ HPyV8	2010 (9)	Skin of Trichodysplasia spinulosa immunocompromise patient	Heart, lungs, liver, small intestine, colon, allograft of kidney transplant patient, bronchus, colon tissue,	Trichodysplasia spinulosa (48)	70–84% (150, 170, 172, 174)
HPyV9	2011(10)	Serum sample of kidney transplant recipient	Skin of Merkel cell carcinoma patient, Skin of healthy people	N.A	20–70% (153, 170, 172, 174-176)
HPyV10	2012 (11, 157)	Stool sample (MW)/ Condyloma specimen (10w)	Forehead swap, stool, respiratory samples	Genital wart, Hypogammaglobulinemia, infections, Myelokathexis (WHIM) syndrome (11, 156)	42–99% (170, 177, 178)
STLPyV/ HPyV11	2013 (12)	Stool sample	N.A	N.A	68–70% (160)

HPyV12	2013 (13)	Liver	Colon, rectum, stool	N.A	23–33% (13, 175)
NJPyV/ HPyV13	2014 (14)	Muscle biopsy of pancreatic transplant recipient	Muscle's endothelia cells and skin	Vasculitic myopathy in pancreatic transplant recipients (14)	N.A
LIPyV	2017 (15)	Skin, eye brown hair	N.A	NA	N.A

3.1.4 **Human polyomaviruses and associated diseases**

Most immunocompetent people with a primary HPyV infection have a nonspecific influenza-like or subclinical course. However, HPyV can replicate and cause specific syndromes in the immunocompromised leading to considerable morbidities. Some HPyV-associated diseases are regularly found in specific patient populations. This suggests that viral determinants interacting within the patient or organs influence the pathological outcome. The reasons for the associations are not clearly understood (179). Different pathologies have been associated with HPyV infections. For instance, BKPyV has been associated with BKPyV-associated nephropathy, however, its association with other diseases such as cancers and autoimmune disease is still been disputed. This presents a hurdle in linking some specific pathologies to HPyVs. However, a classification of HPyVs-associated pathologies has been described (16, 161)

The notions of HPyVs pathologies

Oncogenic HPyV pathology is characterized by HPyV infecting host cells that allow EVGR expression, but insufficient genome replication and LVGR expression to permit rapid viral release and host cell lysis. This concept has been clearly described in the MCPyV-associated Merkel cell carcinoma, whereby truncation of the LTag and/or Vp1 and MCPyV integration into the host's genome results in the impairment of MCPyV genome replication and LVGR expression (7). BKPyV LTag uncoupling of host cell activation, LVGR expression and BKPyV release may be associated with urothelial cancers (180).

Cytopathic HPyV pathology is due to rapid viral replication, but with little or no inflammation, resulting to cell lysis. This is seen in JCPyV-associated PML due to an uncontrolled JCPyV replication in the brain of HIV/AIDS patients. The cytopathology is evidence by the loss of the myelin producing cells, leading to neurological deficits and disease progression.

Cytopathic-inflammatory HPyV pathology is characterized by high-level HPyV replication and inflammatory response as a result of cell lysis and necrotic infiltrates of lymphocytes and granulocytes. This pathological pattern is seen in BKPyV-associated PyVAN stage B in kidney allografts.

Immune-reconstituting inflammatory syndrome (IRIS) HPyV pathology is due to high inflammatory response to HPyV antigens, usually during a hasty recovery of the cellular immune response (161, 181). The prototypes are JCPyV-associated PML, which worsen after the commencement of highly active antiretroviral therapy (HAART) (182) or following removal of natalizumab from the plasma in multiple sclerosis patients (183).

Autoimmune HPyV pathology characterized by responding to 'self' which may be provoked by HPyV antigens. For instance, PyV-LTag interaction with histones and DNA may trigger antibodies against histones or DNA (184-186).

	Transforming	Cytopathic	Cytopathic-Inflammatory	Immune reconstituting	Autoimmune
Concept	• Uncoupling of host cell activation and lytic PyV late-gene expression?	• Predominant PyV replication	• PyV replication *plus* immune response	• Predominant immune response	• Aberrant immune response (triggered by PyV DNA:histone or DNA:LTag)?
Example	• Urothelia carcinoma (p53:BKPyV LTag) • Prostate neoplasia	• PML by JCPyV in AIDS • PyVAN-A by BKPyV	• PyVAN-B by BKPyV (interstitial nephritis) • Ureteral stenosis	• PML by JCPyV in AIDS after HAART • HC by BKPyV in BMT	• Systemic lupus by BKPyV?

Fig. 6: Proposed pattern of HPyV disease during the course of viral replication and immune response adapted from (161). HAART; a colon (:) represents interaction; BMT, bone marrow transplantation; BKPyVAN-A and BKPyVAN-B, BKPyVAN patterns A and B.

BKPyV, JCPyV, MCPyV, TSPyV and HPyV7 are known to cause diseases in humans (3). Moreover, HPyV6 have been linked to certain diseases (48, 141), while other HPyVs, such as HPyV10 and NJPyV were initially detected from diseased specimens or patients (11, 156) .

3.1.4.1 Human polyomaviruses associated non-cancerous diseases

3.1.4.1.1 BKPyV-associated diseases

Polyomavirus associated nephropathy (PyVAN)

BKPyV was associated with PyVAN in renal transplant recipients. Approximately 1%-10% of kidney transplant recipients progress to PyVAN and up to 90% of these patients lose their allograft (187-189). PyVAN is a type of acute interstitial nephritis (190) characterized by denudation of the basement membrane **(Fig. 7A)** and necrosis of the proximal tubules due to BKPyV lytic infection and replication in kidney epithelial cells (191). The introduction of new and more potent immunosuppressive drugs suggests a relationship between immune surveillance disruption and BKPyV

reactivation (189, 192, 193). Currently, PyVAN diagnosis includes renal biopsy analyses by immunohistochemistry (IHC) **(Fig. 7B)** or histopathology to either determine cytopathic effects as a result of lysis of renal tubular epithelia cells. In addition, cytology of urine can be performed to detect decoy epithelial cells, which contain intranuclear viral inclusion bodies and PCR to detect and quantify viral load in urine or blood samples (194, 195). Urine cytology and PCR could be useful for early detection of reactivation and PyVAN (187), thus permitting faster diagnosis and intervention with a consequential increase in graft survival (196).

The risk factor for PyVAN is unclear, however combination of contributing factors from the virus, patient and graft determine the susceptibility to BKPyV. The major risk factor to PyVAN is suggested to be due to an overall degree of immunosuppression and not a particular immunosuppressive treatment (197, 198). More so, the development of PyVAN is not dependent on immunosuppression alone, since the manifestation of the disease is relatively uncommon in non-renal solid organ recipients and other immunosuppressed individuals (161). Furthermore, specific human leukocyte antigen (HLA) mismatches or donor seropositive, older age and male gender are all potential risk factors (187, 199-201).

Until now, there is no specific and effective BKPyV anti-viral treatment available. Nonetheless, decreasing immunosuppression to permit the immune system to regain control over the infection is the most common first line intervention, although may not be effective in all patients, and may expose the patient to a potential risk of acute graft rejection. Cidofovir, an antiviral nucleoside analogue, inhibits BKPyV DNA replication *in vitro* (202), but it is nephrotoxic at high doses (203). Studies revealed that at low doses and combined with the reduction of immunosuppression, Cidofovir can be effective in preventing BKPyV reactivation without causing significant nephrotoxicity (204-207). Additionally, the immunosuppressive drug leflunomide has shown some limited successes (207-209). Other antiviral treatments are fluoroquinolone antibiotics and intravenous immunoglobulin (IVIG) (197, 210-212). More detailed studies are required to evaluate the efficacies of these drugs. Lastly, rapid clearing of the viral source after nephrectomy and then re-transplantation could be another treatment approach.

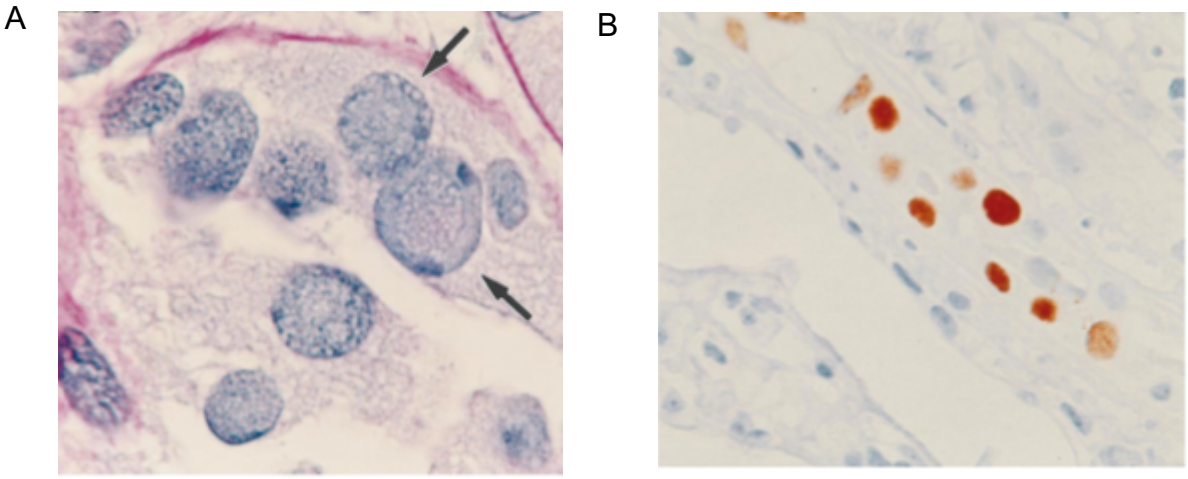


Fig. 7: BKPyV nephropathy (213). (A) Periodic acid Schiff stained section showing tubular epithelia cells with intranuclear viral inclusion bodies (arrows). (B) Immunohistochemistry (IHC) showing a formalin fixed and paraffin embedded tissue section showing BKPyV LTag staining (brown color) with SV40 LTag antibody, which cross reacts with BKPyV LTag.

Polyomavirus-associated hemorrhagic cystitis (PyVHC)

BKPyV is also associated with PyVHC affecting 10%-25% of bone marrow transplant (BMT) patients (214). The disease is probably due to BKPyV reactivation in possibly urothelial cells and renal tubular epithelia cells, resulting in high-level BKPyV replication due to lack of immune surveillance. Furthermore, there is prominent denudation of the cell's basement membrane, leading to hemorrhage and inflammation (215). The disease is characterized by flank pain, dysuria with or without signs of insufficiency. Additionally, it can be associated with varying grades (grades I-IV) of hematuria (215). A classical BKPyV-associated PyVHC occurs after more than 10 days post-transplantation. PyVHC is not life-threatening, but associated with significant morbidity (216-219). The most commonly applied diagnostic methods for BKPyV-associated PyVHC are PCR detection of viral DNA, IHC to detected LTag expression **(Fig. 8A)** and urine cytology for the detection of decoy epithelia cells **(Fig. 8B)** (220). Other sophisticated-diagnostic methods include; magnetic resonance imaging (MRI) or computer tomography (CT) to evaluate the bladder features. The bladder features in infected patients is characterized as; bladder wall thickening, increase mucosa enhancement, mural edema, small bladder capacity and

intraluminal clot (221). Like PyVAN, reduction of immunosuppression can be used as a treatment option for PyVHC, whereas exogenous BKPyV-specific T-cell application is being investigated as future therapies (214). Furthermore, antivirals including fluoroquinolone antibiotics, leflunomide, mTOR inhibitors and cidofovir are currently being evaluated for efficacy in treating BKPyV-associated diseases (222-225).

A B

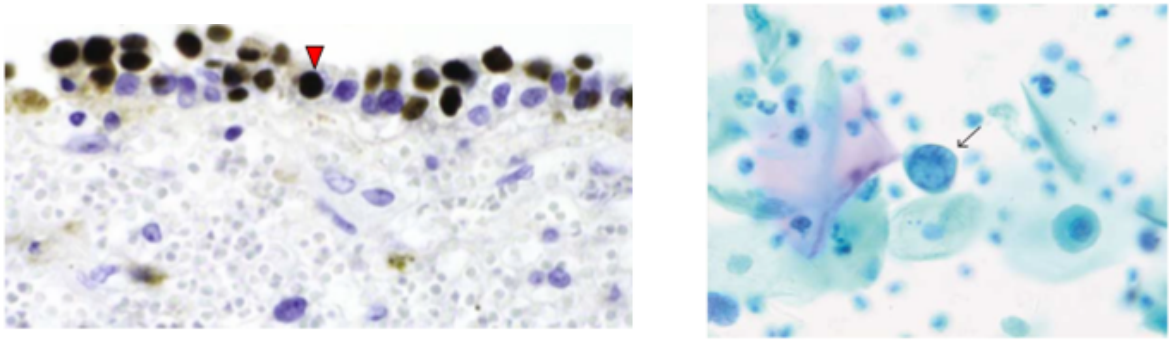


Fig. 8: (A) Bladder biopsy from a patient with PyVHC (www.pathology.mc.duke.edu) depicting denudation and BKPyV infected cells (red arrow head) with enlarged nucleus. (B) PyVHC (220), depicting the characteristic BKPyV-infected cells with decoy cells (arrow) showing enlarged nucleus containing a large intranuclear inclusion BKPyV-associated diseases.

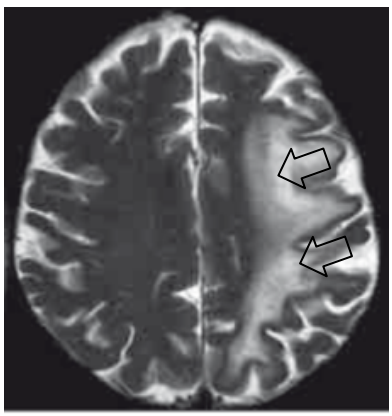
Progressive multifocal leukoencephalopathy (PML)

JCPyV is associated with PML affecting the immunocompromised due to HIV/AIDS, cancer, transplantation and treatment of multiple sclerosis patients with Natalizumab (4, 226-228). A high activity of JCPyV IgG activity is speculative of recent infection with JCPyV, and indicative for an increased risk for PML development in AIDS patients (229). Initially, PML was considered to be a very rare disease associated with immunocompromised patients, usually those with leukemia and lymphomas (93). With the rise of the HIV/AIDS epidemic, PML resulting from JCPyV infections have drastically risen. Up to 85% of PML cases occur in AIDS patients (230, 231), since not all AIDS-affected patients or even those with very low CD4+ T cells counts come down with PML (232). However, only few sporadic PML cases develop without an underlying immunosuppression (233). In essence, PML is now regarded an AIDS-defining disease (93).

PML is clinically characterized by dementia or confusion, speech and vision impairment, varying degree of akinesia or paralysis, in concurrence with immunosuppression and lack of increase in intracranial pressure (94, 233). The progression to PML disease is rapid, with clinically-intensifying symptoms and death within 3-6 months (234), which has however, been improved with the introduction of high-active anti-retroviral therapy (HAART) (235). With respect to the cellular level, PML results as lysis of oligodendrocytes, the myelin producing cells of the brain, leading to lesion in the brain stem, cerebellum and cerebrum and especially along the white and gray matter junction **(Fig. 9A)** (236, 237). JCPyV DNA sequences have been detected in mononuclear cells and B cells within the brain (238, 239). Enlarged JCPyV Tag-expressing oligodendrocytes with crystalline arrays of viral particles **(Fig. 9B**) and nuclear inclusion bodies are other features of PML (230, 231). The high frequency of JCPyV-positive PBMCs and widely distributed brain lesions indicate that lymphocytes can serve as a reservoir and a dissemination vehicle (231, 240, 241). Whether JCPyV infection of the brain is as a result of reactivation of latent infection or *de novo* invasion of the virus is unclear (81, 242).

The association between PML caused by JCPyV and HIV/AIDS is not clearly understood. However, the link between these two virus-associated disease could be explained by several hypotheses (230, 243, 244). Firstly, HIV establishes a thorough immunosuppression in the host by decreasing JCPyV-specific CD4+ T-cells leading to an uncontrollable JCPyV replication. Specifically, infection of HIV may cause break down of the blood-brain barrier, permitting entry of JCPyV-harboring B cells into the brain. Furthermore, certain cytokines produced as a result of HIV infection could initiate cell signaling events that may activate the JCPyV promoter. Lastly, the HIV Tat protein may activate JCPyV promoters in vitro and increase gene expression (245). Currently, there is no specific antiviral drug to treat PML. However, clinicians now focus on immune reconstitution-based concept.

A



B

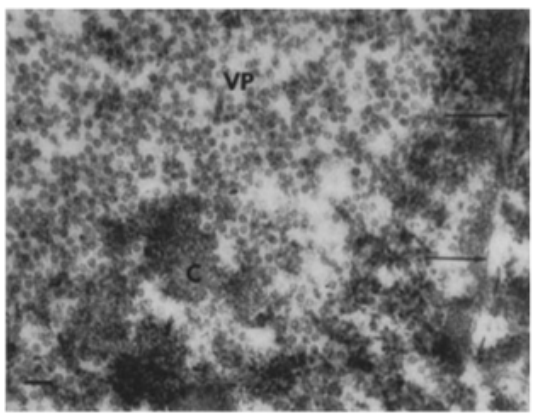


Fig. 8: (A) T2-weighted image of brain from patient with PML (246). The image shows multiple hyper-intense subcortical lesions in the left hemisphere (arrows). (B) Particle of JCPyV in PML lesion (226). Viral particles (VP) scattered all over major areas of the nucleoplasm. Arrows indicate nuclear membrane, while C shows nuclear chromatin.

JCPyV-associated PyVAN

The first case of JCPyV-associated PyVAN was identified in a 15-year old with kidney transplantation (247). Kidney biopsy from the patient stained positive for LTag **(Fig. 10)** and quantitative PCR revealed high JCPyV load in the urine (247). Generally, JCPyVAN is found in approximately 1% of kidney transplant patients (248) and about 6% of HIV patients present with JCPyV replication in the kidney (249). Most patients with JCPyVAN have normal kidney function indicating that JCPyV has a more protracted course or it is apparently less aggressive compared to BKPyVAN. In addition, BKPyVAN and JCPyVAN have another difference based on the severity of the histological pattern and strong viremia associated with BKPyVAN (187). Conversely, patients shedding large amounts of JCPyV in urine or patients with parenchyma involvement have been reported to show low viremia levels. This could be due to the basic differences between the biology of JCPyV and BKPyV, which has not been resolved in detail (250, 251). With the debilitating diseases caused by JCPyV and BKPyV, there is an urgent need for the development of specific antiviral therapies.

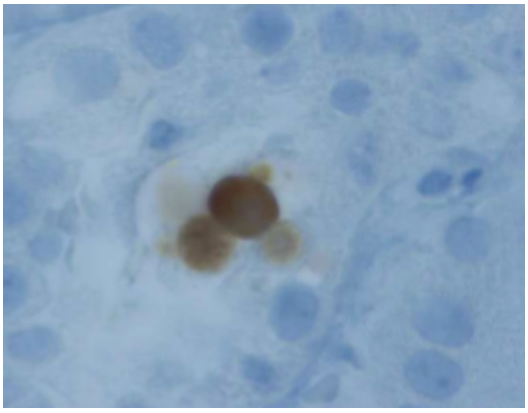


Fig. 10: Kidney allograft biopsy stained positive for LTag using SV40 LTag antibody which cross reacts with JCPyV LTag.

3.1.4.1.2 TSPyV-associated disease

Trychodysplasia spinulosa

TSPyV was initially detected in plucked facial spines from an immunosuppressed patient with a rare skin disease called *Trichodysplasia spinulosa* (TS) (9). Development of the disease may be due to uncontrollable replication of TSPyV in the inner root sheath cells. Analysis of the lesions by EM consecutively showed regularly-spaced 38-45nm intranuclear, crystalloid-organized viral particles (144). The disease is characterized by reddened flesh-colored papules often centrally located in the face and other places of the body, usually with spines made of keratin **(Fig. 11)**. TSPyV has been exclusively reported in the immunocompromised including patients with acute lymphocytic leukemia patients and solid organ transplant recipients (143, 252, 253).

TSPyV is diagnosed by histopathological examination of biopsies from lesioned skins showing enlarged hair follicles with perinuclear globules of the inner root sheath cells and evident eosinophilia. Additionally, TSPyV DNA and viral load is measured by quantitative PCR (qPCR). So far, all TS lesioned samples were positive for TSPyV with a high viral load ranging between 10^4-10^7 genome copies per cell (9, 144). Occasionally, samples from normal skin of TS patients, plucked eyebrow hairs and skin swabs from unaffected people were TSPyV DNA-positive with a viral load of below 100 copies per cell (9, 144). There is evidence that TS is caused by TSPyV (9, 144, 252). There is no antiviral treatment available for TS. However, restoring the immune function in immunocompromised patients resulted in complete resolution of TS symptoms, although with a higher risk of graft rejection. Significant reduction of TS symptoms has been revealed in a number of TS patients treated with 1%-3% cidofovir cream (144). In addition, oral valganciclovir, a guanosine analog, showed a modest effect in some TS patients (144). The exact mechanism of action of these drugs is yet to be elucidated.

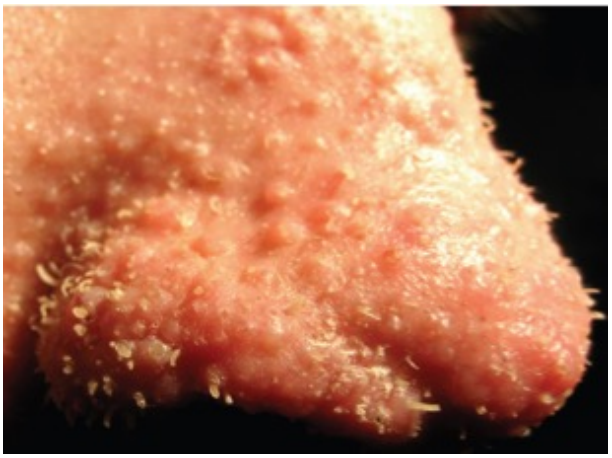


Fig. 11: Features of TS (9). The nose of a patient suffering from TS with numerous papules and spicules.

Potential diseases associated with novel HPyVs

DNA sequences of novel HPyVs, such as HyPV6, HPyV7, HPyV9, HPyV10, HPyV12 and NJPyV, were detected in patient's samples, but not much is known with respect to their potential disease-causing role in humans (7, 10, 13, 14, 156). On the other hand MCPyV is associated with a malignant disease, MCC (reviewed below) (7). Furthermore, KIPyV and WUPyV are detected mostly in respiratory secretions from patients with respiratory diseases, but so far the association of these viruses with respiratory diseases is unclear due to co-detection with other respiratory viruses (254). Nonetheless, high-throughput sequencing of isolated RNA from a nasal swab sample from a healthy and a one-month-old boy with severe respiratory symptoms revealed transcripts of KIPyV, whereas sequences of other viruses were not detected, except of few reads from human Boca virus (18). This indicates a potential role of KIPyV in respiratory pathogenesis. On the other hand, WUPyV may also cause respiratory pathogenesis due to a reported case of WUPyV associated with bronchitis (18). HPyV6 and HPyV7 are chronically shed from the skin (8). Both viruses have been reported to be associated with dyskeratotic dermatoses and pruritic rash in immunosuppressed patients (141, 255). The patients´ skin infection demonstrated brown plaques, scales and pruritus with a characteristic parakeratosis pattern (peacock plumage) (141, 255). More so, the skin biopsies presented with characteristic patterns of parakeratosis and dyskeratosis, which revealed abundance

of viral capsid, Tag, higher viral load compared to healthy skin and intact viral particles (141). HPyV6 DNA has been detected in epithelial neoplasms and the involvement of HPyV6 in the neoplasms cannot be excluded (256). Additionally, DNA of HPyV6 has been found in the CSF of an HIV/AIDS patient with PML (257) and in lymph nodes of a 51-old woman having angiolymphoid hyperplasia with kimura disease or eosinophilia (18). A hematopoietic stem cell transplant patient was diagnosed with HPyV7 viruria (18). Whether these viruses are associated with the above diseases is left to be determined. The small sample size was the major limitation to the above case reports. Probably, larger case control studies would shed more-light on disease-association of these viruses.

3.1.4.2 Human polyomaviruses and cancers

Since 1953, when MPyV was demonstrated to cause tumors in newborn mice, polyomaviruses have been suspected to be an etiologic agent of cancers. The role of SV40 as an etiological agent of cancer has been debated intensively. SV40 has been reported to induce tumors in animal models, transform human cells and detected in many human tumors. Moreover, epidemiological studies failed to show significant seroconversion in the human population (258, 259). Polyomaviruses associated-oncogenesis has been solely attributed to the LTag (28, 260, 261). LTag can interact with tumor suppressor proteins pRB, p107, p130 and p53, to interfere with cellular processes such as protein degradation, transcription, DNA repair, telomerase activity, stimulates angiogenesis, apoptosis and cell migration (260, 262). Interaction with p53 and retinoblastoma proteins required the C-terminal domain and the conserved LXCXE motif. MCPyV-derived MCC cells express LTag with a C-terminal truncation of the p53 binding domain, indicating cellular transformation by MCPyV is independent on p53. (263). Moreover, the sTag may also contribute to transformation by inactivating PP2A, and thus activating anti-apoptotic and proliferating pathways, and enhancing oncogenesis (18, 260, 264). The interaction of sTag with PP2A depends on the conserved motif, $CXCX_2C$ motif. This PP2A binding motif is conserved in all polyomaviruses (177, 265, 266), but whether this motif bind PP2A in all polyomaviruses is left to be tested. One could suggest that, polyomaviruses are potential oncoviruses, since they all encode the oncogenic proteins, LTag and sTag (Fig. 3). However, avian polyomaviruses which also encode LTag and sTag do not

seem to induce tumors, which could be due to the lack of the pRB consensus binding sequence (LXCXE) found in the LTag of mammalian polyomaviruses (28, 267). Cell culture studies reported agno detection in cancers and the probable contribution of agnoprotein to tumorigenesis (261, 268).

3.1.4.2.1 MCPyV-associated Merkel cell carcinoma

Out of the 14 HPyVs, MCPyV is in fact the only HPyV possibly associated with human cancer (7, 269). This virus is associated with about 80% of MCC (7, 269). MCC is a cutaneous tumor cause mostly by MCPyV. Cyril Toker was the first to described MCC in 1972 (270). He found a bluish-red or flesh-colored glassy painless nodule **(Fig. 12)** having solid trabeculae in 5 different places (lip, face, leg, forearm and buttock) of 2 elderly men, who died due to the tumors, and 3 elderly women (270). The name trabecular carcinoma of the skin was used by Toker to described these tumors, thinking that they depict an endocrine carcinoma of the sweat gland (270). MCC is classified as a group 2A carcinogenic biological agents (271). Epidemiological data revealed that this disease is affecting elderly people,

fair-skinned persons or the immunocompromised (270, 272, 273). More than 90% of MCC lesions are commonly found in sun-exposed areas, with half in the face, neck or extremities, it can also occur in perianal region, truck and genitalia although at a significantly reduced frequency (<10%) (270). The origin of primary MCC is unknown. However, the origin is thought to occur in Merkel cells, which is a type of mechanoreceptor cells found in the stratum basale in the epidermis. Excluding Merkel cells, the disease has also been suggested to originate from skin-derived precursors, and epidermal and dermal stem cells (274). Truncation of MCPyV LTag in MCPyV-positive MCC, promotes oncogenic transformation of MCPyV (275). The oncogenic potentials of MCPyV to other tumors have been extensively reviewed and documented in animal models and cell culture systems (48, 87, 91, 130, 134, 261, 263, 276-278). Moreover, the virus was not detected in squamous cell carcinoma or keratoacanthoma (279). Furthermore, MCPyV DNA was investigated in 10 small cell

lung carcinoma patients and 5 MCC patients, but was only detected in 2/5 MCC patients (280). In addition, investigation of MCPyV roles in 30 lung carcinomas, 28 prostrate samples and 25 childhood brain tumors revealed no evidence of MCPyV in these cancers (129, 281). More so, MCPyV DNA was found in 10 MCCs from a study with 1241 tumors (282). However, MCPyV was not found in ovary, breast, bone, melanoma, basal cell carcinoma, soft tissue, uterine cervix and large bowel cancers (282). These suggest that MCPyV has no roles in the oncogenesis of other cancers except MCC. Furthermore, the role of MCPyV in MCC is complimented with the fact that MCC is dependent on Tag expression and patients with MCC, usually demonstrate very high levels of MCPyV antibody titters compared to the general population, suggesting that infection of MCPyV is associated with MCC.

Presently, MCC is treated by using the following therapeutic strategies; chemotherapy, radiotherapy or surgery with little effect in the setting of metastasis (283). An immunotherapy based on programmed death 1 (PD-1), revealed good response in MCC patients with and without MCPyV DNA detection (284). Lastly, probable MCC prevention could be via a MCPyV VP1-based vaccine (285). The need for specific anti-viral treatment for MCPyV is warranted.

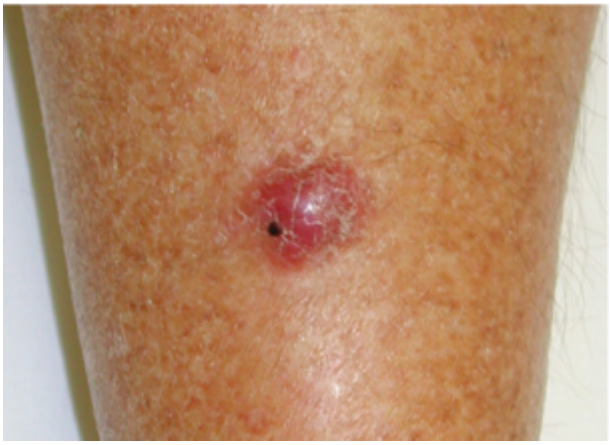


Fig. 12: Merkel cell carcinoma nodule on the leg of a patient (270).

The oncogenic potentials of JCPyV, BKPyV with tumors have been extensively reviewed and documented in animal models and cell culture systems (48, 87, 91, 130, 134, 166, 261, 263, 276-278). BKPyV and JCPyV are classified as a group 2B carcinogenic biological agents with possible carcinogenicity to humans (286). The

DNA, RNA, and proteins of BKPyV and JCPyV have been detected in many tumor tissues, but also in control non-cancer tissues (87, 91, 130). Thus, the potential role of BKPyV and JCPyV in human tumors is controversial. The presence of JCPyV in colorectal and CNS tumors may increase the risk of getting the tumors, while BKPyV may be associated with prostate and renal tumors (166, 278, 287). Harald zur Hausen, suggested that JCPyV-associated colorectal tumors may be due to other polyomaviruses consumed in meat (288). Several bovine polyomaviruses related to MCPyV, HPyV6, HPyV7 and other mammalian PyVs have been identified in beef samples (289, 290). Thus far, these viruses have not been detected in human colorectal biopsies. The association of novel HPyVs with cancers have been scarcely reported with only a few cases where DNA or protein of some novel HPyVs have been detected in tumor tissues **(Table 2)**. Based on the current knowledge, convincing evidence of the role of HPyVs excluding MCPyV in tumorigenesis is lacking (28, 171).

Table 2: Prevalence of novel HPyVs in human cancers, adapted from (28). Immunohistochemistry (IHC), fluorescence in situ hybridization (FISH).

Tumors	Number of samples	Method	Number of positive samples with HPyVs	References
Melanoma	18	PCR and IHC	HPyV6: 18, HPyV7: 17, TSPyV: 4, HPyV9: 1, HPyv10: 12	(256)
Squamous cell carcinoma (SCC)	63	PCR	HPyV6: 2, HPyV7: 1	(291)
Basal cell carcinoma	50	PCR	HPyV6: 2, HPyV7: 2	(291)
Melanoma	47	PCR	HPyV6: 2, HPyV7: 2	(291)
Basal cell carcinoma	41	PCR	HPyV6: 3	(292)
SCC	52	PCR	HPyV6: 2	(292)
SCC in situ	8	PCR	HPyV6: 1	(292)
Keratoacanthoma	42	PCR	HPyV6: 2	(292)
Actinic keratosis (AK)	31	PCR	HPyV6: 1	(292)
Breast cancer	54	PCR	HPyV6: 1, HPyV7: 1	(293)
Merkel cell carcinoma	1	Deep sequencing	HPyV6: 1, HPyV7: 1, HPyV9: 1	(294)
SCC AK	142	Deep sequencing	HPyV6: 1	(295)
Primary cutaneous B-cell lymphomas	130	PCR	HPyV6: 6, HPyV7: 1	(296)
Lung cancer	20	PCR	KIPyV: 9	(128)

Skin lesions from cutaneous T-cells lymphomas patients	39	PCR	HPyV6: 11, HPyV7: 5	(297)
Thymic epithelia tumours	37	PCR, FISH, IHC	HPyV7: 23	(139)
Thymic hyperplasias	20	PCR, FISH, IHC	HPyV7: 14	(139)

3.1.5 **Systems for HPyV propagation and infection**

Bacterially expressed Vp1 pentamers represent the smallest *in vitro* model used for studying viral entry. Pentameric subunits of Vp1 have been successfully used to recapitulate early events of TSPyV and JCPyV life cycle, such as structural studies of Vp1-receptor complexes, receptor specificity, as well as intracellular trafficking routes (298-300). Vp1 of PyV can be self-assembled into VLPs (301, 302). Baculovirus and yeast expression systems have been used to produce PyV VLPs (302). These VLPs were applied in PyV seroprevalence rates screenings (19, 301, 303, 304). Pseudo-viruses (PsVs) have been used for the analysis of PyV primary host cell tropism, such as viral attachment and entry. They are produced by co-transfection of mammalian cell lines such as HEK293TT (293TT) with expression vectors encoding PyV capsid proteins and vectors that encode a reporter protein, such as red fluorescent protein (RFP), green fluorescent protein (GFP) or Gaussia luciferase. This allows vivid quantification of the PsV in the target cells. Although these systems are important in revealing PyV's primary host cell tropism, the cellular factors required for viral expression and replication in the host cells cannot be analyzed using VLPs or Vp1 pentamers, since they lack the promoter of the virus which determines secondary host tropism.

So far, only BKPyV and JCPyV can be efficiently replicated in cell culture (305). Efficient propagation of archetype and rearranged variants of BKPyV and JCPyV have been conducted in the 293TT human embryonic kidney cell line (306). Viral replication, late protein expression and production of viral progeny were

demonstrated following transfection of episomal HPyVs DNA in 293TT cells. The 293TT cells carry an integrated plasmid encoding the cDNA of the SV40 LTag and a copy of the SV40 whole genome (307). The cell line constitutively expresses the SV40 LTag, which drives BKPyV and JCPyV DNA replication from their ori, but failed to significantly support replication from the ori of KIPyV, WUPyV and MCPyV. The successful replication of BKPyV and JCPyV in 293TT could be due to the fact that SV40 T antigens coding sequences have higher similarity to the BKPyV and JCPyV T antigens than the T antigens of KIPyV, WUPyV and MCPyV (306). Although large-scale propagation of BKPyV and JCPyV could be attained in these 293TT cells, this cell line is not a suitable model for the study of HPyVs` life cycle, since it has been transformed. Presently, to study BKPyV *in vitro* infection, human renal proximal tubule epithelial cells (RPTEC) are used as the state-of-the-art model. Nevertheless, the proliferative potential and the experimental design are restricted, since RPTECs could enter into replicative senescence. Rearranged BKPyV variants have been reported to replicate in a broad range of different cell lines including brain and melanoma cancer cell lines, monkey kidney cell lines (Vero and CV-1 cell lines) and human ovarian (302, 308, 309) and salivary gland cells (85, 86, 308). JCPyV has a major tropism in vivo for glial cells, while productively infecting oligodendrocytes and astrocytes to a lesser extent (85, 310). JCPyV infection has been mostly studied in the human fetal glial cell line, SVG which constitutively expresses SV40 LTag (311). A subpopulation of SVG p12 cells provided by ATCC has been reported by Henriksen et al., (312) to be productively infected with BKPyV (312). A subclone of the original SVG human glial cell line, SVG-A cells, which do not contain BKPyV, can be used as an alternative (312). Furthermore, efficient replication of JCPyV archetype and rearranged variants has been reported in Cos-7 which is a CV-1 cell line transformed with SV40 origin-defective mutant (311, 313).

MCPyV have been reported to be propagated in 293-4T cells, a 293TT cell line co-transfected with plasmids expressing MCPyV sTag and LTag (314). In addition, replication of MCPyV in HEK293 cells stably expressing MCPyV sTag or LTag only has also been reported (315). These results suggest that expression of specific T antigens is a limiting factor for HPyV propagation, thus specific T antigens expressing cells could be used as PyV-permissive cell culture systems for other unculturable HPyVs. Moreover, MCPyV have been replicated in cell lines not expressing MCPyV

T antigens such as, H1299 (human non-small cell lung carcinoma cell line) and PFSK-1 (neuroectodermal tumor cell line) and in HEK293 cells, but without serial infection (315, 316). This indicates that these cell lines contain the required transcription factors necessary for MCPyV replication, but lack the necessary receptor(s) and co-receptor(s) for attachment and entry.

The host cells naturally infected by HPyVs are so far unknown with the exception of MCPyV whose natural host cell has been demonstrated to be human dermal fibroblasts (317). Animal models to study HPyV infection are hampered by the narrow host range and cell specificity of HPyVs. For instance, JCPyV inoculation of hamsters or mice resulted in tumor formation, but failed to recapitulate the JCPyV demyelinating disease present in humans (318). To date, small animal models available for HPyV infection include humanized mice, xenografts, infectious mouse PyV models and transgenic mouse models (305). Available culture systems have focused mainly on the propagation of BKPyV and JCPyV with the focus on understanding the primary host cell tropism, that is receptors and co-receptors for JCPyV and BKPyV attachment and entry (319-321). Furthermore, most studies used PsVs and Vp1 pentamers, which lack important components of the viral genome. The study of viral attachment and entry receptors is important, but cellular and viral factors that interact with the HPyV-NCCRs will determine the secondary host cell tropism and potential host cells for HPyVs propagation.

The NCCR of HPyVs defines the secondary host cell tropism and is critical for all HPyVs biology (persistence, latency, reactivation). The activity of HPyV-NCCRs, especially those of novel HPyVs in cell culture is scare, whereas that of BKPyV and JCPyV have been frequently analyzed. Examining the NCCR activity of novel HPyVs in cell culture will shed lights on HPyVs cell tropism and cell lines where they could be productively replicated. The genome of HPyVs are similar however, the HPyV-NCCRs are different with respect to length, LTAg-binding sites and TFBS. NCCR rearrangement has been studied in detailed for the BKPyV and JCPyV, and less for the novel HPyVs. Therefore, understanding the effect of HPyV-NCCR rearrangement in novel HPyVs, and the detailed mechanistic role of Sp1 binding sites and LTag-binding sites in the regulation of EVGR and LVGR expression, would go a long way in deciphering possible avenues for therapeutic strategies against the growing family

of HPyVs. The role of Sp1 and LTag-binding motif on EVGR and LVGR expression were investigated in BKPyV ww as a model, which can be extrapolated to other HPyVs with similar Sp1 and LTag-binding sites organization.

4 AIMS OF THE THESIS

The aims of the thesis are:

1. Compare the NCCR activity of the 13 known HPyVs in different host cell lines with respect to EVGR and LVGR expression.
2. Characterize the effect of exogenous LTag expression in HPyVs EVGR and LVGR expression.
3. Characterize the effect of different host cell lines on EVGR and LVGR expression.
4. Examine the role of predicted LTag-binding motif in the BKPyV ww NCCR with respect to EVGR and LVGR expression.

4.1 Hypotheses

Question 1

Are there differences in HPyV-NCCRs activity with respect to EVGR and LVGR expression in the same host cells?

H1: There are differences when present in the same host cell.

H0: The are no differences.

Approach

We used a bi-directional reporter vector to analyze the NCCR-driven EVGR and LVGR expression of the 13 HPyVs in HEK293 as a reference cell line.

Interpretations

- The HPyV-NCCRs which differ in length and TFBS will sense and interpret the same host cell's transcription factors make-up to drive gene expression.
- The HPyV-NCCRs may have the same activity in the same host cells.

Caveat: The techniques used may not be sensitive enough to measure differences.

Question 2

Are there differences in HPyV-NCCRs activity with respect to EVGR and LVGR expression in different host cells?

H1: There are differences when present in different host cells.

H0: The are no differences.

Approach

We used the bidirectional reporter vector to analyze the NCCR-driven EVGR and LVGR expression of the 13 HPyVs in different cell lines.

Interpretations

- The HPyV-NCCRs which differ in length and TFBS will sense and interpret different host cells' transcription factors make-up to drive gene expression.
- The HPyV-NCCRs may have the same activity in the different host cells.

Caveat: The techniques used may not be sensitive enough to measure differences.

Question 3

Does SV40 T-antigens have effects on HPyV-NCCRs driven EVGR and LVGR expression?

H1: SV40 T-antigens have effects on HPyV-NCCRs driven EVGR and LVGR expression.

H0: SV40 T-antigens have no effect on HPyV-NCCRs driven EVGR and LVGR expression

Approach

We examined the effect of T-antigens on the EVGR and LVGR expression of HPyVs in HEK293-derivative cell lines that constitutively expressing SV40 sTag and LTag.

Interpretations

- The HPyV-NCCRs which differ in number and organization of LTag-binding sites will sense and response to SV40 T-antigens.
- The HPyV-NCCRs may not response to SV40 T-antigens.

Caveat: Our techniques used may not be sensitive enough to measure HPyV-NCCRs' response to SV40 T-antigens.

Question 4

Does SV40 T-antigens have different effects on HPyV-NCCRs driven EVGR and LVGR expression?

H1: SV40 T-antigens have different effects on HPyV-NCCRs driven EVGR and LVGR expression.

H0: SV40 T-antigens have the same effect on HPyV-NCCRs driven EVGR and LVGR expression.

Approach

We examined the effect of T-antigens on the EVGR and LVGR expression of HPyVs in HEK293-derivative cell lines that constitutively expressing SV40 sTag and LTag.

Interpretations

- The HPyV-NCCRs which differ in number and organization of LTag-binding sites will sense and response to SV40 T-antigens differently.
- The HPyV-NCCRs may response the same to SV40 T-antigens.

Caveat: Our techniques used may not be sensitive enough to measure the difference in HPyV-NCCRs' response to SV40 T-antigens.

Question 5

Does mutageneses of LTag-binding sites of BKPyV-NCCR have an effect on EVGR and LVGR expression in HEK293 cells and HEK293-derivative cell lines that constitutively expressing SV40 LTag in different amounts?

H1: There is an effect on EVGR and LVGR expression.

H0: There is no effect on EVGR and LVGR expression.

Approach

We used the bi-directional reporter to compare the BKPyV ww NCCR-driven EVGR and LVGR expression to designed BKPyV ww NCCR defective in one or more LTag-binding sites.

Interpretations

- The number and position of LTag-binding sites will affect the NCCR-driven EVGR and LVGR expression.
- The number and position of LTag-binding sites has no role with respect to EVGR and LVGR expression

Caveat: Our techniques might not be sensitive to measure the EVGR and LVGR expression of the different LTag-binding site mutants.

4.2 Rationales

Increasingly, the HPyVs are considered to be emerging opportunistic viruses. Very little is known about the basic biology of these viruses, especially the newly discovered HPyVs. Information on the route of transmission, associated diseases, medical treatment, life cycle, viral gene regulation, cell tropism and replication in cell culture are lacking. Furthermore, the NCCR activity which can be used as an indicator of HPyV cell tropism and replication in cell culture is lacking for the newly discovered HPyVs. Propagation of these viruses in cell culture is scarce; only BKPyV and JCPyV can be efficiently replicated in cell culture. Nonetheless, many cell culture studies have focused on the mechanisms of HPyV primary host cell entry by using Vp1 pentamers, VLPs and pseudoviruses, which do not include all aspects of the viral replication cycle. Although primary host cell tropism is important, cellular factors or early viral proteins that interact with the NCCR will affect viral expression and replication, thus determining the secondary host cell tropism. The NCCR of HPyVs functions as a bi-directional promoter, regulating the EVGR and LVGR expression and determining the permissiveness of HPyVs in their host cells. Analyzing NCCR activity of the 13 HPyVs, the roles of LTag and Sp1 binding motif with respect to EVGR and LVGR expression, would throw lights on the basic biological aspects of HPyVs, help in the development of cell culture systems for propagating HPyVs, identification of cell tropism and open avenues for new therapeutic approaches for eliminating these viruses, especially in the context of immunosuppression.

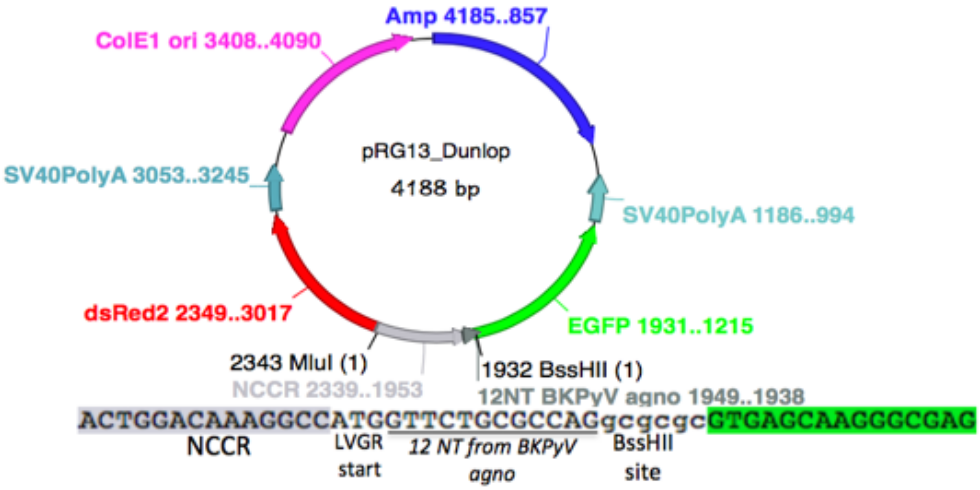
5 RESULTS

5.1 Designing a smaller bi-directional reporter vector lacking some sequences compared to the previous pHRG reporter vector

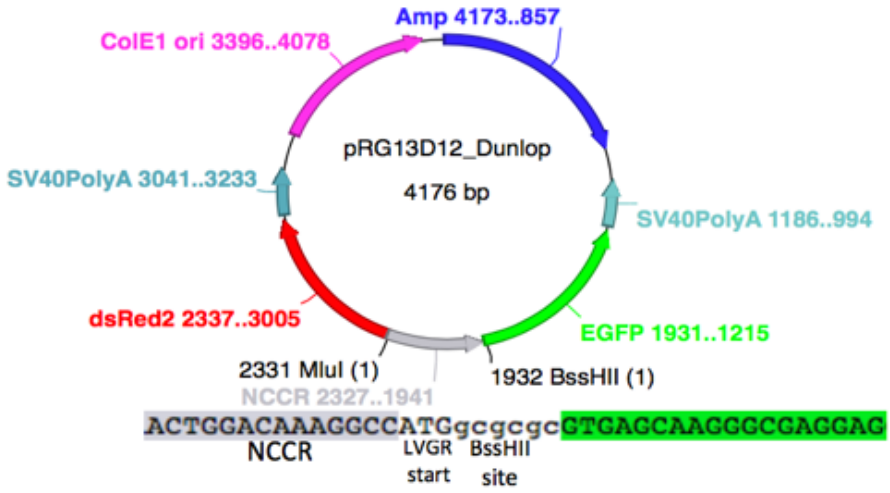
Transfection of large plasmids has been reported to have lower efficiency than small ones in human cell lines**.** We designed a new-smaller bi-directional reporter pRG13 recapitulating the HPyV genome **(Fig. 1A)** based on the phRG reporter described in Gosert et al. (25)**.** Compared to phRG (25), the reporter lacks the beta globulin polyadenylation signal, Hygromycin gene and an SV40 early promoter, which could compete with the NCCR. The late leader sequence encoding the agno protein of JCPyV has been reported to contain potential regulatory elements. Deletion or mutagenesis analyses of this sequence significantly decreased EVGR expression and thus viral replication (322). Furthermore, clinical BKPyV variants with agno gene deletion have been identified to be non-functional (323). PhRG and the newly designed pRG13 contain a 12-nucleotide sequence of BKPyV ww agno gene (nucleotides number 4-15) immediately after the NCCR and the start codon of EGFP **(Fig. 1A).** We deleted the 12-nucleotide sequence in pRG13 (BKPyV-Dunlop) **(Fig. 1B)** and phRG (BKPyV-Dunlop) and transfected HEK293 cells. Two days post-transfection, we quantified the EVGR and LVGR expression by flow cytometry and calculated the mean fluorescence intensity (MFI) of EVGR (red) and LVGR (green) expression **(Material and Methods in section 5.2)**. Surprisingly, both reporter vectors showed reduction in EVGR expression in favor of LVGR expression of more than 5-fold compared to the respective undeleted ones **(Fig. 1C, D).** We suspected the 12-nucleotide sequence harbored vital regulatory TFBS. To visualize this, we used the MatInspector tool (324) to search the 12-nucleotide sequence for the presence of TFBS. We identified putative TFBS including one for nuclear factor 1 (NF1), which we called NF1-7 representing the 7^{th} NF1 binding site with respect to the six NF1 bindings sites in the BKPyV ww NCCR (**Fig. 2A**). Since many studies have reported NF1 to be an important regulator of viral expression (325, 326), we made phRG (BKPyV-DUN) constructs with mutations within the 12-nucleotides sequence that would affect the core binding site of NF1 (NF1-7_mut1, NF1-7_mut2

and NF1-7_mut4) and another that would not affect the core NF1 binding site (NF1-7_mut3) (**Fig. 2A**). The core NF1 binding site is shown on the right of Fig. 3A. The result revealed, that mutagenesis of the predicted NF1-7 site decreased EVGR expression to ~30%, while the LVGR expression increased ~2.6-fold, but not up to the levels of the deletion mutant (**Fig. 2B, C**). This indicates that other factors within the deletion may be involved in the regulation of EVGR and LVGR expression. On the other hand, mutation, which did not affect NF1 binding, had little or no effect on EVGR or LVGR expression. This result suggested that the immediate BKPyV agno coding sequence (nucleotides 4-15) contained an NF1 binding site, which may cooperate with other factors to regulate EVGR and LVGR expression of BKPyV. Since the 12-nucleotide sequence could have a regulatory function and novel HPyVs do not encode agno, we analyzed the HPyVs-NCCR activity in a bi-directional reporter pRG13D12 (D12 indicates pRG13 lacking the 12-nucleotide sequence) in section 5.2.

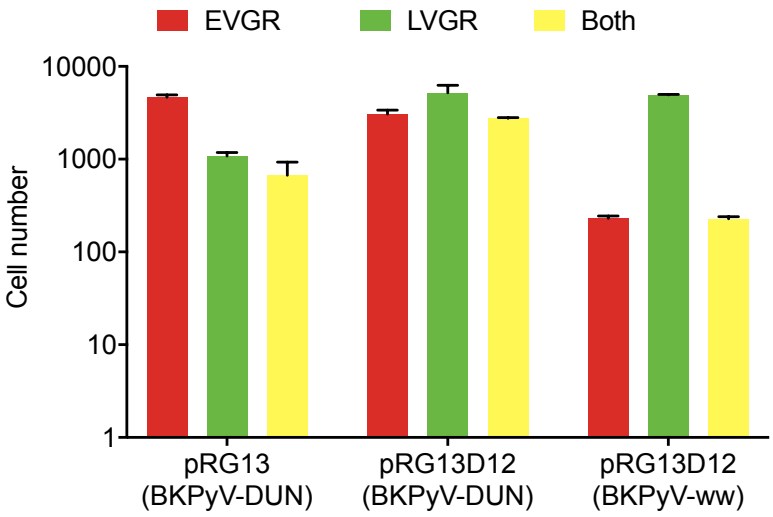
A


Amp 4185..857
ColE1 ori 3408..4090
pRG13_Dunlop
4188 bp
SV40PolyA 3053..3245
SV40PolyA 1186..994
dsRed2 2349..3017
EGFP 1931..1215
2343 MluI (1)
NCCR 2339..1953
1932 BssHII (1)
12NT BKPyV agno 1949..1938
ACTGGACAAAGGCCATGGTTCTGCGCCAGgcgcgcGTGAGCAAGGGCGAG
NCCR
LVGR start
12 NT from BKPyV agno
BssHII site

B


Amp 4173..857
ColE1 ori 3396..4078
pRG13D12_Dunlop
4176 bp
SV40PolyA 3041..3233
SV40PolyA 1186..994
dsRed2 2337..3005
EGFP 1931..1215
2331 MluI (1)
NCCR 2327..1941
1932 BssHII (1)
ACTGGACAAAGGCCATGgcgcgcGTGAGCAAGGGCGAGGAG
NCCR
LVGR start
BssHII site

C



D

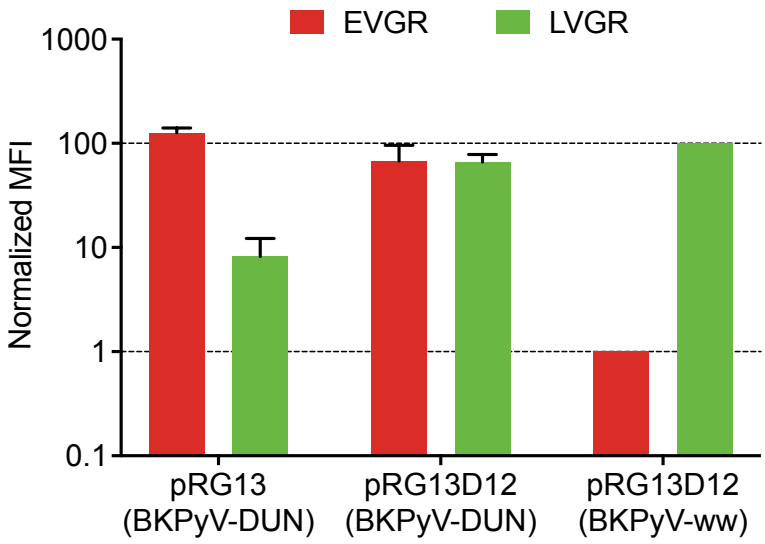
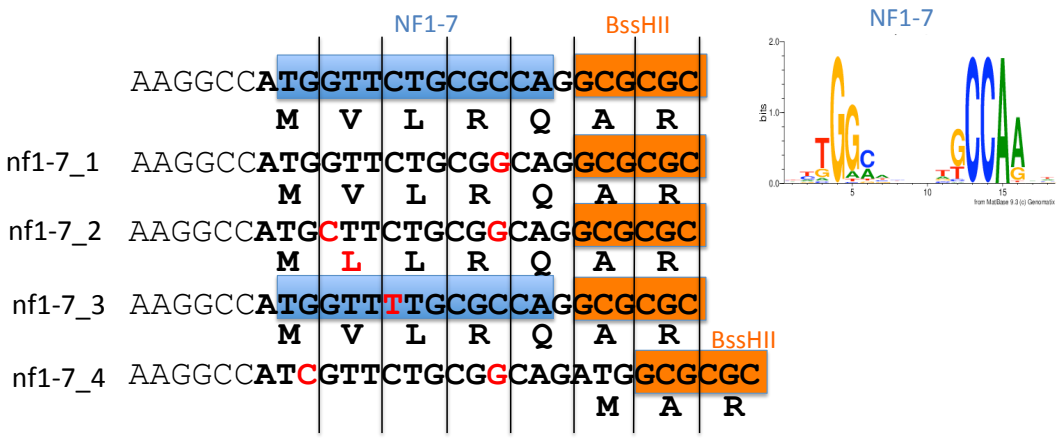


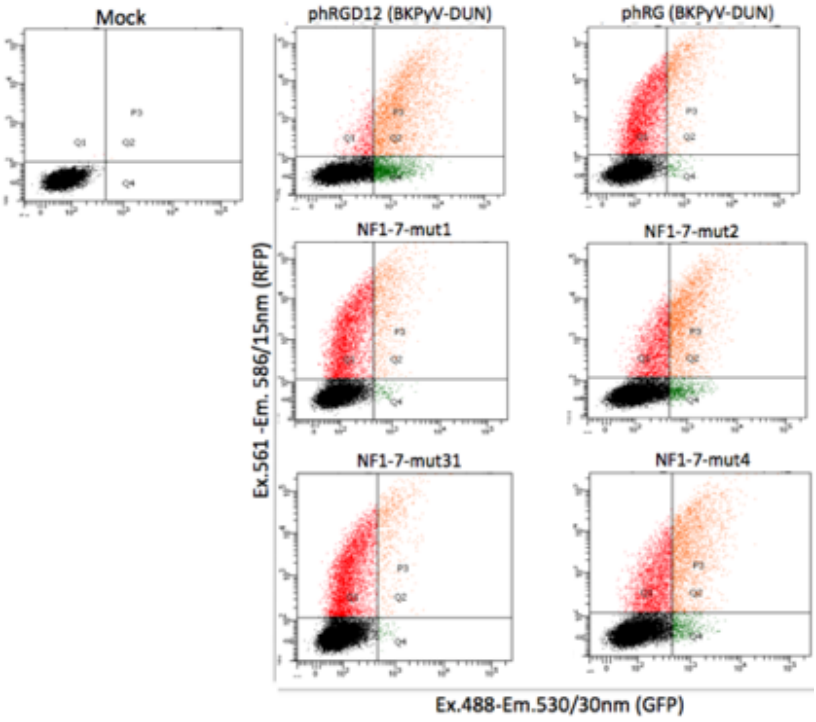
Fig. 1: (A) The bi-directional reporter vector pRG13 containing the BKPyV Dunlop NCCR (grey colored) in the early to late orientation, which was cloned via restriction sites MluI and BssHII; red fluorescence protein, dsRed2 as a marker of EVGR expression; green fluorescence protein, EGFP in opposite orientation, as a marker of LVGR expression; SV40 polyadenylation signals (SV40 polyA) in opposite orientations to stabilize the dsRed2 and EGFP messages; Colicin E1 ori for replication of the plasmid in bacteria; ampicillin resistant gene (Amp) for selection in

bacteria. The sequence below the map shows part of the NCCR, LVGR start codon, the 12-nucleotide sequence of BKPyV ww agno (underlined), BssHII restriction site and part of the EGFP sequence. This reporter contains the 12-nucleotide sequence of BKPyV agno located adjacent to the LVGR start codon (underlined). **(B)** The bidirectional reporter pRG13D12 is similar to the pRG13 except the deletion of the 12-nucleotide sequence of agno. **(C)** Analysis of flow cytometry data of BKPyV ww and dunlop NCCRs cloned in pRG13 and pRG13D12 reporter vectors and transfected in HEK293 cells. Summary plot, red bars, sum of red cells (Q_1+Q_2); green bars, sum of green cells (Q_2+Q_4); yellow bars double-positive cells (Q_2), where Q stands for the quadrants of the flows. **(D)** Normalized mean fluorescence intensity (MFI). The weighted MFI was calculated for each measurement (see formulas in Materials and Methods), late expression was normalized to BKPyVww NCCR (green MFI set as 100) while early expression normalized to BKPyVww NCCR (red MFI set as 1). Results shown are from 3 independent replicates.

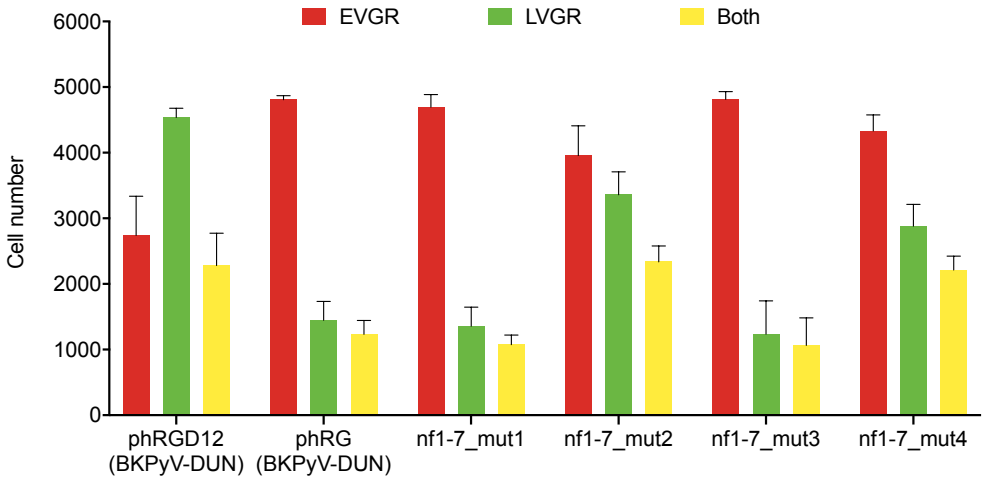
A


NF1-7
BssHII
NF1-7
AAGGCCATGGTTCTGCGCCAGGCGCGC
M V L R Q A R
nf1-7_1 AAGGCCATGGTTCTGCGGCAGGCGCGC
M V L R Q A R
nf1-7_2 AAGGCCATGCTTCTGCGGCAGGCGCGC
M L L R Q A R
nf1-7_3 AAGGCCATGGTTTTGCGCCAGGCGCGC
M V L R Q A R
BssHII
nf1-7_4 AAGGCCATCGTTCTGCGGCAGATGGCGCGC
M A R

B


Mock
phRGD12 (BKPyV-DUN)
phRG (BKPyV-DUN)
NF1-7-mut1
NF1-7-mut2
NF1-7-mut31
NF1-7-mut4
Ex.561 -Em. 586/15nm (RFP)
Ex.488-Em.530/30nm (GFP)

C



D

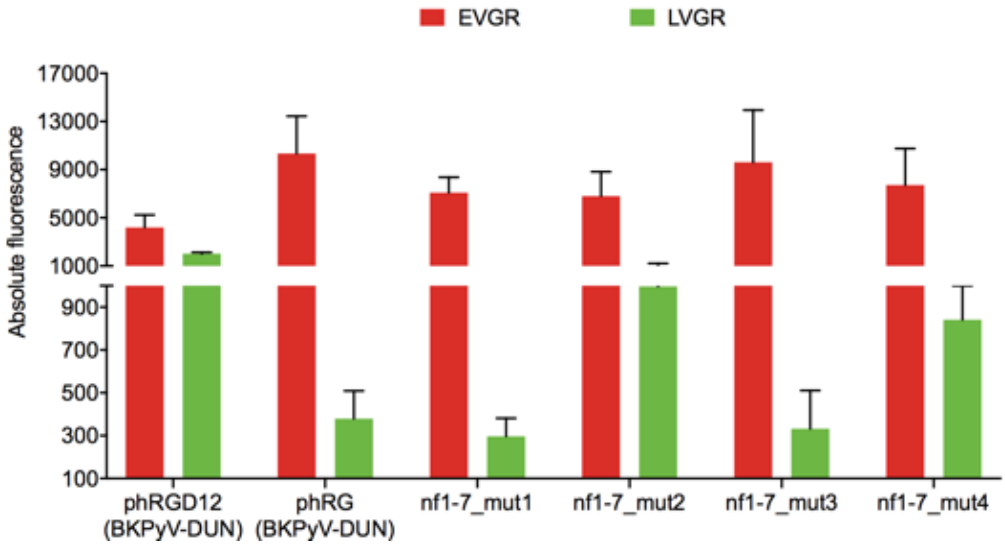


Fig. 2: (A) The 12-nucleotide sequence (shaded blue, indicating nf1 binding) of BKPyV ww agno in the phRG reporter, adjacent is the BssHII site (shaded orange). The bases in red color within the 12-nucleotide sequence are the different mutation substituted within the core NF1 binding sites which can potentially affect nf1 binding (nf1-7_mut1, _mut2, and _ mut4) or in the non-core binding sites which may not affect nf1 binding (nf1-7_mut 3). The alphabets depict the single amino acid codes. On the right of Fig2A is the sequence logo of NF1 core binding site shown in capital letters. **(B)** Analysis of flow cytometry of BKPyV Dunlop NCCR cloned in phRGD12

reporter vector (12-nucleotide sequence deleted as explained above) and also cloned in phRG reporter vector including nf1 NCCR mutants. Transfection was done in HEK293 cells. x-axis, EGFP fluorescence, y-axis, dsRed2 fluorescence, 10,000 mock transfected cells were gated for the live gate while 5,000 transfected cells were gated for the P3 (Q_1, Q_2 and Q_4) gate. Q_1, Q_4 and Q_2 depict cells expressing red, green and both, respectively. Ex. excitation wavelength, Em, emission wavelength. **(C)** Summary plot, red bars, sum of red cells (Q_1+Q_2); green bars, sum of green cells (Q_2+Q_4); yellow bars double-positive cells (Q_2). **(D)** Plot of weighted MFI was calculated for each measurement (see formulas in Materials and Methods).

This manuscript is under review.

5.2 Novel human polyomavirus non-coding control regions differ in bi-directional gene expression according to host cell, large T antigen expression and clinically occurring rearrangements

There is a rapidly expanding family of HPyVs with 13 HPyVs identified so far, suggesting many more would be identified. The question of HPyVs' host cell tropism, replication in cell cultures, route of transmission and associated diseases is raised. HPyVs widely infect humans and show differences in their associated diseases and tissue tropisms. Many aspects of the basic biology of these viruses, especially those of the novel HPyVs, are unknown. Presently, cell lines for propagating the novel HPyVs are lacking, thus hindering basic research on these emerging viruses. Currently, only BKPyV and JCPyV can be efficiently propagated in cell culture (305, 306). Attempts to efficiently propagate MCPyV in cell culture were futile, as only low-level virion production by transfection/infection of HEK293 cells was achieved with no serial transmission (17). To understand the primary host cell tropism, many studies have looked at the entry mechanisms of HPyVs into host cells by using VLPs, Vp1 pentamers and pseudoviruses (327). The NCCR activity which can be used as an indicator of HPyV secondary host cell tropism and replication in cell culture, has not been thoroughly investigated, especially for the recently discovered HPyVs. The fundamental feature of the genome organization is similar among the family of HPyVs. However, the NCCR sequences differ among HPyVs with respect to lengt,

LTag-binding sites and TFBS, which may have influence on host cell tropism and viral replication. We hypothesized that the HPyVs will demonstrate different activity in cells derived from the same or different tissues. We designed a smaller bi-directional reporter vector based on the one described in Gosert et al., (25). The bi-directional reporter vector contains the red fluorescence protein (dsRed2), which serves as a marker of EVGR, and the green fluorescence protein (EGFP) as a marker of LVGR expression. We tested the NCCR activity of the 13 HPyVs including (rr)-NCCRs from HPyV7 and HPyV9 from patients and MCPyV point mutant from MCC using HEK293 as our basic cell line. The early viral proteins (sTag and LTag) are involved in the regulation of EVGR and LVGR expression and viral genome replication. To characterize the effect of Tags on EVGR and LVGR expression, we analyzed our HPyV-NCCR reporter constructs in HEK293T cells constitutively expressing SV40 Tags and HEK293MCT constitutively expressing MCPyV Tags. Next, to get an indication of secondary host cell tropism, we investigated the NCCR activity of the 13 HPyVs in different cell lines originating from kidney, skin, lung, cervix, brain and colon. Quantification of EVGR and LVGR expression was performed by fluorescence microscopy and flow cytometry.

TITLE:

Novel Human Polyomavirus non-coding control regions differ in bi-directional gene expression according to host cell, large T-antigen expression, and clinically occurring rearrangements

Elvis T. Ajuh[1], Tobias Bethge[1*], Emma Kraus[2,3], Nicole Fischer[3], Fabian H. Weissbach[1], Rainer Gosert[1,4], Hans H. Hirsch[1,4,5#]

RUNNING TITLE: Human Polyomaviruses differ in NCCR-controlled bi-directional viral gene expression

[1]Transplantation & Clinical Virology, Department Biomedicine (Haus Petersplatz), University of Basel, Basel, Switzerland

[2] Institute of Medical Microbiology, Virology and Hygiene, University Medical Centre Hamburg-Eppendorf, Martinistrasse 52, D-20246 Hamburg, Germany

[3] Heinrich-Pette Institute, Leibniz Institute for Experimental Virology, Hamburg, Germany

[4] Division Infection Diagnostics, Department Biomedicine (Haus Petersplatz), University of Basel, Basel, Switzerland

[4] Infectious Diseases & Hospital Epidemiology, University Hospital Basel, Basel, Switzerland

[*] Present address: Centre of Laboratory Medicine, Kantonsspital Aarau, Tellstrasse, 5001 Aarau, Switzerland

KEY WORDS: polyomavirus, non-coding control region, early viral gene region, late viral gene region, T antigen, bi-directional

27

28 **#Correspondence:**

29 Hans H. Hirsch, MD, MSc

30 Transplantation & Clinical Virology,

31 Department Biomedicine (Haus Petersplatz),

32 University of Basel, Petersplatz 10, CH-4051 Basel, Switzerland

33 Phone: +41 61 207 32 62

34 Fax: +41 61 207 3283

35 E-mail: hans.hirsch@unibas.ch

36

ABSTRACT (244 WORDS)

Human polyomavirus (HPyV) DNA genomes contain three regions denoted early viral gene region (EVGR) encoding the regulatory T-antigens, late viral gene region (LVGR) encoding the structural capsid proteins Vp1-3, and non-coding control region (NCCR). The NCCR harbours the origin of viral genome replication and bi-directional promoter/enhancer functions governing EVGR- and LVGR-expression from opposite DNA strands. Despite principle similarities, HPyV-NCCRs differ in length, sequence, and architecture. To functionally compare HPyV-NCCRs, sequences from human isolates were inserted into a bi-directional reporter vector using dsRed2 for EVGR- and EGFP for LVGR-expression, respectively. Transfecting HPyV-NCCR reporter vector into human embryonic kidney (HEK)-293 cells and flow cytometry results normalized to archetype BKPyV-NCCR revealed a hierarchy of EVGR-expression levels with MCPyV-, HPyV12-, and STLPyV-NCCRs conferring stronger, and HPyV6-, HPyV9-, and HPyV10-NCCRs weaker levels, while LVGR-expression showed less variability. Transfection of HEK293T cells expressing the SV40-large T-antigen (LTag) increased EVGR-expression compared to HEK293 levels for most HPyV-NCCRs, which correlated with the number of LTag-binding sites (Spearman r 0.625; $p<0.05$). LTag activation was specifically confirmed for 2 different MCPyV-NCCRs in 293MCT cells expressing the cognate MCPyV-LTag. Clinically occurring NCCR-sequence rearrangements of HPyV7-PITT1 and -2, and HPyV9-UF1 were found to increase EVGR-expression compared to the respective HPyV archetype, but which was partly host cell type-specific. HPyV-NCCR expression in different cell lines derived from skin (A375), cervix (HeLaNT), lung (A549), brain (Hs683), and colon (SW480) demonstrated that host cell properties significantly modulate baseline the HPyV-NCCR activity.

AUTHOR'S SUMMARY (140 WORDS)

HPyV-NCCRs integrate essential viral functions with respect to host cell specificity, persistence, viral replication, and disease. Here, we show that HPyV-NCCRs differ not only in sequence length,

number and position of LTag- and common transcription factor-binding sites, but also confer differences in bi-directional viral gene expression. Importantly, EVGR-reporter expression was significantly modulated by LTag-expression and by host cell properties. For the first time, clinical sequence variants of HPyV7- and HPyV9-NCCRs containing deletions and insertions were associated with increased EVGR-expression similar to BKPyV- and JCPyV-rearrangements emphasizing that HPyV-NCCR sequences are major determinants not only of host cell tropism, but also pathogenicity. These results are essential for understanding secondary HPyV cell tropism beyond HPyV surface receptors, for identifying key viral and host factors shaping the viral life cycle, and for developing pre-clinical models of HPyV persistence, replication, and disease, to identify suitable antiviral targets.

INTRODUCTION

Polyomaviruses (PyVs) belong to the *Polyomaviridae* family, which is characterized by a restricted host range they can productively infect (1). PyV virions are non-enveloped icosahedral particles of approximately 45 nm diameter containing a circular, double-stranded DNA genome of about 5 kb (1, 2). More than 70 PyVs have been identified to date, and at least 13 are known as human polyomaviruses (HPyVs) (3). In 1971, isolation of BKPyV and JCPyV by cell culture was reported from urine of a kidney transplant patient and from brain tissue of a patient with progressive multifocal leukoencephalopathy (PML), respectively (4, 5). In the past decade, molecular techniques have led to the discovery of 11 novel HPyVs called Karolinska institute polyomavirus (KIPyV) (6), Washington University polyomavirus (WUPyV) (7), Merkel cell polyomavirus (MCPyV) (8), human polyomavirus 6 (HPyV6) (9), human polyomavirus 7 (HPyV7) (9), *Trichodysplasia spinulosa* polyomavirus (TSPyV) (10), human polyomavirus 9 (HPyV9) (11), human polyomavirus 10 (HPyV10) (12), St.Louis polyomavirus (STLPyV) (13), human polyomavirus 12 (HPyV12) (14), and New Jersey polyomavirus (NJPyV) (15). While this work was in progress, Lyon IARC PyV (LIPyV) has been described, awaiting further confirmation as HPyV (16).

The molecular detection of HPyV genomes was complemented by serological evidence of infection detecting HPyV capsid-specific antibodies. The results indicated that HPyV infections frequently occur during childhood reaching high seroprevalence rates of 60% - 90% in the general adult population, with average co-exposure rates of 7 HPyVs (17-19). Despite this high rate, clinical symptoms or signs of HPyV infection have not been identified. In fact, only 5 HPyVs have been consistently linked to disease: BKPyV to nephropathy and hemorrhagic cystitis (20), JCPyV to PML and nephropathy (21), MCPyV to Merkel cell carcinoma (8); TSPyV to *trichodysplasia spinulosa* (10), and HPyV7 to pruritic hyperproliferative keratinopathy (22), whereby disease occurs almost exclusively in inherited, acquired, or therapeutic immunodeficiency states including transplantation, HIV-AIDS, autoimmune disease, and cancer/chemotherapy (3, 23). Evidence of HPyV disease is emerging for KIPyV (24), WUPyV (25, 26), HPyV6 (27, 28), HPyV10 (12) and

NJPyV-2013 (15) due to dedicated studies correlating cytopathic alterations and virus infection by specific immunohistochemistry.

Despite this plethora of different HPyVs, the genome can be divided into three common regions: Early viral gene region (EVGR) encoding the regulatory large T antigen (LTag) and small T antigen (sTag) as well as various spliced derivatives; late viral gene region (LVGR) encoding the structural Vp capsid proteins; non-coding control region (NCCR) harbouring the origin of viral replication and bi-directional promoter/enhancer functions governing EVGR and LVGR-expression from opposite DNA strands through numerous transcription factor binding sites (TFBS) (1, 3, 29). Of note, rearrangements of the archetype NCCR of BKPyV and JCPyV have been identified in patients with nephropathy (30) and PML (21, 31), respectively, and were shown to confer higher EVGR-expression, augmented viral replication capacity, and increased cytopathology compared to their archetype counterparts (30, 32, 33). However, systematic functional comparisons of novel HPyV replication and their rearranged variants, have not been reported so far. In fact, cell culture propagation of the novel HPyVs has been difficult, which severely limits studies of viral replication and antivirals targets. However, HPyV – host cell interactions have been studied at the level of viral capsid – receptors defining primary HPyV cell tropism and entry have been characterized (29, 34-36). Yet, the NCCR defines a second level of host cell tropism intranuclearly, which is critical for coordinating and directing essential steps of HPyVs biology from latency/persistence to progression through the viral life cycle (37). To experimentally overcome the NCCR bottle neck, researchers have resorted to viral recombinants carrying genomes with hybrid NCCRs e.g. between SV40 and JCPyV (38, 39) or provided SV40 EVGR proteins like the LTag *in trans* (40-43). To understand the role of TFBS in archetype and rearranged HPyV-NCCRs, we have chosen archetype BKPyV-NCCR as a model, and introduced inactivating point mutations in 28 common TFBS (44). We identified three phenotypic groups of i) strong, ii) intermediate, or iii) low EVGR-expression and corresponding viral replication capacity (44). Interestingly, a prominent role emerged for binding sites for rather common host cell factors such as Sp1, Ets1, and NF1 (44).

125 However, as revealed for two key Sp1 sites in the EVGR and the LVGR promoter, respectively,
126 their location, directionality, and affinity resulted in graded activation of EVGR-expression at the
127 expense of LVGR-expression (45). When examining the NCCRs of the HPyVs, we found
128 differences not only in NCCR length, but also in number and composition of common TFBS and
129 LTag binding sites. We therefore hypothesized that these NCCR differences give rise to different
130 bi-directional EVGR- and LVGR-expression patterns. To this end, our results indicate the
131 presence of a hierarchy of HPyV EVGR-expression, which is modulated by host cell, LTag
132 expression, and clinically occurring NCCR rearrangements.

RESULTS

HPyV-NCCRs confer different strength of EVGR-expression

Given the prominent role of Sp1, Ets1, NF1, and LTag in the archetype BKPyV-NCCR (44, 45), we compared the archetype NCCRs of BKPyV, JCPyV and 11 novel HPyVs and found not only differences in overall length, but also in number and composition of these binding sites (46) (**Fig.1**). LTag-binding sites were predicted in all HPyV-NCCRs and preferentially located in the EVGR-promoter region with an average of 5 ranging from 1 and 2 for HPyV6- and HPyV7- to 10 and 12 for TSPyV- and MCPyV-NCCRs. NF1-sites were preferentially predicted in the LVGR, but some HPyV-NCCRs had one and less (MCPyV, HPyV6, HPyV7, HPyV9). Sp1-, Ets1-, and Spi-B-sites were found in all HPyV-NCCRs and often in clusters (**Fig.1**).

We therefore hypothesized that these differences in NCCR result in different bi-directional EVGR- and LVGR-expression. To test this hypothesis, we compared 13 archetype HPyV-NCCRs using a novel bi-fluorescent reporter vector pRG13D12 recapitulating the PyV genome organization regarding bi-directional EVGR and LVGR (**Fig.2A**). To validate its suitability **(Fig. 2B),** NCCRs of the well-characterized archetype BKPyV(ww) and the rearranged BKPyV-Dunlop strain were inserted in forward orientation or the BKPyV-Dunlop strain NCCR was inserted in reversed orientation, the resulting reporter constructs transfected into HEK293 cells, and analysed by flow cytometry at 48 hours post transfection (hpt). Indeed, the rearranged BKPyV-NCCR(DUN) conferred significantly stronger EVGR-expression (red) compared to archetype BKPyV(ww) NCCR (**Fig. 2C**) and the BKPyV-Dunlop NCCR cloned in reversed orientation displayed opposite expression compared to the BKPyV-Dunlop cloned in forward orientation, in line with previous results (30, 44)**.** Quantification of the number of fluorescent cells (**Fig. 2D**) and the mean fluorescent intensity (MFI) indicated that the simplified reporter vector pRG13D12 captured these differences (**Fig. 2E**), hence being suitable for a principle comparative analysis of HPyV-NCCRs.

To that end, 13 archetype HPyV-NCCR sequences were inserted into pRG13D12, verified by

159 sequencing, and transfected into human embryonic kidney HEK293 cells. EVGR- and LVGR-
160 expression were quantified by flow cytometry and the results normalized to the archetype BKPyV-
161 NCCR(ww). The results demonstrated that EVGR-expression in HEK293 cells varied over more
162 than 3 orders of magnitude (**Fig. 3A**). MCPyV- and HPyV12-NCCR were located at the upper end
163 of the EVGR-responses, whereas HPyV6- and HPyV9-NCCRs were found at the lower end. The
164 corresponding LVGR-expression also showed some variability, but tended to be within the same
165 order of magnitude (**Fig. 3A**). The results indicated that HPyV-NCCRs give rise to a hierarchy of
166 EVGR-expression levels in HEK293, which were higher, similar, and lower than the corresponding
167 archetype BKPyV-NCCR activity in HEK293.
168
169 **Large T-antigen activates HPyV NCCR-controlled EVGR-expression**
170 Since LTag is a major regulatory protein encoded in the EVGR, exerting key functions on viral
171 replication and gene expression, part of which are mediated directly through LTag-binding sites
172 (**Fig.1**, red triangles), it was of interest to investigate the effect of LTag on HPyV-NCCR reporter
173 expression. To ensure comparable conditions for all 13 HPyV-NCCRs, the well-characterized
174 HEK293 derivative 293T cells were chosen, which constitutively express SV40-LTag (45).
175 Transfection and flow cytometry analysis 48 hpt showed that the overall hierarchy of EVGR-
176 expression was only little changed, whereby MCPyV and HPyV12 were in the higher group, and
177 HPyV6 and HPyV7 remained in the lower group (**Fig. 3B**). However, JCPyV EVGR-expression
178 appeared to be most responsive to SV40-LTag and now located among those HPyVs conferring
179 equal or higher NCCR-driven EVGR-expression than BKPyV. Compared to HEK293, EVGR-
180 expression in HEK293T cells increased more than 10-fold for some HPyV-NCCRs including
181 BKPyV, whereas only little change was observed for HPyV6. LVGR-expression also increased
182 but to lesser extent (**Fig.3C**). Thus, most, but not all HPyV-NCCR were able to respond to SV40-
183 LTag *in trans* with significantly increased EVGR-expression.
184

As noted above, the LTag-binding sites differed in number and location among the different HPyV-NCCRs (**Fig.1**). HPyV6- and HPyV7-NCCR contained the lowest number (1 and 2 sites, respectively) of LTag-binding sites, while MCPyV-NCCR harbours the highest number (12 sites) of LTag binding sites, which appeared to partly correlate with the EVGR-expression levels in HEK293T cells (Spearman correlation, Spearman r 0.625; $p<0.05$; **Fig.3D**). Of note, HPyV12 and TSPyV deviated by showing significantly higher and lower levels, respectively. Together, the data indicated that SV40-LTag *in trans* provided an important stimulus to the basal EVGR-expression of most HPyV-EVGR, but also suggested that other factors related to the primary HPyV-NCCR sequence mattered.

Autologous MCPyV-LTag increases MCPyV-NCCR EVGR-expression

Although SV40-LTag has considerable homology to the LTag encoded in HPyV genomes (29, 37), differences in amino acids and in the host range domain has been noted, which can be grouped in one of seven clades (47). We wondered whether or not the known strong and pleiotropic action of SV40-LTag in 293T cells could be an appropriate indicator and surrogate of the cognate viral LTag. We addressed this question for MCPyV by examining two MCPyV-NCCR reporter constructs in 293MCT (48) expressing the cognate MCPyV-LTag (**Fig.4A**). One NCCR had been detected in healthy skin and was tested above (MCPyV-R17b) (9), while the other one had been detected in a Merkel cell carcinoma carrying 1 C->G substitution and 1 ‘A’ base deletion in an A-rich sequence stretch (MCPyV-MCVw156) (49). The results demonstrated increased EVGR-expression from two different MCPyV-NCCRs in 293MCT cells, thus independently supporting the results obtained with SV40-LTag in 293T cells (**Fig. 4BCD**).

Rearranged NCCR patient variants increase EVGR-expression

Clinical NCCR sequence variants of the novel HPyVs have been described (22, 28, 50), but their effect on bi-directional EVGR- and LVGR-expression has not been compared with the respective

211 archetype NCCRs. We therefore examined the rearranged HPyV7-NCCR variants -PITT1 and -
212 PITT2 (**Fig.5A**) detected in two lung transplant patients with hyper-proliferative keratinopathy (22).
213 The data demonstrated a significant increase of HPyV7-PITT1 and -PITT2-EVGR-expression in
214 the skin cell line SW480, while only a trend to higher levels was observed in HEK293 (**Fig.5A**).
215 For the rearranged HPyV9-UF1 (**Fig.5B**), a significant increase of EVGR-expression over
216 archetype HPyV9-NCCR activity was observed in both, HEK293 and in the lung cancer line A549,
217 whereas little change in LVGR-expression was noted (**Fig.5B**). These results indicated that
218 naturally occurring NCCR rearrangements of the novel HPyV7 and -9 are able to confer increased
219 EVGR-expression, but that this effect also depended on the host cell context.
220
221 **Role of host cells for bi-directional NCCR-expression**
222 Given the role of the host cell context suggested above, we analysed the NCCR-driven EVGR-
223 and LVGR-expression of the 13 HPyV-NCCR reporter constructs in different cell lines derived
224 from skin, lung, cervix, brain and colon (**Table 1, Table 2**). Indeed, HPyV-NCCR showed
225 differences in EVGR- and LVGR-expression according to the host cell lines tested (**Table 1**). Thus,
226 MCPyV demonstrated the strongest EVGR-expression in the skin-derived A375 cells, and
227 intermediate levels in the epithelial cell lines derived from kidney and colon, whereas cervix, lung,
228 and brain cell lines showed rather low EVGR-expression. HPyV12-NCCR conferred stronger
229 expression in brain and colon, followed by lung, cervix and skin cell lines, while remaining relatively
230 low in kidney-derived cell lines. JCPyV-NCCR EVGR-expression was highest in brain-derived
231 cells, whereas BKPyV-NCCR showed highest EVGR-expression in kidney, followed by lung- and
232 colon-derived cell lines. KIPyV-NCCR driven EVGR-expression was highest in cervix and brain
233 followed by lung, while WUPyV-NCCR was highest in the colon-derived cell line (**Table 1**).
234 Complementary LVGR-expression levels were observed for these HPyV-NCCRs, some of which
235 decreased as the EVGR-expression levels were higher. Taken together, the data demonstrate
236 that the bi-directional activity of a given HPyV can substantially differ in different host cells.

237 **DISCUSSION**

238 The results of the present study demonstrate that HPyV-NCCRs not only differ in sequence length,
239 number and position of binding sites for LTag and rather common factors like Sp1, NF1, and Ets1,
240 but also confer significant differences in viral gene expression. This difference in basal EVGR
241 expression is most impressively captured in HEK293 cells by the high levels seen for MCPyV- and
242 HPyV12-NCCRs at the upper end, and which range to the 2 orders of magnitude lower levels seen
243 for HPyV6-NCCRs. LVGR-expression was generally strong, but varied less among the different
244 archetype HPyV-NCCRs as described for the archetype BKPyV-NCCR(ww) (30, 32).

245 The functionality of the bi-directional MPyV-NCCR reporter activity was further addressed by
246 providing the SV40 homolog of the EVGR-encoded LTag *in trans* (29, 51) in 293T cells, which
247 showed increased EVGR-expression levels for practically all HPyV-NCCRs. This indicates that
248 the reporter construct recapitulated an essential NCCR response described in detail for the viral
249 life cycle of polyomavirus representatives such as SV40 or BKPyV (1, 29, 30, 44). LVGR-
250 expression was also increased, although to a lesser extent in line with the bi-directional balance
251 of EVGR- versus LVGR-expression (45). Although SV40-LTag has been recognized and exploited
252 as a strong pleiotropic activator of HPyV gene expression, the autologous LTag of some HPyVs
253 might confer only selective responses as reported for the effect of JCPyV-LTag on its cognate
254 JCPyV-NCCR that are not necessarily captured by the SV40- or BKPyV-NCCR (52). Since we
255 could not assess the specific impact of each of the 13 HPyV-LTag orthologues on their respective
256 NCCR, our observations were supported by the impact of MCPyV-LTag in 293MCT (48)
257 demonstrating a similar activation of EVGR-expression for two different MCPyV-NCCRs, one from
258 a healthy control and one from a Merkel cell carcinoma bearing a G-to-C point mutation.

259 The role of the HPyV-NCCR sequences as differential determinants of EVGR- and LVGR-
260 expression was further strengthened by the correlation of the number of LTag-binding sites and
261 the level of EVGR expression in 293T cells. However, there were also two notable exceptions

such as HPyV12-NCCR, which showed much higher EVGR-expression than predicted from the number of LTag-binding sites, whereas the opposite was true for the TSPyV-NCCR EVGR-response. This suggested that other regulatory elements were critical in determining the EVGR-response. In this context, the higher expression of rearranged BKPyV(DUN)-NCCR compared to the archetype BKPyVww should be noted, which carries a deletion of one LTag site together with the high-affinity *SP1-4* site in the LVGR-promoter (45). Deletions and insertions in HPyV7-PITT1, HPyV-PITT2, and HPyV9-UF1 were also associated with increased EVGR-expression despite rather small alterations of the primary sequence. Thus, HPyV-NCCR sequences are key determinants of EVGR-activity and suggests for the first time that the pathogenicity of the respective variants would likely be increased in susceptible host cells as shown before for clinical variants of BKPyV- and JCPyV-NCCRs (30, 33).

Finally, we noted that HPyV-NCCR expression levels differed in different host cell lines emphasizing that the combined cellular make-up of transcription factors and other regulatory proteins is essential in sensing and interpreting the HPyV-NCCRs with respect to viral gene expression and persistence (37). As this work was in progress, Moens and colleagues independently reported the characterization of HPyV-NCCRs in different cell lines using a unidirectional luciferase reporter assay, which also identified higher EVGR-expression by MCPyV- and HPyV12-NCCRs as well as by TSPyV-NCCR (53), but the latter was not confirmed in our study. Although uni-directional reporter assays have been commonly and successfully used to study uni-directional cell and viral promoters, they remain challenging for a bi-directional gene expression organisation within a small DNA sequence of approximately 500bp (**Fig.1**). Part of this stems from the difficulty to generate well-justified and controlled truncations of the intricate intertwined and competing functional elements of the bi-directional HPyV-NCCR (44, 45), which may-be difficult to accomplish given the pronounced effects of NCCR point mutations (44) and even simple rearrangements seen. Together with our previous analyses (30, 33), we are confident that the bi-directional reporter assays used here recapitulates the HPyV-NCCR activity, which can

288 be robustly enumerated in single cells using flow cytometry and includes estimates of non-
289 transfected non-fluorescent cells (44, 45).

290 Our comprehensive analysis of the bi-directional HPyV-NCCR reporter expression using flow
291 cytometry provides also estimates for different model cell lines derived from kidney, skin, lung,
292 cervix, colon, and brain. The results suggest that there may be underlying mechanisms linked to
293 cell differentiation and the respective HPyV-NCCR as expected for a co-evolutionary relationship.
294 Thus, EVGR expression of MCPyV-NCCR was even stronger in skin-derived cell line A375 than
295 in the kidney derived reference HEK293 or 293T derivative, whereas that of BKPyV-NCCR was
296 strongest in kidney-derived, and that of WUPyV in colon-derived cell lines (**Table 1**). Although
297 identifying these promoting or restricting factors in the respective cell lines require further study,
298 their non-random expression profile is intriguing and suggests a first functional characterization of
299 secondary HPyV host cell tropism through this bi-directional reporter assay. The combination of
300 host cell and HPyV-NCCR reporter vector should be amenable to molecular screening approaches
301 using loss-of-function e.g. through shRNA knock-down libraries, or gain-of-function approaches
302 e.g. through expression library, or corresponding small compound libraries. Clearly, our correlation
303 of functional response and NCCR sequence does presently not allow to simply delineate
304 candidate sequences or factors, but requires fine mapping of major transcription factors binding
305 sites as done for the BKPyV-NCCR (44) through further systematic work. In this regard, the current
306 deficiencies of straight-forward HPyV culture and replication models must be acknowledged as a
307 limitation. This obstacle is somewhat at odds with the rather frequent detection of certain HPyVs
308 such as HPyV6 and HPyV7 on healthy skin and the rather poor NCCR-expression levels observed
309 here as well as by others, and suggest that important host cell factors must be identified first before
310 cell culture work can be successful.

311 The high number of HPyVs in humans continues to surprise (54), as comprehensive serological
312 studies using specific Vp1-based IgG detection indicate that most humans have been infected

313 with more than one (17-19, 55), whereby serological evidence points to an average of 6 to 7 HPyV
314 co-infections (19), some which may co-exist in the same organ as shown for skin and kidney, and
315 may show direct positive and negative viral interactions (54).

316 In summary, the HPyV-NCCR mediates key functions of polyomavirus biology including
317 persistence of the episomal viral genome in the host cell nucleus as well as timing and sequential
318 steps of the viral replication cycle. Despite some limitations, our results are informative and
319 represent an important step towards understanding secondary HPyV cell tropism beyond HPyV
320 surface receptors. Using this approach raises novel perspective of identifying key viral and host
321 factors shaping the viral life cycle, and for developing models of HPyV infection, replication,
322 disease to identify suitable antiviral targets.

323

324 **ACKNOWLEDGEMENT**

325 This research was supported by an appointment grant of the University of Basel to HHH.
326 The funders played no role in the topic, results, interpretation and writing of this
327 manuscript.

328 **MATERIALS AND METHODS**

329 **HPyV NCCR and reporter constructs**

330 Based on the phRG1 reporter vector (30) recapitulating the principle HPyV genome organisation
331 with respect to EVGR and LVGR (**Fig. 2A**), a smaller bi-directional reporter pRG13D12 was
332 constructed, in which the expression cassette for hygromycin resistance was removed, and the
333 reporter genes were placed upstream of SV40 polyadenylation sites instead of those from beta-
334 globin **(Fig. 2B)**. The HPyV-NCCRs were chemically synthesized in pUC57 (Eurogentec S.A,
335 Belgium) (**Table S1**) and excised using the restriction enzymes BssHII and MluI (New England
336 Biolabs, England), and cloned into the corresponding restriction sites of pRG13D12. HPyV-NCCR
337 constructs were verified by Sanger sequencing for correct NCCR sequence and orientation using
338 the 3130 Genetic Analyzer (Applied Biosystems, Switzerland).
339
340 **Cell lines**

341 The cell lines, their origin and providers, and the standard culture conditions are listed in **Table 2**.
342 All cell lines were cultured in a HERA cell-150 incubator (Thermo Fischer Scientific, Switzerland)
343 at a temperature of 37°C and 5% CO_2.
344
345 **Transfection**

346 Cell transfection was done in a ratio of 1:3 DNA to transfection reagent using 3µL Lipofectamine
347 2000 (Thermo Fisher Scientific) and 1µg HPyV NCCR reporter constructs in a 12 well plate, except
348 for SW480 cells that were transfected with 3µL Lipofectamine3000 (Thermo Fisher Scientific) and
349 1µg HPyV NCCR reporter constructs. For each well, 3µl of lipofectamine2000 (or
350 lipofectamine3000) was diluted in 100µl of Opti-Mem (Thermo Fisher Scientific) medium, flicked,
351 spun down and incubated at room temperature (RT) for 5 minutes. 1µg of plasmid DNA was diluted
352 in 100 µl of Opti-Mem medium for each well. Transfection reagents and plasmid DNA were then
353 mixed, incubated for 5 minutes at RT and added to the cells with a confluency ranging between

70-90%. 24 h post-transfection (hpt), medium was replaced with DMEM with 10% FBS or DMEM-H with 10% FBS complemented with 2mM Glutamine. 48 hpt, fluorescence images were taken by fluorescent microscopy and EVGR (RFP) and LVGR (GFP) expression was quantified by flow cytometry.

Flow cytometer-based quantification

Cells were washed once at 48 hpt with 1mL PBS/2.5 mM EDTA (BioConcept, Switzerland) and then treated with trypsin (without phenol red)/0.25mg/mL EDTA (Lonza, MD, USA). The cells were suspended in 1mL DMEM (without phenol red) containing 20mM HEPES, 2mM Glutamine, 1% FBS and 1% Penicillin/Streptomycin (P/S) (Thermo Fischer Scientific, Switzerland) and transferred to 5 mL polystyrene round-bottom tubes (BD, Franklin Lakes, NJ, USA). Prior to each measurement, DAPI (D8417, Sigma-Aldrich, St. Louis, MO, USA) was added to mark dead cells in a final concentration of 1 ng/mL. A Fortessa Cytometer (Becton-Dickinson, Franklin Lakes, NJ, USA) was used as follows: For RFP, excitation 561 nm (Yellow-Green Laser), emission 586/15 nm; for GFP, excitation 488 nm (Blue Laser), emission 530/30 nm; DNA staining with DAPI, excitation 405 nm (Violet Laser), emission 450/50 nm. Transfection with pUC19 was used to set the gates for non-fluorescent transfection control cells in quadrant Q3, the cell numbers in the corresponding quadrants were determined (Q1 for red-only cells; Q2 for red and green cells; Q3 for non-fluorescent cells; Q4 for green-only cells). The number of fluorescence cells for red, green and both were calculated as follows: red = number of cells in Q1 and Q2, green = number of cells in Q2 and Q4, red and green = number of cells in Q2. The mean fluorescence intensities (MFI) for EVGR (red) and LVGR (green) expression were calculated by inserting the cell number (N), mean fluorescence (I) of the quadrants Q1 (red cells only), Q2 (red and green cells), and Q4 (green cells only) into the respective formulas: $\mathrm{MFI(red)} = \frac{(N_{Q1}*I_{Q1})+(N_{Q2}*I_{Q2})}{(N_{Q1}+N_{Q2}+N_{Q4})}$; $\mathrm{MFI(green)} = \frac{(N_{Q2}*I_{Q2})+(N_{Q4}*I_{Q4})}{(N_{Q1}+N_{Q2}+N_{Q4})}$.

EVGR-expression of all HPyV constructs was normalized to the EVGR (red) and LVGR (green)

expression of the archetype BKPyV-NCCR(ww) set as 1 and 100%, respectively. Dot plot images from flow cytometer were processed in Adobe Illustrator CS4 14.0.0.

LT-Ag expression in 293 cells stably transfected with MCPyV early gene region (293MCT cells). 293 cells were transfected with 100ng linearised plasmid DNA encoding for MCPyV early gene region. After puromycin selection cells were lysed and analyzed by immunoblotting for MCPyV LTag expression using the mouse monoclonal MCPyV LTag antibody Cm2B4. Equal protein amounts loaded (30µg) were ensured by reincubating the membrane with an anti-actin antibody, Chemicon Cat.No.1501 (lower Blot). Parental cell line 293 negative for MCPyV LTag expression was used as a control on the immunoblot.

In silico analysis of HPyV NCCRs

The MatInspector software tool (Genomatix, Munich, Germany)(46) was used to search for potential transcription factor- and LTag-binding sites within the HPyV NCCR. The NCCR sequences were uploaded into the MatInspector software tool using the settings described for BKPyV (45). For the identification of potential LTag binding sites the consensus sequence GRGGC (where R = A or G) was used allowing no mismatch.

Statistics

The mean and standard deviations were calculated from three independent transfections using Microsoft Office Excel for Mac 2011 and GraphPad Prism software (version 7.0c for Mac OS). Multiple comparisons 2-way ANOVA was used for calculating the statistical differences between NCCR reporters construct expression in different cell lines, and 2-way comparisons as appropriate. Correlation analysis was done by non-parametric Wilcoxon test, or log-linear Spearman regression.

APPENDICES

Table S1: NCCR sequences of HPyVs used in this study. The GenBank accession number and NCCR sequence including the start codons (CAT and ATG) of EVGR and LVGR, respectively are shown for each HPyV.

HPyVs	GenBank	NCCR sequences
BKPyV-DUN	KP412983	CATTTTTGCAAAAATTGCAAAAGAATAGGGATTTCCCCAAATAGTTTTGCTAGGCCTCAGAAAAAGCCTCCACACCCTTACTACTTGAGAGAAAGGGTGGAGGCAGAGGC GGCCTCGGCCTCTTATATATTATAAAAAAAAAAGGCCACAGGGAGGAGCTGCTTACCCATGGAATGCAGCCAAACCATGACCTCAGGAAGGAAAGTGCATGACTCACAGG GGAATGCAGCCAAACCATGACCTCAGGAAGGAAAGTGCATGACTCACAGGGAGGAGCTGCTTACCCATGGAATGCAGCCAAACCATGACCTCAGGAAGGAAAGTGCAT GACAGACATGTTTTGCGAGCCTAGGAATCTTGGCCTTGTCCCCAGTTAAACTGGACAAAGGCCATG
BKPyVww	JQ513592	CATTTTTGCAAAAATTGCAAAAGAATAGGGATTTCCCCAAATATTTTGCTAGGCCTCAGAAAAAGCCTCCACACCCTTACTACTTGAGAGAAAGGGTGGAGGCAGAGGC GGCCTCGGCCTCTTATATATTATAAAAAAAAAGGCCACAGGGAGGAGCTGCTAACCCATGGAATGTAGCCAAACCATGACCTCAGGAAGGAAAGTGCATGACTGGGCAG CCAGCCAGTGGCAGTTAATAGTGAAACCCCGCCCCTGAAATTCTAAATAAACACAAGAGGAAGTGGAAACTGGCCAAAGGAGTGGAAAGCAGCCAGACAGACATGTTT TGCGAGCCTAGGAATCTTGGCCTTGTCCCCAGTTAAACTGGACAAAGGCCATG
JCPyV	AB038249	CATTTTAGCTTTTTGCAGCAAAAAATTAGTGCAAAAAAGGGAAAAACAAGGGAATTTCCCTGGCCTCCTAAAAAGCCTCCACGCCCTTACTACTTCTGAGTAAGCTTGGAG GCGGAGGCGGCCTCGGCCTCCTGTATATATAAAAAAAAGGGAAGGTAGGGAGGAGCTGGCTAAAACTGGATGGCTGCCAGCCAAGCATGAGCTCATACCTAGGGAGCC AACCAGCTGACAGCCAGAGGGAGCCCTGGCTGCATGCCACTGGCAGTTATAGTGAAACCCCTCCCATAGTCCTTAATCACAAGTAAACAAAGCACAAGGGGAAGTGGAA AGCAGCCAGGGGAACATGTTTTGCGAGCCAGAGCTGTTTTGGCTTGTCACCAGCTGGCCATG
KIPyV	EF127906	CATTTTGCCTCTCTAGGCCTCTCAAATGCCTCTGCAGGCCCTCTCCTTCTTCTAGAAAAAGCTGGGGCTTTTTTGGCCTCTGGCCTCCTGTAATATAGAAAAAAAGGGCACA GTGATTGACAGTTGTGTATACAAGCATGTGTGGTATGTTTAGTGTGTAAGCCAATAAAGTTAAAGGTCACTACTGTGGGTGGTGACACCTGATACCGGCCGGAACTAGTTG CTACGTGCCACACAATAGCTTCACTCTTGGCGTGAAGCCAACTTCCTGGGCCGTGAGCCAGCTTCCTGCGGCCTTGTGTTTTTTACCAACACACCTGGTGAACTTCTACTG TCCTTGACACAGGTAAGACTGGGGACCCTTTGTAGGCCAAAGGAGAGTGAAGGGTAACTGAAATGCTAAGACTGTAAGTTCTAATCCTAGTATTTCAGTTCGGGGATGTT GGCGCCATCGTCTCGAACCTGGCCTGCATACCTTTGGATATAGAGGGTCACCAATTTTTATTTTGTTTTTAGATG
WUPyV	NC_009539	CATTTTGCCTTCTTTAGCCTCAAGGCGCCTCAGCAAGGCCCTCTGCTTTTAGTTCAGAAAGTTGAGGCTTTTTAGGCCTCAGGCCTCCTTATTATAATAAAAAAAAAGCTAAG CATGATTGACAGTGTGGGCTAACCAAAAGCACAAGAACAAAGCTTTTAGCCAATTAGCAGCCACAAGGTGGAGCAAAAGTATTAAGTTTCACTGTTATGTGCAGGAAT GTGCAGCTGTGACCTTTTAAAGTTTCCGGGCACGGCGCCAACTTCCTGGGCCTGGTGCCATACCAACACAGCTGCTGAGCTTCCGGAATACAATACTGGTGCCCTTTGTAA GTGTTTTACAGGTAAGTAAGGCCTACAACAGGGCTTATTTGTACTATAAGTTAATGGGGGCCCTTTGTAGTCCAGCGGAAAGTGAAGGGTGGCTTAACAGAGACGTCCTT GGGTTCAAACCTAAGGGTGCCATAAGCAACATTACATTAATGTTGTGACATCTCCAGTCGGGGGTATTGGCCTATAGGAAACCCTAGGGCTCTATAAGCAGCATACATAT GTTGTGACATCTCCGTTGAGTCTGGGGGTATTGGTGCTACCGTCTCGAACCTAGCCGACGCCGTTGGATATAAAGGGTCACCATTTTTATTTCAGATG
MCPyV-R17b	HM011556	CATCTTGTCTATATGCAGAAGGAGTTTGCAGAAAGAGCAGAGGAGCAAATGAGCTACCTCACTAAGGAGTGGTTTTTATACTGCAGTTTCCCGCCCTTGGGATCTGCCCTT AGATACTGCCTTTTTTGCTAATTAAGCCTCTTAAGCCTCAGAGGCCTCTCTCTTTTTTTTCCAGAGGCCTCGGAGGCTAGGAGCCCCAAGCCTCTGCCAACTTGAAAAAAAA AAGTCACCTAGGCAGCCAAGTTGTGGTTACATGATTGAACTTTTATTGCTGCAGGGTTTCTGGCATTGACTCATTTCCTGGAGAGGCGGAGTTTGACTGATAAACAAAACTT TTTTTCTTTCTGTTTGGGAGGGAGACGGAAGACTCTTAACTTTTTTTCAACAAGGGAGGCCCGGAGGCTTTTTTTTCTCTTACAAAGGGAGGAGGACATTAAAAGAGTAAG GTATCCTTATTTATTTTTCAGGATG
MCPyV-MCV156	HM355825	CATCTTGTCTATATGCAGAAGGAGTTTGCAGAAACAGCAGAGGAGCAAATGAGCTACCTCACTAAGGAGTGGTTTTTATACTGCAGTTTCCCGCCCTTGGGATCTGCCCTT AGATACTGCCTTTTTTGCTAATTAAGCCTCTTAAGCCTCAGAGGCCTCTCTCTTTTTTTTCCAGAGGCCTCGGAGGCTAGGAGCCCCAAGCCTCTGCCAACTTGAAAAAAAA AGTCACCTAGGCAGCCAAGTTGTGGTTACATGATTGAACTTTTATTGCTGCAGGGTTTCTGGCATTGACTCATTTCCTGGAGAGGCGGAGTTTGACTGATAAACAAAACTT TTTTTCTTTCTGTTTGGGAGGGAGACGGAAGACTCTTAACTTTTTTTCAACAAGGGAGGCCCGGAGGCTTTTTTTTCTCTTACAAAGGGAGGAGGACATTAAAAGAGTAAG TATCCTTATTTATTTTTCAGGATG
HPyV6	HM011563	CATTTTGCTGCTTGAGTAGTCTCTGAATGAAATGTTAAAATCCCAGACAGCAGGTGAAATCCCTGTGCAGGCTTACGCAACTGGGCAGGGCCATTTAAGCTCCCTTATCTC TCTAATTATAGGAGGCAATAAAGGCCACCAGGCCACCTCAATGTATGAGAAAAAAGGAGAGGAAATAGGGCAACCAAAGGTCAAAAGGAAGTTATTAGGCGAGAAAAG CCCGCCAAATTTTTCCCAGTCATAACTGAGGTTGACCACCGTTGACACAACACCCTAATTAGTAAGTTCTTCAATTTCTGTTTTATTTTGTTTAACTCTTCAGTGACCTCAC CGCCTGTTGACCCTGGTCTTAATTTCTACTTAACAGGTAAGCCATG
HPyV7	HM011567	CATTTTCCCTTTCTCTCTGTAATTTCTGAAGAGAAAGTGGTGGCAGGTGCAGCTGACACCCTTTAAATAGTCTTCTCTCTGCAATAGAAGGGTGGAGGCAAGAGGCCACCA AGCCACCTACATTAGGAAAAAAAAAGGTGGAGGAGGAAATGTCTGGGTTACTGGGTCAGACAGGATATGATTAGCATAATATTAGCCCGCCAAAAGCTGTGGTTTTTATG CAGATGAGGTGTGACTTTGGTGGAGGACCTATCTCTCTATCTGTAAGTAAGAATTTTTCACTTTTTAACAAATTTTAACCTCTTAGTAACTGTAGTGTTCTCTCTCAGA AAGTCCTTGATATTGTTGGCCATTTGGGTGCTGTGTCAAGGACTTAGGTAAGAGAATG
HPyV7-PITT1	KJ733012	CATTTTCCCTTTCTCTCTGTAATTTCTGAAGAGAAAGTGGTGGCAGGTGCAGCTGACACCCTTTAAATAGTCTTCTCTCTGCAATAGAAGGGTGGAGGCAACAGGCCACCA AGCCACCTACATTAGGAAAAAAAAAGGTGGAGGAGGAAATGTCTGGGTTACTAGGTCAGACAGGATATGATTCAGACAGGATATGATTAGCATAATATTAGCCCGCCAAA AAGTGGTTTTTATGCAGATGAGGTTTCACCTTTGGTGGAGGACCTATCTCTCTATCTGTAAGTAAGAATTTTTCACTTTTTAACAAATTTTAACCTCTTAGTAACTGTAGTTTT CTCTCTCCTCAGAAAGTCCTTGATATTGTTGGCCATTTGGGTGCTGTGTCAAGGACTTAGGTAAGAGAATG
HPyV7-PITT2	KJ733013	CATTTTCCCTTTCTCTCTGTAATTTCTGAAGAGAAAGTGGTGGCAGGTGCAGCTGACACCCTTTAAATAGTCTTCTCTCTGCAATAGAAGGGTGGAGGCAAAAGGCCACCA AGCCACCTACATTAGGAAAAAAAAAGGTGGAGGAGGAAATGTGTCAGACAGGATATGATTAGCATAATATTAGCCCGCCAAAAAGTGGTTTTTATGCAGATGAGGTTTCA CCTTTGGTGGAGGACCTATCTCTCTATCTGTAAGTAAGAATTTTTCACTTTTTAACAAATTTTAACCTCTTAGTAACTGTAGTTTTCTCTCTCCTCAGAAAGTCCTTGATATTG TTGGCCATTTGGGTGCTGTGTCAAGGACTTAGGTAAGAGAATG
TSPyV	GU989205	CATTTTGCTGAATGCACCAGAAGACAGGTAAGGGGAGAAATGAAGAAATGAAGATGGGAGCTTTTTGGAGCCCTTTTAAAGCCTCTGTGTGCCTCAATTAATTTTCATCA GATTTGGCTTACATAAAAGGGAGGAAGTGCCAAGCCTCAAGAGGCCTCCGGAGGCCTCTCCCTCTATGTCTGTGTCAAGAGGGGCAGGAAGCCTCAGCCTCCTGGTAA TACTGAGAAAAAACAAGATTTCCATAGTAACCCCCTATTAGGAATTTGAGGAACTATACTTAGAGTGTTGACATAAATAGGTCATTGTTTGCTTAAGACCGTTATCCTGAG GAAGCTCGTGCGCCACCTATCATTTGTTATGGAAAATGCTGTGTCAGTTAGTCTCAAGGGGATGTTTGTCATACTGCCACAAACACAGGAAGGCAGCCAAGAGACAAGAG GAAGCCAGCCAAATCCCGTGATTGTTTTTGTATATTCTTAAGATAGCACTTGTCTAACTGCTCTCATCCTGTCCGTCCGGTCTCCATAACAACACATCCTCCCGCTCTGTGTT CTCACGCGGCAAATTCAAATTCTAATGGTAAGTTTTTCTTATTTTTCAGGTAAAATG
HPyV9	HQ696595	CATGGCTCTGCAAAAAGTAAAATAAGTCTTACTACCTGAGAATCAAGTTAATTAAGTTTCAAATAGGTGGAGAACTTCTTACCTTAATGAGTTTTGGAAAAGCCTCCAAAG CCTCTCTCTTTGTTGCAACAAGAGAAGGAGGCAAGGGGCCCCTGGCCTCTTATACTATCAAAAAAAACCTGTGTTGCCATAGTGATTTGAGTGACAGGGAGATAACTTT GCAAATTGAGGAACTAAATGACAGGTTATTTTGCAGAATCAACTCTAGAGGAAACTGTGGTATCTGTGGTATCTACTGCTGCCAGTTGACCTGGGGACTTTTGGTGTTGGT ACATTCCGGTGTGTGACAAACAAGTGCATTGCGGTTTCCGTTGCTAGGTGGCGCTTAGCAACCCCTCTAGATACCAGTAGCTAAGGGGGAAGTGAAATCATTTTGCCATC TAAGGTCAACAGAGAAGAAGCCCGCCTAAACTAAACGCGCCAATTTTGAACTTGGGTAAGATATG
HPyV9-UF1	KC831440	CATGGCTCTGCAAAAAGTAAAATAAGTCTTACTACCTGAGAATCAAGTTAATTAAGTTTCAAATAGGTGGAGAACTTCTTACCTTAATGAGTTTTGGAAAAGCCTCCAAAG CCTCTCTCTTTGTTGCAACAAGAGAAGGAGGCAAGGGGCCCCTGGCCTCTTATACTATCAAAAAAAACCTGTGTTGCCATAGTGATTTTGAGTGACAGGGAGATAACTTT GCAAATTGAGGAACTAAATGACAGGTTATTTTGCAGAATCAACTCTAGAGGAAACTGTGGTATCTACCTGCTCCACTTGACCGCCCGGGGACTTTTTGGTGACATTCCG GTGTGTGACAAACAAGTGCATTGCGGTTTCCGTTGCTAGGTGGCGCTTAGCAACCCCTCTAGATACCAGTAGCTAAGGGGGAAGTGAAATCATAGCAACCTAGAGGAAG CTCATTTCATCAGCGCCGCCCAGAAGCCCGCCTTTTTTAAACCCGCCCATTTTGAACTTGGGTAAGATATG
HPyV10	JX262162	CATCCTTGCTGAATTTGCAAGTAGTAAAAAGTTTGCAGACGCGGTAAAGATGGCTCCCAGAGTCCTTCCTCTTTTCACCGGAAAGACAGAGGCTAGGCCTCCGGCCCCCG GCTTATATAGAAAAAATTTTAGCTTATTGTTTTGCTACTTAACCTCAGGTAGGTCAACAGCTATTGTTGGCAAGCTATTGTTGGCAAGTATTGGTATTAATCACCCAGACAA CTCAGAAGTTTCCACCTCTTGGAGGCGGTCCAGAGTTAACCTGTGACTGTTGGCGGGAAGCCAATAACAGCAACTTTGACATTTCATCACGAGCCCTTTAAACGCCCTCTA GGCGGAGACGGAAGACAATTGACTCTTGGCACGGACGGAAGAGGAATGCCGTCTGCTCACCTTAGTTAACACGAGATTTTCTTTGGAAATACTCCAAAGTACAGTAAGT ATATG
STLPyV	JX463183	CATTGCTGCAGAGTCAACTTCTGTCTCTAAAGAGGCTGAGGCTTAGGCCCCAGGCCCCGCATATATATGAGAAAAAAGGTAGCTGTGTGACTCACATTATTTATTGTCTA TTGCATAGTTTCATCACCTCAGTGATTCTTGGCAGTTTGCCACCTGGCTGAATCAGAGTGTGGAGCAGCTGCTGCGTGACCTGGAAAGCTGTCAAACCACGGAAAAGAGC GCCAAGGCAACCAATAGACTTTGGCTCTGTGCCCCTAGGCACACCTTACTTCCGGATATTATATAGACCATAATATTTAGTCAGCTGAGAGTAGTAACTTCGGCAGCGAGC CAACAACAGCATTATGTAAGTAAATG
HPyV12	JX308829	CATTCTGCTAAATTAGTGCAGTACAGTGTACGTGCCTGGTACCAAAAATGGCCTCCAAAAGCCTCTCCCCTTTTTTATAGGAAGGACGGGGGCTGGGTGCCTCTTGCCTCC CAGTAATAAAAAAAAAAGGAGGTGCCTAAGTACATGCCTGAGTGCGGGTTTAACTAGAGTTCATTATTGATTAATGTGTGTACTGGACATGCCTGTTCCCAGGCAGACGG CCAAGTTCCACAGGCAGACGGCCAAGTTCCTGGGCAGTCTGCCAAGTTTCTATTTTGGACACGCGCACTTCCTGTTCCTAACCGCGGTCAGAACAAGTTAGAACAACAACC AGCCCATAACCTCTTTTTTTTTTTAAGTAAAAAAAAAAGAAGAAAAAGAATCTTAACCCTCTGTCATCAAATCCTGAACCTTCTAGAAGATCTACTAAAGTAAGTCTTTTTTT TTTTTAGATG
NJPyV	KF954417	CATTTACCTGCCTAGAAGACCTTGCAGCCCGCAGCGAAGCAAGTAAGTGCAAGTGTGGGGCTAATTAGGTCTCTCTCCTTTTTATAAGATTGAGGTAGAGGCAAGAGGCC TCCTGCCTCACCACAAATTAAAAAAAAAACATCTCTATGTTTGAGCCCTCCTGACTGGTTGCCACCAAGTCCGGGAAGGCAGCCAAGGTCAGGGCTGCCTGCCAATGTGG TTAATTATTACACTTGGGAACCACAGGCGGCGGGTGGCTGCTCGCCAAACACAGCAGTAAATTGAAGGAAGTTGACTCTGTGTGATCATTGTGACCAGCCAGGCGAAACA GAGACACGCGACCAAACAAGAGAAAGCAAGTTACTGTTATTTCTTGTAAGTAAAACTTAATTTTTTAATTCTTTTTAGAATG

Table 1: NCCR expression in different cell lines derived from kidney, skin, lung, cervix, colon, brain and colon.

EVGR- and LVGR-expression in the transfected cell lines indicated.

Top row lists mean fluorescence intensity of BKPyV(ww) in the indicated cell lines, which was used for normalization to 1 for EVGR, and as 100 for LVGR levels.

EVGR

	Transfected cell lines							
	Kidney	Kidney	Kidney	Skin	Lung	Cervix	Brain	Colon
NCCR constructs	HEK293	HEK293T	HEK293TT	A375	A549	HeLa NT	HS683	SW480
BKPyVww-E (Ab. MFI)	*49.6 (10.7)*	*553.5 (87.4)*	*136.1 (12.1)*	*10.1 (6.0)*	*19.4 (7.3)*	*2.0 (0.5)*	*7.6 (3.6)*	*13.6 (1.1)*
BKPyVww-E	1 (0.2)	1.0 (0.4)	1.0 (0.2)	1.0 (0.7)	1.0 (0.4)	1.0 (0.3)	1.0 (0.5)	1.0 (0.1)
JCPyV-E	0.3 (0.2)	1.2 (0.2)	2.5 (0.7)	0.2 (0.1)	0.9 (0.3)	4.1 (1.6)	5.0 (4.8)	3.7 (0.4)
KIPyV-E	1.5 (0.9)	1.9 (0.5)	4.9 (2.6)	3.6 (0.6)	7.0 (1.4)	31.0 (5.0)	5.5 (3.1.0)	0.6 (0.02)
WUPyV-E	0.3 (0.1)	0.3 (0.1)	0.2 (0.1)	2.5 (1.0)	1.4 (0.7)	4.7 (0.7)	2.5 (1.6)	17.3 (12.3)
MCPyV-E	7.5 (3.2)	15.6 (3.5)	12.9 (2.4)	60.6 (15.1)	0.6 (0.1)	2.1 (0.7)	2.1 (0.5)	5.5 (2.4)
HPyV6-E	0.04 (0.01)	0.04 (0.02)	0.1 (0.1)	0.7 (0.4)	0.1 (0.03)	0.1 (0.1)	0.2 (0.2)	0.2 (0.1)
HPyV7-E	0.4 (0.4)	0.1 (0.04)	0.3 (0.3)	0.6 (0.4)	0.1 (0.1)	0.5 (0.6)	2.5 (1.2)	0.2 (0.1)
TSPyV-E	0.2 (0.1)	0.5 (0.1)	0.4 (0.5)	2.0 (0.4)	0.9 (0.2)	0.5 (0.1)	1.0 (0.7)	0.2 (0.1)
HPyV9-E	0.02 (0.01)	0.2 (0.2)	1.1 (0.4)	0.4 (0.1)	0.3 (0.1)	0.4 (0.2)	2.7 (1.7)	10.6 (2.3)
HPyV10-E	0.3 (0.3)	0.3 (0.2)	0.9 (0.5)	0.4 (0.1)	1.3 (0.2)	5.9 (0.1)	2.8 (0.3)	7.8 (4.8)
STLPyV-E	3.6 (3.2)	1.0 (0.2)	3.7 (2.2)	1.7 (0.3)	0.8 (0.2)	1.6 (0.02)	0.6 (0.1)	0.3 (0.1)
HPyV12-E	1.5 (0.9)	6.2 (1.0)	11.1 (7.9)	21.3 (7.2)	22.2 (8.2)	22.1 (4.5)	37.3 (19.0)	31.0 (12.8)
NJPyV-E	0.5 (0.5)	1.5 (0.5)	1.4 (0.6)	0.7 (0.3)	3.6 (2.5)	0.2 (0.2)	0.9 (0.5)	0.5 (0.2)

LVGR

NCCR constructs	HEK293	HEK293T	HEK293TT	A375	A549	HeLa NT	HS683	SW480
BKPyVww-L (Ab. MFI)	*2883.7 (494.6)*	*15950.3 (844)*	*22727 (1356.1)*	*3372.7 (784.0)*	*4005.2 (1166.1)*	*2049.0 (12.7)*	*4112.0 (37.2)*	*5202.8 (1343.8)*
BKPyVww-L	100.0 (17.1)	100.0 (5.4)	100.0 (6.7)	100.0 (23.0)	100.0 (29.1)	100.0 (0.6)	100.0 (0.9)	100.0 (25.0)
JCPyV-L	113.6 (27.1)	134.1 (11.3)	178.3 (36.8)	102.0 (12.2)	106.5 (9.6)	111.8 (26.0)	122.7 (32.5)	127.5 (10.6)
KIPyV-L	77.7 (26.9)	53.5 (4.6)	65.7 (5.6)	68.3 (3.6	72.2 (2.8)	173.2 (15.1)	83.2 (22.0)	83.1 (2.9)
WUPyV-L	62.8 (18.5)	44.9 (4.1)	26.7 (3.4)	64.3 (6.4)	97.3 (25.0)	188.03 (28.1)	104.9 (13.9)	95.7 (24.0)
MCPyV-L	46.3 (14.0)	32.3 (4.2)	38.4 (4.1)	40.3 (5.7)	44.1 (3.4)	74.0 (14.5)	44.8 (2.6)	55.7 (9.1)
HPyV6-L	50.6 (27.6)	24.5 (2.4)	29.0 (2.4)	13.8 (1.1)	29.0 (4.6)	79.8 (22.0)	32.9 (17.3)	47.1 (9.5)
HPyV7-L	47.3 (13.4)	16.2 (1.6)	19.2 (1.6)	13.2 (0.9)	22.7 (1.7)	44.0 (4.0)	27.2 (5.4)	34.5 (6.84)
TSPyV-L	113.2 (45.2)	63.7 (16.6)	78.8 (16.5)	83.9 (19.5)	99.3 (16.6)	159.0 (16.0)	120.0 (14.3)	144.0 (14.47)
HPyV9-L	79.4 (48.0)	61.2 (5.6)	74.2 (5.8)	23.8 (3.7)	21.8 (4.1)	70.6 (13.3)	32.3 (8.2)	41.4 (17.4)
HPyV10-L	69.4 (31.7)	51.9 (10.7)	61.8 (10.5)	23.3 (3.6)	54.3 (3.0)	51.6 (4.8)	62.5 (10.8)	50.4 (7.2)
STLPyV-L	49.0 (28.6)	34.0 (4.1)	39.7 (17.3)	12.1 (2.7)	21.1 (2.4)	33.1 (3.2)	23.8 (10.2)	48.3 (5.4)
HPyV12-L	96.2 (42.3)	34.1 (17.6)	44.0 (5.5)	74.0 (7.3)	137.2 (38.1)	150.0 (21.4)	102.6 (19.0)	100.9 (4.8)
NJPyV-L	90.0 (30.4)	43.6 (11.8)	55.2 (13.0)	19.8 (4.8)	78.7 (8.7)	71.8 (18.0)	86.8 (15.9)	66.0 (10.4)

418 **Table 2: Cell lines and culture conditions**
419
420

Cells transfected with NCCR reporter contructs

Cell lines	Culture media supplemented with (10% FBS, 2mM glutamine)	Transfection reagents	Origin	Cell lines source
HEK293	DMEM-H	Lipofectamin2000	Human embroynic kidney cells	ATCC® CRL-1573™
HEK293T	DMEM-H	Lipofectamine2000	Human embroynic kidney cells expressing SV40 LTag and sTag	ATCC® CRL-3216™
HEK293TT	DMEM-H plus 400µg/ml hygromycin B	Lipofectamin2000	Human embroynic kidney cells expressing SV40 LTag and sTag	NCI-Frederick Repository Service, Frederick, Maryland, USA
HEK293MCT	DMEM-H	Lipofectamine2000	Human embroynic kidney cells expressing MCPyV LTag and sTag	Gift from Prof. Nicole Fischer from Institute for Microbiology and Virology, University Medical Center Eppendorf, Hamburg, Germany
A375	DMEM-H	Lipofectamin2000	Human malignant melanoma cells	ATCC® CRL-1619™
A549	EMEM	Lipofectamine3000	Human epithelial lung adenocarcinoma	ATCC-CCL-185
HeLa NT	DMEM-H	Lipofectamin2000	Human cervical carcinoma cells	Gift from Natalia Teterina from Laboratory of Viral Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland, USA
HS683	DMEM-H	Omni avalanche	Human brain glioma cells	ATCC® HTB-138™
SW480	EMEM	Lipofectamin2000	Human colon adenocarcinoma cells	ATCC-CCL-228

FIGURE LEGENDS

Fig. 1: *In silico* prediction of common transcription factor- and LTag-binding sites in the NCCR of 13 HPyVs. The BKPyVww NCCR was used as the reference as described in Materials & Methods (45). HPyV-NCCR length is shown in the brackets in base pairs (bp). The direction of EVGR and LVGR transcription and the start codon (ATG) is depicted in red and green arrows, respectively. TSS, transcription start sites in dashed arrows, inr, initiation element in shaded grey rectangle; DPE, downstream promoter element in blue shaded rectangle; binding sites are shown as red triangle, LTag; yellow triangles, Sp1; green triangles, NFkB; black square, TATA-box; grey square, TATA-like element; blue triangles, NF1; orange triangles, Ets1; light orange triangles, Spi-B; ori, origins of replication in pink rectangles.

Fig. 2: Rearranged BKPyV Dunlop NCCR showed higher EVGR-expression compared to BKPyVww NCCR in HEK293 cells.

(A) Schematic representation of HPyV genome showing the noncoding control region (NCCR), early viral gene region (EVGR) in red encoding large and small T antigens (Tags), alternatives splice Tags; micro RNAs in blue arrow; late viral gene region (LVGR) encoding structural proteins (Vp1, Vp2, and Vp3), and the agnoprotein (agno) only in BKPyV and JCPyV.

(B) Representation of the bidirectional reporter vector, pRG13D12 containing the NCCR (in grey) in the early to late orientation cloned via restriction sites MluI and BssHII; red fluorescence protein, dsRed2 as a marker of EVGR-expression; green fluorescence protein, EGFP in opposite orientation, as a marker of LVGR-expression; SV40 polyadenylation signals (SV40 polyA) for the dsRed2 and EGFP expression cassette; E1 ori for bacterial plasmid replication; ampicillin resistant gene (Amp) for selecting *E.coli* transformants.

(C) Flow cytometry of HEK293 cells 2 dpt with pRG13D12 reporter vector containing the archetype BKPyVww-NCCR or the BKPyV(DUN)-NCCR-Fw cloned in forward orientation or BKPyV(DUN)-

446 NCCR-Rev cloned in reversed orientation. x-axis, EGFP fluorescence, y-axis, dsRed2
447 fluorescence, 10,000 control transfected cells were gated for the live gate while 5,000 transfected
448 cells were gated for the P3 (Q_1, Q_2 and Q_4) gate. Q_1, Q_4 and Q_2 depict cells expressing red, green
449 and both respectively. Ex, excitation wavelength; Em, emission wavelength.
450 (D) Quantification of cells: red bars, sum of red cells (Q_1+Q_2); green bars, sum of green cells
451 (Q_2+Q_4); yellow bars red- and green-fluorescence double-positive cells (Q_2); grey bars non-
452 fluorescence cells (Q3, negative). Mean with SD of three independent replicates shown.
453 (E) Normalized mean fluorescence intensity (MFI). The weighted MFI was calculated for each
454 measurement (see formulas in Materials and Methods), late expression was normalized to
455 BKPyVww NCCR (green MFI set as 100) while early expression normalized to BKPyVww NCCR
456 (red MFI set as 1). Mean with SD of three independent replicates shown.
457
458 **Fig. 3: HPyV-NCCR differ in EVGR-expression in HEK293 and 293T cell lines.**
459 (A) HEK293 normalized to archetype BKPyV(ww); (B) HEK293T cells normalized to archetype
460 BKPyV(ww); (C) Fold change in 293T versus HEK293 cells; error bars (SD of three independent
461 replicates); (D) Log-linear regression (y-log, x-linear) of LTag-binding sites and EVGR-expression
462 in 293T cells.
463
464 **Fig.4. MCPyV-NCCR response in MVPyV T-antigen expressing 293MCT.**
465 (A) MCPyV-LTag expression in 293MCT cells
466 (B) Flow cytometry of MCPyV-17b and MCPyV-MCVw156;
467 (C) EVGR-expression presented as fold change in HEK293 cells;
468 (D) LVGR-expression expressed as fold change in HEK293 cells; mean with SD of three
469 independent replicates shown; Wilcoxon-t-test.
470
471 **Fig. 5 Clinical NCCR rearrangements of HPyV7 and HPyV9**

(A) HPyV7-PITT1 carries a 16 bp insertion (ins); HPyV7-PITT2 carries a 12 bp deletion (del, dotted line) compared to the archetype HPyV7-NCCR shown in Fig.1 EVGR- and LVGR-expression shown as fold change in HEK293 or SW480 cells, respectively. Mean with SD of three independent replicates shown; Wilcoxon t-test.

(B) HPyV9-UF1 carries an insertion providing 3 additional Sp1 sites compared to archetype HPyV7 shown in Fig 1. EVGR- and LVGR-expression shown as fold change in HEK293 or A549 cells, respectively. Mean with SD of three independent replicates shown; Wilcoxon t-test.

Figure1

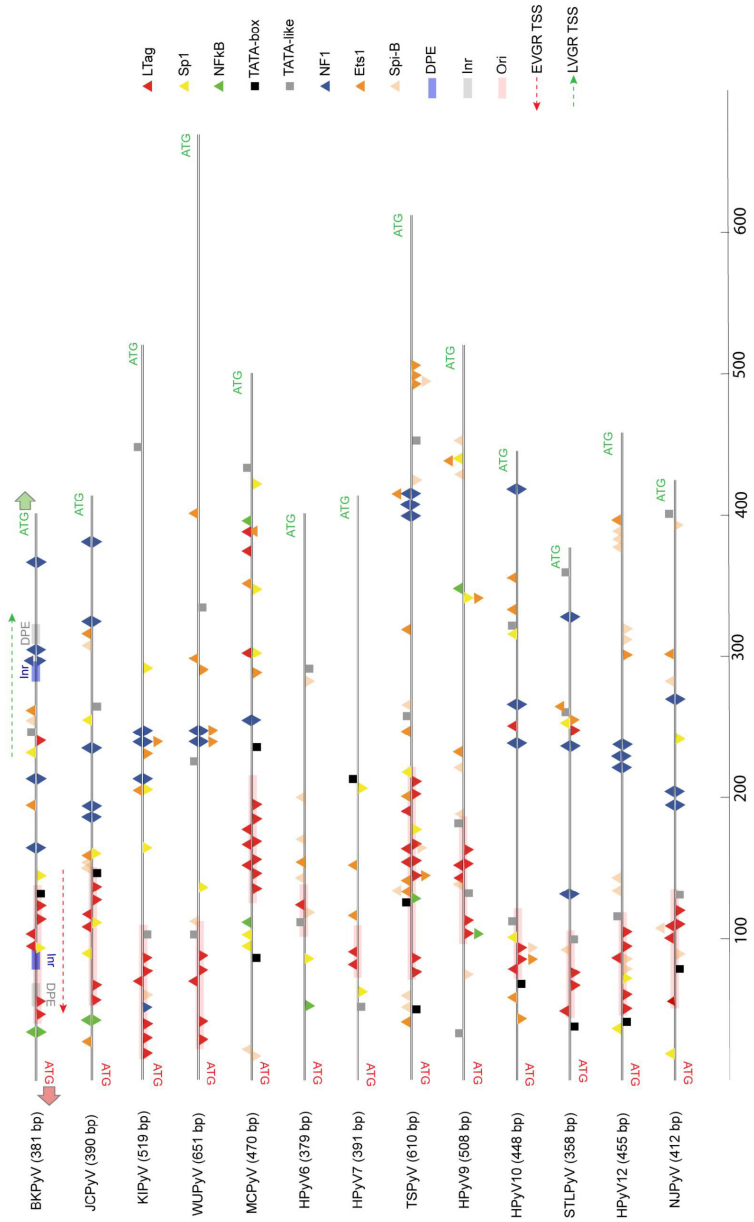
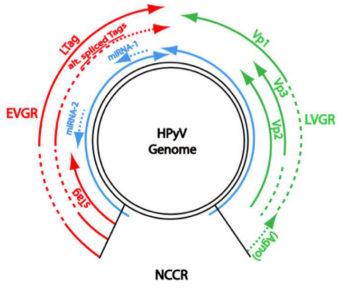
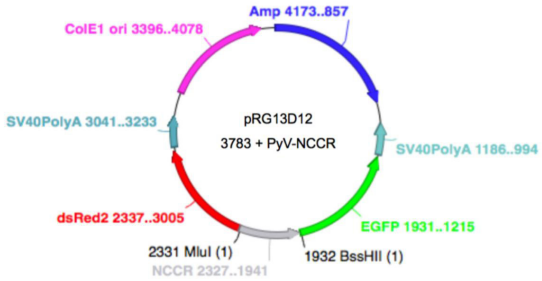

LTag
Sp1
NFkB
TATA-box
TATA-like
NF1
Ets1
Spi-B
DPE
Inr
Ori
EVGR TSS
LVGR TSS
ATG
Inr
DPE
100
200
300
400
500
600
BKPyV (381 bp)
JCPyV (390 bp)
KIPyV (519 bp)
WUPyV (651 bp)
MCPyV (470 bp)
HPyV6 (379 bp)
HPyV7 (391 bp)
TSPyV (610 bp)
HPyV9 (508 bp)
HPyV10 (448 bp)
STLPyV (358 bp)
HPyV12 (455 bp)
NJPyV (412 bp)

Figure 2

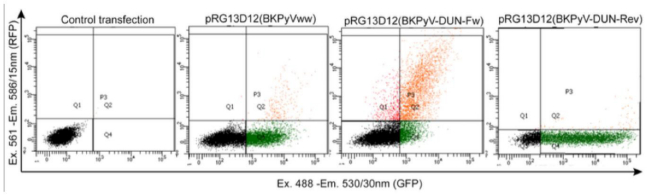
A


LTag
alt. spliced Tags
miRNA-1
miRNA-2
EVGR
RGTP
HPyV
Genome
Vp1
Vp3
Vp2
LVGR
(agnoprotein)
NCCR

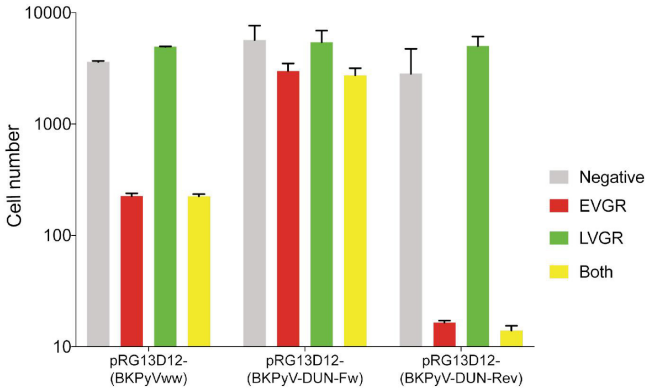
B


Amp 4173..857
ColE1 ori 3396..4078
SV40PolyA 3041..3233
pRG13D12
3783 + PyV-NCCR
SV40PolyA 1186..994
dsRed2 2337..3005
EGFP 1931..1215
2331 MluI (1)
1932 BssHII (1)
NCCR 2327..1941

529 C


Control transfection
pRG13D12(BKPyVww)
pRG13D12(BKPyV-DUN-Fw)
pRG13D12(BKPyV-DUN-Rev)
Ex. 561 -Em. 586/15nm (RFP)
Ex. 488 -Em. 530/30nm (GFP)

D


10000
1000
100
10
Cell number
pRG13D12-
(BKPyVww)
pRG13D12-
(BKPyV-DUN-Fw)
pRG13D12-
(BKPyV-DUN-Rev)
Negative
EVGR
LVGR
Both

E

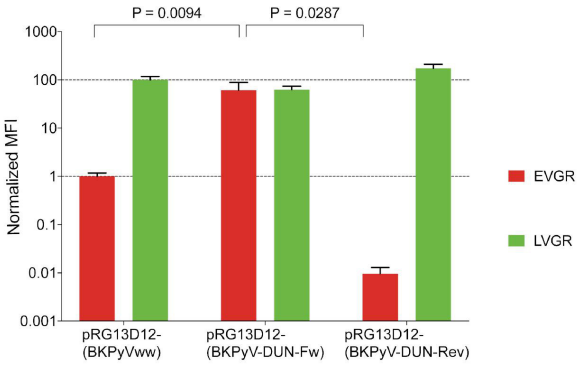
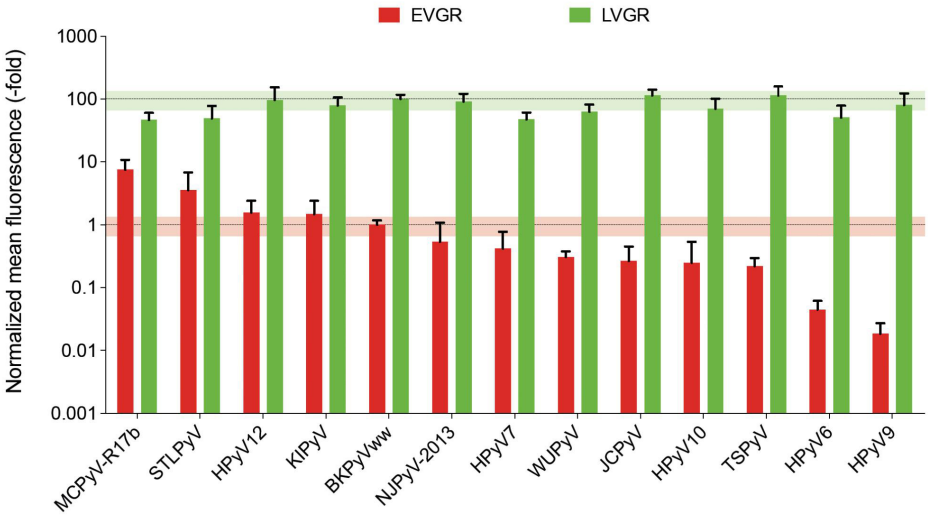
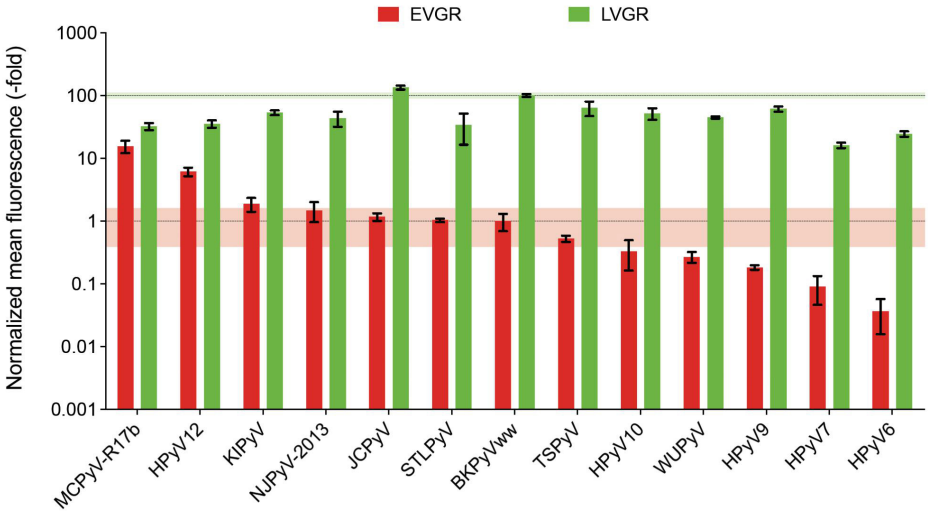

P = 0.0094
P = 0.0287
1000
100
10
1
0.1
0.01
0.001
Normalized MFI
pRG13D12-
(BKPyVww)
pRG13D12-
(BKPyV-DUN-Fw)
pRG13D12-
(BKPyV-DUN-Rev)
EVGR
LVGR

Figure 3

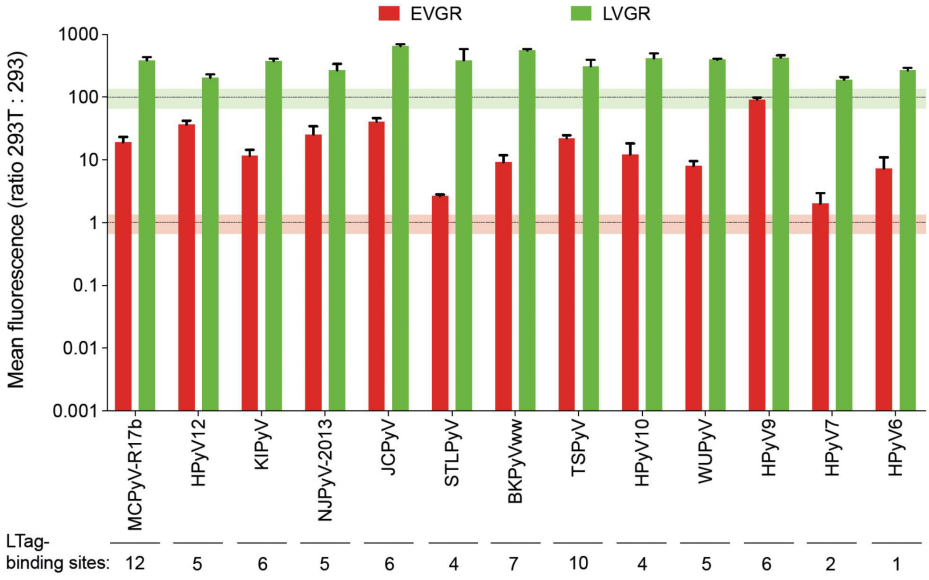
A


EVGR
LVGR
Normalized mean fluorescence (-fold)
1000
100
10
1
0.1
0.01
0.001
MCPyV-R17b
STLPyV
HPyV12
KIPyV
BKPyVww
NJPyV-2013
HPyV7
WUPyV
JCPyV
HPyV10
TSPyV
HPyV6
HPyV9

B


EVGR
LVGR
Normalized mean fluorescence (-fold)
1000
100
10
1
0.1
0.01
0.001
MCPyV-R17b
HPyV12
KIPyV
NJPyV-2013
JCPyV
STLPyV
BKPyVww
TSPyV
HPyV10
WUPyV
HPyV9
HPyV7
HPyV6

C


EVGR
LVGR
Mean fluorescence (ratio 293T : 293)
1000
100
10
1
0.1
0.01
0.001
MCPyV-R17b
HPyV12
KIPyV
NJPyV-2013
JCPyV
STLPyV
BKPyVww
TSPyV
HPyV10
WUPyV
HPyV9
HPyV7
HPyV6
LTag-binding sites:
12
5
6
5
6
4
7
10
4
5
6
2
1

D

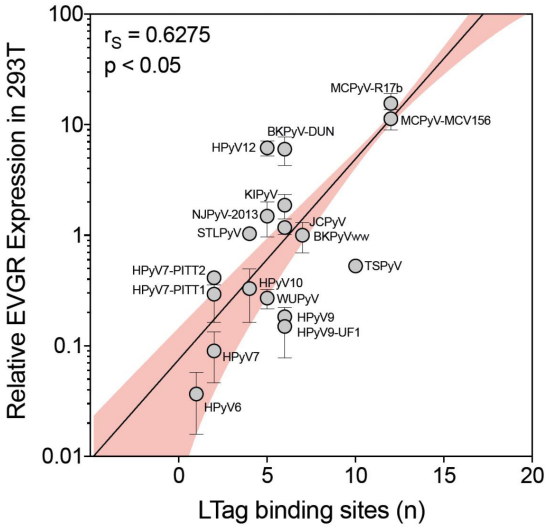
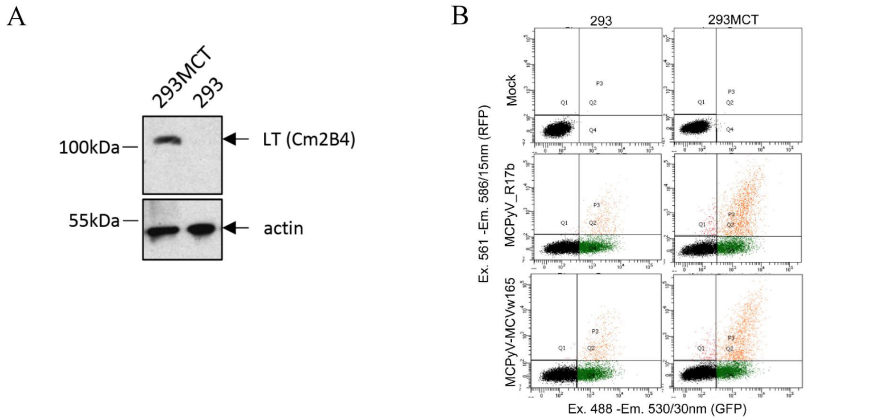
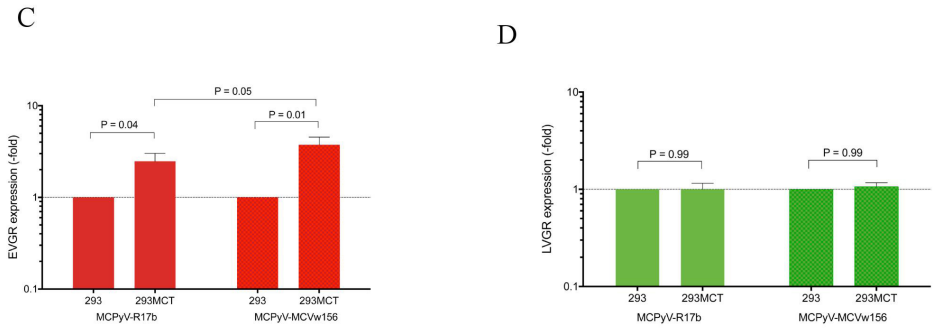

r_S = 0.6275
p < 0.05
Relative EVGR Expression in 293T
LTag binding sites (n)
100
10
1
0.1
0.01
0
5
10
15
20
MCPyV-R17b
MCPyV-MCV156
BKPyV-DUN
HPyV12
KIPyV
NJPyV-2013
JCPyV
STLPyV
BKPyVww
TSPyV
HPyV7-PITT2
HPyV7-PITT1
HPyV10
WUPyV
HPyV9
HPyV9-UF1
HPyV7
HPyV6

Figure 4


A
293MCT
293
100kDa
LT (Cm2B4)
55kDa
actin
B
293
293MCT
Mock
MCPyV_R17b
MCPyV-MCVw165
Ex. 561 -Em. 586/15nm (RFP)
Ex. 488 -Em. 530/30nm (GFP)

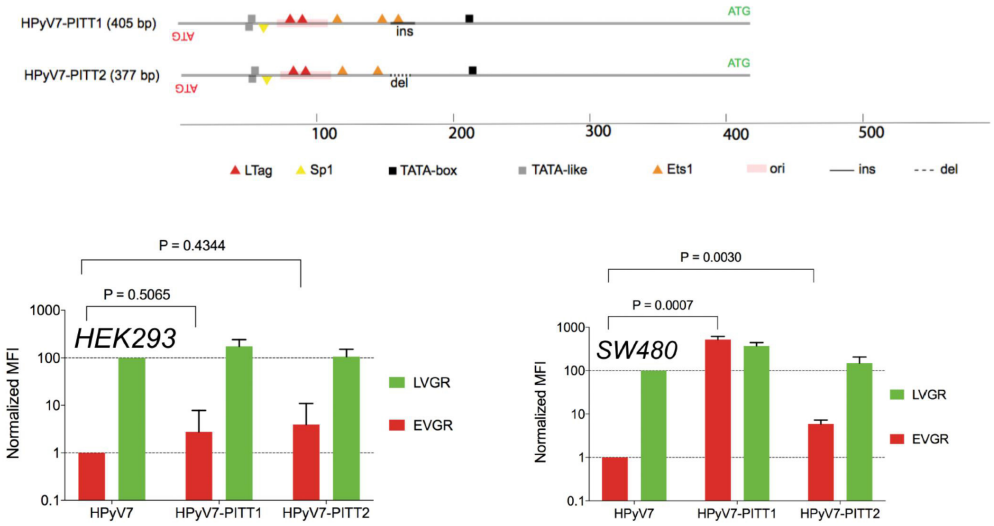

C
P = 0.05
P = 0.04
P = 0.01
EVGR expression (-fold)
10
1
0.1
293
293MCT
293
293MCT
MCPyV-R17b
MCPyV-MCVw156
D
P = 0.99
P = 0.99
LVGR expression (-fold)
10
1
0.1
293
293MCT
293
293MCT
MCPyV-R17b
MCPyV-MCVw156

549 Figure 5

550

551 A HPyV7

552

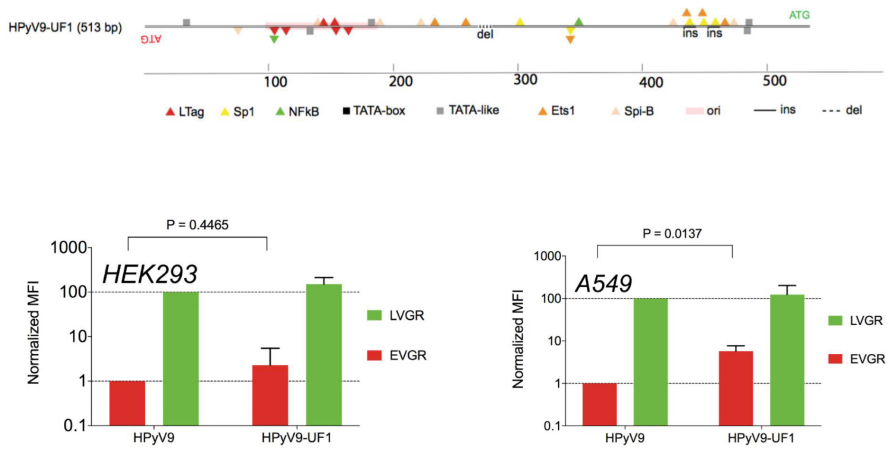


553

554

555 B HPyV9

556



557

1. DeCaprio JA, Imperiale MJ, Major EO. 2013. Polyomaviruses, p 1633-1661. *In* Knipe DM, Howley P (ed), Fields Virology, 6th Edition, 6th Edition ed.
2. Polyomaviridae Study Group of the International Committee on Taxonomy of V, Calvignac-Spencer S, Feltkamp MC, Daugherty MD, Moens U, Ramqvist T, Johne R, Ehlers B. 2016. A taxonomy update for the family Polyomaviridae. Arch Virol 161:1739-50.
3. Greenlee JE, Hirsch HH. 2017. Polyomaviruses, p 599-623. *In* Microbiology AASf (ed), Clinical Virology, 4th Edition, 4th Edition ed doi:Book DOI: 10.1128/9781555819439, asmscience.org.
4. Gardner SD, Field AM, Coleman DV, Hulme B. 1971. New human papovavirus (B.K.) isolated from urine after renal transplantation. Lancet 1:1253-57.
5. Padgett BL, Walker DL, ZuRhein GM, Eckroade RJ, Dessel BH. 1971. Cultivation of papova-like virus from human brain with progressive multifocal leucoencephalopathy. Lancet 1:1257-60.
6. Bialasiewicz S, Whiley DM, Lambert SB, Wang D, Nissen MD, Sloots TP. 2007. A newly reported human polyomavirus, KI virus, is present in the respiratory tract of Australian children. J Clin Virol 40:15-8.
7. Gaynor AM, Nissen MD, Whiley DM, Mackay IM, Lambert SB, Wu G, Brennan DC, Storch GA, Sloots TP, Wang D. 2007. Identification of a novel polyomavirus from patients with acute respiratory tract infections. PLoS Pathog 3:e64.
8. Feng H, Shuda M, Chang Y, Moore PS. 2008. Clonal integration of a polyomavirus in human Merkel cell carcinoma. Science 319:1096-100.
9. Schowalter RM, Pastrana DV, Pumphrey KA, Moyer AL, Buck CB. 2010. Merkel cell polyomavirus and two previously unknown polyomaviruses are chronically shed from human skin. Cell Host Microbe 7:509-15.

584 10. van der Meijden E, Janssens RW, Lauber C, Bouwes Bavinck JN, Gorbalenya AE,
585 Feltkamp MC. 2010. Discovery of a new human polyomavirus associated with
586 trichodysplasia spinulosa in an immunocompromized patient. PLoS Pathog 6:e1001024.
587 11. Scuda N, Hofmann J, Calvignac-Spencer S, Ruprecht K, Liman P, Kuhn J, Hengel H,
588 Ehlers B. 2011. A novel human polyomavirus closely related to the african green
589 monkey-derived lymphotropic polyomavirus. J Virol 85:4586-90.
590 12. Siebrasse EA, Reyes A, Lim ES, Zhao G, Mkakosya RS, Manary MJ, Gordon JI, Wang
591 D. 2012. Identification of MW polyomavirus, a novel polyomavirus in human stool. J Virol
592 86:10321-6.
593 13. Lim ES, Reyes A, Antonio M, Saha D, Ikumapayi UN, Adeyemi M, Stine OC, Skelton R,
594 Brennan DC, Mkakosya RS, Manary MJ, Gordon JI, Wang D. 2013. Discovery of STL
595 polyomavirus, a polyomavirus of ancestral recombinant origin that encodes a unique T
596 antigen by alternative splicing. Virology 436:295-303.
597 14. Korup S, Rietscher J, Calvignac-Spencer S, Trusch F, Hofmann J, Moens U, Sauer I,
598 Voigt S, Schmuck R, Ehlers B. 2013. Identification of a novel human polyomavirus in
599 organs of the gastrointestinal tract. PLoS One 8:e58021.
600 15. Mishra N, Pereira M, Rhodes RH, An P, Pipas JM, Jain K, Kapoor A, Briese T, Faust PL,
601 Lipkin WI. 2014. Identification of a novel polyomavirus in a pancreatic transplant recipient
602 with retinal blindness and vasculitic myopathy. J Infect Dis 210:1595-9.
603 16. Gheit T, Dutta S, Oliver J, Robitaille A, Hampras S, Combes JD, McKay-Chopin S, Le
604 Calvez-Kelm F, Fenske N, Cherpelis B, Giuliano AR, Franceschi S, McKay J, Rollison
605 DE, Tommasino M. 2017. Isolation and characterization of a novel putative human
606 polyomavirus. Virology 506:45-54.
607 17. Kean JM, Rao S, Wang M, Garcea RL. 2009. Seroepidemiology of human
608 polyomaviruses. PLoS Pathog 5:e1000363.

609 18. Kardas P, Leboeuf C, Hirsch HH. 2015. Optimizing JC and BK polyomavirus IgG testing
610 for seroepidemiology and patient counseling. J Clin Virol 71:28-33.
611 19. Gossai A, Waterboer T, Nelson HH, Michel A, Willhauck-Fleckenstein M, Farzan SF,
612 Hoen AG, Christensen BC, Kelsey KT, Marsit CJ, Pawlita M, Karagas MR. 2016.
613 Seroepidemiology of Human Polyomaviruses in a US Population. Am J Epidemiol
614 183:61-9.
615 20. Hirsch HH, Knowles W, Dickenmann M, Passweg J, Klimkait T, Mihatsch MJ, Steiger J.
616 2002. Prospective study of polyomavirus type BK replication and nephropathy in renal-
617 transplant recipients. N Engl J Med 347:488-496.
618 21. Hirsch HH, Kardas P, Kranz D, Leboeuf C. 2013. The human JC polyomavirus (JCPyV):
619 virological background and clinical implications. APMIS 121:685–727.
620 22. Ho J, Jedrych JJ, Feng H, Natalie AA, Grandinetti L, Mirvish E, Crespo MM, Yadav D,
621 Fasanella KE, Proksell S, Kuan SF, Pastrana DV, Buck CB, Shuda Y, Moore PS, Chang
622 Y. 2015. Human polyomavirus 7-associated pruritic rash and viremia in transplant
623 recipients. J Infect Dis 211:1560-5.
624 23. Hirsch HH, Babel N, Comoli P, Friman V, Ginevri F, Jardine A, Lautenschlager I,
625 Legendre C, Midtvedt K, Munoz P, Randhawa P, Rinaldo CH, Wieszek A, Hosts
626 ESGoliC. 2014. European perspective on human polyomavirus infection, replication and
627 disease in solid organ transplantation. Clin Microbiol Infect 20 Suppl 7:74-88.
628 24. Siebrasse EA, Nguyen NL, Smith C, Simmonds P, Wang D. 2014. Immunohistochemical
629 detection of KI polyomavirus in lung and spleen. Virology 468-470C:178-184.
630 25. Siebrasse EA, Nguyen NL, Willby MJ, Erdman DD, Menegus MA, Wang D. 2016.
631 Multiorgan WU Polyomavirus Infection in Bone Marrow Transplant Recipient. Emerg
632 Infect Dis 22:24-31.

633 26. Siebrasse EA, Pastrana DV, Nguyen NL, Wang A, Roth MJ, Holland SM, Freeman AF,
634 McDyer J, Buck CB, Wang D. 2015. WU polyomavirus in respiratory epithelial cells from
635 lung transplant patient with Job syndrome. Emerg Infect Dis 21:103-6.
636 27. Schrama D, Groesser L, Ugurel S, Hafner C, Pastrana DV, Buck CB, Cerroni L, Theiler
637 A, Becker JC. 2014. Presence of human polyomavirus 6 in mutation-specific BRAF
638 inhibitor-induced epithelial proliferations. JAMA Dermatol 150:1180-6.
639 28. Nguyen KD, Lee EE, Yue Y, Stork J, Pock L, North JP, Vandergriff T, Cockerell C, Hosler
640 GA, Pastrana DV, Buck CB, Wang RC. 2017. Human polyomavirus 6 and 7 are
641 associated with pruritic and dyskeratotic dermatoses. J Am Acad Dermatol 76:932-940
642 e3.
643 29. DeCaprio JA, Garcea RL. 2013. A cornucopia of human polyomaviruses. Nat Rev
644 Microbiol 11:264-76.
645 30. Gosert R, Rinaldo CH, Funk GA, Egli A, Ramos E, Drachenberg CB, Hirsch HH. 2008.
646 Polyomavirus BK with rearranged noncoding control region emerge in vivo in renal
647 transplant patients and increase viral replication and cytopathology. J Exp Med 205:841-
648 52.
649 31. Ferenczy MW, Marshall LJ, Nelson CD, Atwood WJ, Nath A, Khalili K, Major EO. 2012.
650 Molecular Biology, Epidemiology, and Pathogenesis of Progressive Multifocal
651 Leukoencephalopathy, the JC Virus-Induced Demyelinating Disease of the Human Brain.
652 Clin Microbiol Rev 25:471-506.
653 32. Olsen GH, Hirsch HH, Rinaldo CH. 2009. Functional analysis of polyomavirus BK non-
654 coding control region quasispecies from kidney transplant recipients. J Med Virol
655 81:1959-67.
656 33. Gosert R, Kardas P, Major EO, Hirsch HH. 2010. Rearranged JC virus noncoding control
657 regions found in progressive multifocal leukoencephalopathy patient samples increase
658 virus early gene expression and replication rate. J Virol 84:10448-56.

659 34. Pastrana DV, Ray U, Magaldi TG, Schowalter RM, Cuburu N, Buck CB. 2016. Erratum
660 for Pastrana et al., BK Polyomavirus Genotypes Represent Distinct Serotypes with
661 Distinct Entry Tropism. J Virol 90:624.
662 35. Khan ZM, Liu Y, Neu U, Gilbert M, Ehlers B, Feizi T, Stehle T. 2014. Crystallographic
663 and glycan microarray analysis of human polyomavirus 9 VP1 identifies N-glycolyl
664 neuraminic acid as a receptor candidate. J Virol 88:6100-11.
665 36. Stroh LJ, Neu U, Blaum BS, Buch MH, Garcea RL, Stehle T. 2014. Structure analysis of
666 the major capsid proteins of human polyomaviruses 6 and 7 reveals an obstructed sialic
667 Acid binding site. J Virol 88:10831-9.
668 37. Imperiale MJ, Jiang M. 2016. Polyomavirus Persistence. Annu Rev Virol 3:517-532.
669 38. Bollag B, Chuke WF, Frisque RJ. 1989. Hybrid genomes of the polyomaviruses JC virus,
670 BK virus, and simian virus 40: identification of sequences important for efficient
671 transformation. J Virol 63:863-72.
672 39. Chen BJ, Atwood WJ. 2002. Construction of a novel JCV/SV40 hybrid virus (JCSV)
673 reveals a role for the JCV capsid in viral tropism. Virology 300:282-90.
674 40. Major EO, Miller AE, Mourrain P, Traub RG, de Widt E, Sever J. 1985. Establishment of
675 a line of human fetal glial cells that supports JC virus multiplication. Proc Natl Acad Sci U
676 S A 82:1257-61.
677 41. Tang WJ, Folk WR. 1989. Constitutive expression of simian virus 40 large T antigen in
678 monkey cells activates their capacity to support polyomavirus replication. J Virol 63:5478-
679 82.
680 42. Henriksen S, Tylden GD, Dumoulin A, Sharma BN, Hirsch HH, Rinaldo CH. 2014. The
681 human fetal glial cell line SVG p12 contains infectious BK polyomavirus. J Virol 88:7556-
682 68.
683 43. Grundhoff A, Fischer N. 2015. Merkel cell polyomavirus, a highly prevalent virus with
684 tumorigenic potential. Curr Opin Virol 14:129-37.

685 44. Bethge T, Hachemi HA, Manzetti J, Gosert R, Schaffner W, Hirsch HH. 2015. Sp1 sites
686 in the noncoding control region of BK polyomavirus are key regulators of bidirectional
687 viral early and late gene expression. J Virol 89:3396-411.
688 45. Bethge T, Ajuh E, Hirsch HH. 2016. Imperfect Symmetry of Sp1 and Core Promoter
689 Sequences Regulates Early and Late Virus Gene Expression of the Bidirectional BK
690 Polyomavirus Noncoding Control Region. J Virol 90:10083-10101.
691 46. Cartharius K, Frech K, Grote K, Klocke B, Haltmeier M, Klingenhoff A, Frisch M,
692 Bayerlein M, Werner T. 2005. MatInspector and beyond: promoter analysis based on
693 transcription factor binding sites. Bioinformatics 21:2933-42.
694 47. Ehlers B, Moens U. 2014. Genome analysis of non-human primate polyomaviruses.
695 Infect Genet Evol 26:283-94.
696 48. Neumann F, Borchert S, Schmidt C, Reimer R, Hohenberg H, Fischer N, Grundhoff A.
697 2011. Replication, Gene Expression and Particle Production by a Consensus Merkel Cell
698 Polyomavirus (MCPyV) Genome. PLoS One 6:e29112.
699 49. Martel-Jantin C, Filippone C, Cassar O, Peter M, Tomasic G, Vielh P, Briere J, Petrella T,
700 Aubriot-Lorton MH, Mortier L, Jouvion G, Sastre-Garau X, Robert C, Gessain A. 2012.
701 Genetic variability and integration of Merkel cell polyomavirus in Merkel cell carcinoma.
702 Virology 426:134-42.
703 50. Lednicky JA, Butel JS, Luetke MC, Loeb JC. 2014. Complete genomic sequence of a
704 new Human polyomavirus 9 strain with an altered noncoding control region. Virus Genes
705 49:490-2.
706 51. Pipas JM. 1992. Common and unique features of T antigens encoded by the
707 polyomavirus group. J Virol 66:3979-85.
708 52. Lynch KJ, Frisque RJ. 1990. Identification of critical elements within the JC virus DNA
709 replication origin. J Virol 64:5812-22.

710 53. Moens U, Van Ghelue M, Ludvigsen M, Korup-Schulz S, Ehlers B. 2015. Early and late
711 promoters of BK polyomavirus, Merkel cell polyomavirus, Trichodysplasia spinulosa-
712 associated polyomavirus and human polyomavirus 12 are among the strongest of all
713 known human polyomaviruses in 10 different cell lines. J Gen Virol 96:2293-303.
714 54. Rinaldo CH, Hirsch HH. 2013. The human polyomaviruses: from orphans and mutants to
715 patchwork family. APMIS 121:681–4.
716 55. Ehlers B, Wieland U. 2013. The novel human polyomaviruses HPyV6, 7, 9 and beyond.
717 APMIS 121:783–95.
718

The data demonstrated that the NCCR of HPyVs is different in length, number and position of LTag and TFBS composition, which drive the differential bi-directional expression. This disparity in the NCCRs-driven expression in different cell lines is also due to the transcription factors (TFs) make-up in the cell lines. Some HPyVs presented with high EVGR or LVGR expression in cell lines derived from tissues, where they were reported to have tropism for or detected in. For instance, BKPyV and JCPyV known to have tropism for kidney and brain had highest EVGR in the kidney derived cells HEK293 or the brain derived cells HS683, respectively. MCPyV, suspected to have tropism for the skin, displayed its highest EVGR expression in the skin-derived A375 cells. While HPyV12 discovered from the liver of a colon carcinoma patient revealed a higher EVGR expression in the colon derived SW480 cells. However, others showed high activity in cell lines derived from tissues where they have not so far been detected. This indicates that these HPyVs could have tropism for the body compartments where the cell lines were derived from or may be replicated in, however, this remains to be tested. BKPyV, JCPyV and MCPyV have been efficiently replicated in cell lines expressing SV40 or MCPyV Tags, respectively (305). SV40 Tags activated EVGR expression of HPyVs and MCPyV cognate Tags activated EVGR expression from the MCPyV NCCRs, indicating cell lines expressing cognate Tags of respective HPyVs could be used for their propagation. Analyses of rearranged NCCR variants for BKPyV and JCPyV have been greatly investigated (25, 26). We reported that rearranged NCCR variants of novel HPyVs (HPyV7 and HPyV9) derived from patients increased EVGR as reported for BKPyV and JCPyV (25, 26) reiterating the NCCR as the pathogenicity determinant. Interestingly, even the rearranged NCCR variants of novel HPyVs demonstrated varied activity in different cell lines, indicating the host cell TFs make-up are also involved in the NCCR-driven viral expression. Our data are important for identifying host and viral factors necessary for culturing HPyVs and prediction of HPyVs secondary host cell tropism.

This work was performed together with Christian Baumann.

5.3 Point mutations in the large T antigen binding sites of BK polyomavirus non-coding control region affect gene expression

ABSTRACT

The non-coding control region of all human polyomaviruses contain the origin of viral replication, which is made up of binding motifs for large T antigen (LTag). LTag interacts with these motifs to mediate viral genome replication and activation of late viral gene region (LVGR) expression as already described for the SV40. Since the LTag binding motifs between SV40 and BK polyomavirus (BKPyV) at the ori are the same, we hypothesised that the effect of LTag-binding sites mutageneses in BKPyV NCCR would be similar to that of SV40. We compared the EVGR and LVGR expression of BKPyV archetype (ww) NCCR defective in one or more LTag binding sites using a bidirectional reporter vector in HEK293 cell lines expressing no, small or large amounts of SV40 LTag. Our results demonstrated an increase in early viral gene region (EVGR) expression in all mutants in both HEK293T and HEK293TT cells compared to HEK293 cells. The EVGR expression increase was less in HEK293TT cells compared to HEK293T cells, however, mutants containing mutant A showed a lesser decrease of EVGR expression in HEK293TT cells compared to all other mutants without mutant A. Furthermore, mutants harboring mutant A displayed higher levels of EVGR expression compared to other mutants in both HEK293T and HEK293TT cells. Conversely, LVGR expression increase proportionally in HEK293T and HEK293TT cells; respectively. Nonetheless, mutants harboring A and D mutations, abrogated the further increase of LVGR in HEK293TT cells.

INTRODUCTION

Polyomaviruses (PyVs) belong to the family *Polyomavirdae*, they are double-stranded DNA (dsDNA) viruses with small circular genomes of approximately 5 kb (328). The genome is encapsidated in an icosahedral capsid of approximately 45 nm (1). The genome of PyVs is divided into 3 main regions: The early viral gene region (EVGR), late viral gene region (LVGR) and the non-coding control region (NCCR) (329). The EVGR encodes small and large T antigens (sTag and LTag), while the LVGR encodes viral structural proteins. The LTag possesses multifunctional roles, such as initiation and maintenance of viral genome replication, regulation of EVGR and LVGR expression, and assembly of virions (59, 330-332). Additionally, LTag leads to transformation by interacting with tumour suppressor proteins such as retinoblastoma proteins (pRB) and p53 (333). PyVs' sTag has been reported to be implicated in the transformation role of LTag (18, 260, 264). The NCCR of HPyVs of approximate 400-600 base pairs (bp) in length is located between the EVGR and the LVGR. The NCCR harbours LTag binding motifs mostly at the origin of viral replication (ori), a numerous transcription factor binding sites (TFBSs), enhancers, EVGR and LVGR promoters, which initiate and regulate the bidirectional expression of PyVs. The NCCR of BK polyomavirus (BKPyV) is approximately 400 bp. The archetype (ww) BKPyV NCCR is made up of linear sequence blocks arbitrarily denoted as O_{146} (containing the ori), P_{68}, Q_{39}, R_{63}, and S_{63} (subscript numbers represent number of nucleotides). Two kinds of BKPyV NCCR architectures have been identified, ww- and rearranged (rr-) NCCRs. The rr-NCCRs consist of sequence deletions, duplications, or combinations of both. The rearrangement frequently occurs in the LVGR promoter side, while the EVGR promoter side containing the ori remain intact (22). The rr-NCCRs of BKPyV and JCPyV variants from patients suffering from BKPyV-associated nephropathy (BKPyVAN) and progressive multifocal leukoencephalopathy (PML), have been demonstrated to be associated with increased EVGR expression and viral replication (25, 26). The ori of PyVs is located in the NCCR, proximal to the EVGR side (49). The role of each LTag binding sites within the ori of HPyVs have not been well characterized, unlike in the monkey polyomavirus SV40 (59). The SV40 ori is consist of pentanucleotide sequences (5'-GRGGC-3' or 5'-GCCYC-3'), which are found at the oris. LTag interacts with these

pentanucleotide sequences via the origin-binding domain (OBD). This OBD is about 50% identical among HPyVs and can activate transcription (334). Furthermore, this protein melts and unwind the dsDNA in an ATP-dependent manner. Additionally, other host's factors such as topoisomerase I, polymerase alpha primase complex and replication factor A are recruited by LTag to the site of replication initiation (335, 336). Additional functional domain of the LTag, such as the DnaJ domain plays a vital role in viral genome replication *in vivo* (337), however, dispensable for viral replication *in vitro* (338). The SV40 ori is made up of two LTag binding motifs proximal to the EVGR with the ori referred to as *site I*, and four centrally located LTag binding sites referred to as *site II* (59). The SV40 minimal ori is characterised by three sequence motifs that is, a centrally located four-pentanucleotide repeats (*site II*), a poorly conserved early palindrome (EP) sequence proximal to the early side of the ori, and a conserved AT-rich sequence near the late side of the ori (49). The SV40 ori is made of 2 LTag binding motifs on the early side of the ori called *site I*, and four centrally located LTag binding sites called *site II* within the ori (49). A similar ori organization is found in BKPyV ww, JCPyV ww, Merkel cell polyomavirus, human polyomavirus 9, human polyomavirus 12 (49). As reported for SV40, each base pair of the pentanucleotide sequence is critical for LTag binding and viral replication (59). However, adjacent sequences to these LTag binding sites have little or no influence on the binding of LTag and viral replication (59). This suggests that the pentanucleotide sequence at the viral ori is the fundamental functional unit for LTag binding and viral DNA replication (339). Adjacent LTag binding sites increase the stability of LTag binding. In SV40, the main function of LTag binding to *site I* is the auto-regulation of EVGR expression (59). LTag binds to *site I* and represses the transcription of EVGR, whereas the function of *site II* is to bind and position LTag in the correct orientation and location for its function in viral genome replication (59). The 6 LTag binding sides near the EVGR side of BKPyV ww are similarly organized compared to the 6 LTag binding sites of SV40 (49). The roles of the SV40 LTag motifs in the regulation of the SV40 bidirectional expression have been well described (59), however, very little is known about that of BKPyV. The LTag binding motifs between SV40, BKPyV and JCPyV at the ori are the same, we hypothesize that the role of the LTag binding motifs in BKPyV is similar to that of SV40 (49). In order to elucidate the role of BKPyV LTag-binding motif, we compared in a

bidirectional reporter vector the EVGR and LVGR expression of BKPyV ww NCCR with those of BKPyV ww NCCR defective in one or more LTag-binding motif in HEK293 cells expressing none, a small or large amounts of SV40 LTag. Our results showed the mutants containing mutant A displayed a lesser decrease in HEK293TT cells compared to other mutants. While the LVGR expression proportional increase is abrogated for mutants harboring mutants, A and D.

RESULTS

LTag mutagenesis and mutant naming

We identified by MatInspector software seven LTag binding sites (Fig. 1). Six of these sites are contained in the ori and are similarly organized compared to the six LTag binding sites of SV40 (49). We performed different mutageneses within the LTag-binding sites in the BKPyV ww NCCR by substituting the first two bases of the sites as follows; (G->T, C->A or vice versa), and the mutant NCCRs were reanalyzed by MatInspector for LTag binding, while allowing 2 mismatches. *In silico* analyses indicated no LTag binding after introducing two substitutions. To ease naming of the mutants, we assigned the DNA strand that goes in 5' to 3' direction from the EVGR via the NCCR to the LVGR as the *forward strand.* Likewise, we named the strand that goes in 5' to 3' direction from the LVGR through the NCCR to the EVGR as the *reverse strand.* Two sites were identified on the *forward strand* (pink) (5'-GRGGC-3') and five on the *reverse strand* (yellow) (5'-GCCYC-3') (Fig. 1). The BKPyV contains 6 LTag binding sites within the ori (49). Two LTag binding sites referred to as *site A* proximal to the EVGR (*site I* in the SV40 nomenclature) and four centrally located motifs are found adjacent to an AT-rich sequence referred to as *site B* and *C* (*site II* in the SV40 nomenclature) (49). These 4 sites contained two forward motifs referred to as *site B* and two reverse motifs referred to as *site C* in the BKPyV nomenclature. The *B* and *C sites* are each separated by one nucleotide as shown for SV40 (49)**.** Finally, the LTag binding site outside the ori, referred to as *site D,* is proximal to the LVGR side (Fig.1). All the other LTag binding sites on the *reverse strand* contain a thymine at the fourth position of the reversed motifs 5`GCCYC`3 except LTag binding *site D*, which contains a cysteine instead of a thymine at this position. LTag binding sites in close proximity and on the same strand were mutated together. This provided a total of four possible LTag binding site mutants, which were labelled A (R1 and R2), B (F1 and F2), C (R3 and R4) and D (R5), where 'R' stands for reverse and 'F' for forward, while the numbers are the motifs numbering from EVGR to the LVGR (Table S1). *D site* was the only single site mutated alone (Table S1). Altogether, eight different mutants including combination of the different LTag binding site mutants were made (Table S1). The LTag binding mutagenesis affected some TFBS within the NCCR. However, apart from binding sites for E2F4 and Sp1 (highlighted in yellow)

affected due to D mutation, most of the TFBS affected are not found in the kidney (Table S2).

Effect of LTag binding site mutants on EVGR and LVGR expression in HEK293 cells

To examine the effect of the different NCCRs defective of LTag binding, we cloned them in the bidirectional reporter vector pRG13D12, transfected HEK293 cells and measured the EVGR and LVGR expression by flow cytometry. Mutation of *B*, *C* and *D sites* showed a gain of function for EVGR expression in HEK293 cells, while all other mutant combination containing *A site* mutation presented with a loss of function for EVGR expression compared to the un-mutated archetype. Conversely, mutant combinations containing mutation in *A* or *D site* alone, revealed a loss of function for LVGR expression compared to the archetype. Whereas, the LVGR expression of mutants, *B* and *C sites* showed little or no effect.

Effect of LTag binding site mutants on EVGR and LVGR expression in HEK293-derivative cell lines that constitutively expressing SV40 LTag in different amounts

To measure the effect of the different LTag binding site mutants in the presence of small and large amounts of LTag, we transfected HEK293, HEK293T and HEK293TT cells constitutively expressing none, a small or large quantity of LTag (Fig. 3). Two days post transfection we analyzed the EVGR and LVGR expression by flow cytometry as previously described (34). Furthermore, we calculated the EVGR MFI of each construct normalized to their corresponding EVGR MFI in HEK293 cells set as 1. Our result depicted 3 distinct groups based on EVGR expression patterns (Fig. 3A). We could identify group wwE (mutants with similar EVGR expression pattern as archetype), group A (all mutants harboring mutation A) and group D. Compared to the corresponding expression in HEK293 cells, all groups exhibited a gain of function for EVGR expression in HEK293T cells. The EVGR expression in HEK293TT cells in the 3 groups were as follows; for group wwE the EVGR expression was reduced to the levels in HEK293 cells, while for group A it remained at similar levels to the corresponding levels in HEK293T cells and lastly, for group D the level was reduced

below the corresponding expression in HEK293 cells (Fig.4A). In addition, mutants containing *A site* mutation generally showed higher EVGR expression fold in both cell lines expressing LTag compared to other mutants (Fig. 4A). Mutations in *A site* presented with very slight decrease of EVGR in HEK293TT cells compared to other mutantions (Fig.4A). This implies that abolishing LTag interaction to the *site A* by mutagenesis prevents the EVGR expression decrease in HEK293TT cells as seen for other mutants in combination with *A site* mutant.

We also calculated the LVGR MFI of each construct normalized to their corresponding LVGR MFI in HEK293 cells set as 1. With respect to the LVGR, we were able to identify 2 different groups regarding LVGR expression patterns, in particular group wwL (mutants with similar LVGR expression pattern as archetype) and group AD (all mutants harboring both mutation A and D). Compared to corresponding expression in HEK293 cells both groups revealed a gain of function for LVGR expression in HEK293T and HEK293TT cells. However, all except mutants containing A and D mutations displayed a proportional LVGR expression corresponding to increasing amounts of LTag (Fig. 4B). This indicates that mutation in both *A* and *D sites* together abrogated further LVGR activation in HEK293TT cells. Interestingly, mutants with either A or D alone still permitted the LVGR activation in HEK293TT cells (Fig. 4B), thus indicating a cooperative action of LTag interaction to both sites to increase LVGR expression in the presence of high amounts of LTag expression.

DISCUSSION

Here we report that reduced-EVGR expression in the presence of large amounts of LTag expression can be via LTag binding *site A*, LVGR expression increase can be via both *sites A* and *D.* The oris of the 13 HPyVs has been reported (49). The number, orientation and spacing of LTag binding motifs at the ori of SV40, BKPyV and JCPyV are the same (49). SV40 and BKPyV core oris contain 4 centrally identical LTag binding motifs termed *site II* with respect to SV40 and *site B, C* with respect to BKPyV nomenclature (49). The two other motifs located proximally to the EVGR termed *site I* with regard to SV40 and *site A* with respect to BKPyV nomenclature are separated by 7 A-rich bps. BKPyV has a 7th LTag binding motif termed D motif located proximal to the LVGR side (Fig.1). Other HPyVs such as, MCPyV, HPyV10 and STLPyV also encode LTag binding motifs outside their oris, proximal to the LVGR side (Fig. 1 section 5.1). However, the role of these sites is yet to be elucidated.

Since the roles of the SV40 LTag motifs with respect to regulation of EVGR and LVGR and viral genome replication are well known and the motifs between SV40, BKPyV and JCPyV at the ori are the same, we hypothesize that mutageneses of BKPyV`s LTag-binding sites will show similar effect compared to SV40. We analyzed EVGR and LVGR expression of BKPyV ww NCCR defective in one or more LTag-binding sites compared to BKPyV ww-NCCR in HEK293 cells expressing none, a small or large amount of SV40 LTag.

To get an indication of the effect of the LTag mutants in the absence of LTag, we analyzed the EVGR and LVGR expression in HEK293 cells. All mutations harboring *site A* mutagenesis displayed a loss of function for EVGR, while mutants *B*, *C*, and *D* showed a gain of function for EVGR expression compared to BKPyV ww-NCCR (Fig. 2A), suggesting that mutagenesis of *site A* may results in a loss of activating transcription factor binding sites (TFBS) or gain of repressing TFBS (Table S2). However, the reverse could be plausible for the other mutants. On the other hand, mutants *A*, *D* and mutants harbouring *A site* mutantion exhibited a loss of function for LVGR expression, whereas mutant *B* and *C* has little or no effect, indicating a loss of activating TFBS once the *A* and *D sites* were mutated. Specifically, mutation in the *D*

site showed an *in-silico* loss of Sp1 and E2F4 binding, which were the only affected factors found in the kidney (Table S2). This is in line with the published work of Bethge et al., whereby Sp1-4 core site mutagenesis led to an increase in EVGR in expense of LVGR expression (340). However, our results were not exactly the same compared to that of Bethge et al, which could be due to the fact that the D site mutation did not affect the Sp1-4 core binding site, suggesting Sp1 may still be interacting with the binding site in a reduced affinity.

Next, we investigated the effect of these NCCR-mutants defective in LTag binding in HEK293 cells producing small and large amounts of SV40 LTag, HEK293T and HEK293TT cells, respectively. Interestingly, all mutants including mutant combination where all LTag-binding sites were disrupted, showed higher EVGR expression in HEK293T cells compared to HEK293 cells. Moreover, the EVGR expression pattern between the different mutants were regulated differently in HEK293T and HEK293TT cells, resulting in the depiction of 3 different groups that is groups, wwE, A and D (Fig. 3A). Group A (mutants harbouring mutant *A*) generally shows higher EVGR expression in both HEK293T and HEK293T cells compared to other mutants, indicating that mutation of binding *site A* leads to gain of function for EVGR expression in the context of LTag expression as previously reported for SV40 with *site I* mutation (341). Intriguingly, all mutants including the mutant where all the LTag binding motifs were mutated (ABCD) presented with elevated levels of EVGR expression in HEK293T cell. Further studies are required to examine this elevation, which most probably do not involve LTag interaction to its binding sites, however, may reflect LTag inhibition of EVGR-negative regulatory factors or increase in plasmid copy numbers. However, our preliminary copy number quantification does not correlate with the EVGR expression (Fig. 4C).

Furthermore, we showed that EVGR expression of mutants in group wwE and D increased in HEK239T cells expressing small amounts of LTag and greatly decreased in HEK293TT cells expressing large amounts of LTag compared to mutants containing *A* mutation (group A). In summary, mutating the LTag-binding *site A* or combination of LTag-binding *site A* with other LTag-binding sites (group A) abrogated the EVGR-expression decrease, indicating a specific role of LTag-binding sites with respect to EVGR expression.

In addition, except mutants containing both *A* and *D* mutation (group AD) there was a proportionate increase in LVGR expression in the presence of increasing amounts of LTag expression, which was similarly described for JCPyV LVGR activation by SV40 LTag (342). This suggests that the proportionate increase of LVGR expression can be mediated via sites *A* and *D*, since mutants containing either *A* or *D* or other LTag binding sites still show this proportional increase in LVGR expression. Taken together, while LTag-binding site A can be involved in the decrease of EVGR expression in abundant LTag expression, LVGR expression can be increased by cooperative involvement of LTag-binding sites *A* and *D*. Interestingly, like the EVGR expression of all mutants including mutant ABCD displayed elevated levels of LVGR expression in HEK293T cells. However, the mechanism cannot be deciphered with this current study. However, LTag inhibition of negative LVGR regulatory factors or copy number effect can be suspected. A summary model of the results is depicted in Fig. 5.

Lastly, *site II* in SV40 is reported to play a role in SV40 replication, but not expression (59). Mutation of a corresponding site in BKPyV, sites *B* and *C,* showed little or no effect with respect to EVGR and LVGR expression in the context of LTag expression compared to the archetype, thus corroborating the SV40 results (59). Our current study cannot explain viral replication, Nonetheless, our preliminary copy number quantification does not correlate with the expression, indicating copy numbers might not play a role in the expression pattern depicted.

In this study, we analyzed the effect of SV40 LTag on the BKPyV ww-NCCR defective of LTag binding motifs. However, one cannot exclude that the BKPyV cognate LTag might give different results. Nonetheless, the SV40 and BKPyV LTag have great similarity, and the fact that BKPyV can replicate in Cos-7 cells constitutively expressing SV40 LTag argue against this possibility. Here, EVGR and LVGR expression were analyzed using the bidirectional reporter vector. Further experiments are required to quantify the viral replication potential of the various NCCR mutant defective of LTag binding motif by cloning in a Dunlop backbone and perform viral replication studies in cell culture. However, the fact that BKPyV and SV40 have similar LTag binding sites and show similar EVGR and LVGR expression regulatory mechanism via the LTag binding sites, suggests that results obtained from

studying BKPyV can be extrapolated to other HPyVs such as, JCPyV with similar LTag binding site organization.

MATERIALS AND METHODS

Plasmids

We used the bidirectional reporter vector designed a new-smaller bi-directional reporter pRG13D12 described, similar to the phRG reporter described in Gosert et al (25). The reporter lacks the SV40 early promoter, beta globulin poly-adenylation signals, Hygromycin gene present in the NCCR cloned in phRG (25). PUC57 vector containing HPyV NCCR without the 12-nucleotide sequence. NCCRs defective of LTag binding were purchased from Eurogentec (Eurogentec S.A, Belgium). The NCCRs were excised from the pUC57 vector by using the restriction enzymes BssHII and MluI (New England Biolabs, England). The NCCRs were cloned via the corresponding restriction sites in the bi-directional vector, pRG13D12. Constructs with the right NCCR were verified by the 3130 Genetic Analyzer Sanger sequencer (Applied Biosystems, Switzerland).

Cell lines

The following cell lines, HEK293 cells (CRL1573, ATCC, Manassas, VA, USA), HEK293T cells (CRL3216, ATCC, Manassas, VA, USA), HEK293TT cells (0508000, NCI-Frederick Repository Service, Frederick, Maryland, USA), were grown in Dulbecco's modified Eagle's medium with high glucose (DMEM-H, D5671, Sigma-Aldrich, St. Louis, MO, USA) containing 10% FBS (S0113, Biochrome AG, Berlin, Germany) and 2mM L-glutamine (K0302, Biochrome AG, Berlin, Germany). The medium of HEK293TT cells was supplemented with 400µg/mL hygromycin B (Calbiochem, San Diego, CA, USA). The cells were cultured in a HERA cell-150 incubator (Thermo Fischer Scientific, Switzerland) at a temperature of 37°C and 5% CO_2.

Western blot

HEK293, HEK293T, and HEK293TT cells were lysed in triple-buffer (50 mM Tris-HCl [pH 8.0], 150 mM NaCl, 1% NaDOC, 1% Triton X-100 and 0.1% SDS) containing 1x protease inhibitors. Equal volume of 10µg of cell lysates were separated by SDS-PAGE using running buffer (50 mM Tris base, 199mM Glycine, 0.1% SDS, pH 8.3) at

25mA for 50 mins. The proteins were then electro-transferred (semi dry method) onto 0.45 µm polyvinylidene difluoride (PVDF) membrane (IPFL00010; Millipore/Merck, Darmstadt, Germany) using 1x transfer buffer (10x: 48mM Tris base, 390m M Glycine, 10% SDS) at 70mA for 48 minutes. The membrane was dried, reactivated with 5 mL methanol (Sigma-ALDRICH, Switzerland) and washed with milli-Q H_2O. Odyssey blocking buffer (927-40000; Licor, Lincoln, NE, USA) diluted 1:2 in Tris-buffered saline (TBS) was used to block the membrane at RT for 1 h. Incubation of the membrane was done in the following primary antibodies: monoclonal mouse anti-actin (1:5'000; Abcam, Cambridge, England), polyclonal rabbit anti-LTag (1:10'000; Professor Christine Hanssen Rinaldo, University of North Norway in Tromsö, Norway), diluted in 1:2 Odyssey blocking buffer-TBS-0.1% Tween20 at 4^{o}C overnight. The membrane was then washed at least 5x with TBS-0.1% Tween 20. Next, the membrane was incubated in the following secondary antibodies: donkey anti mouse Alexa Fluor 680 (1:15000; A10038; goat anti-rabbit antibody-IRDye 800CW (1:10,000; 926-32211; Licor), invitrogen) diluted in 1:2 Odyssey blocking buffer TBS, 0.1% Tween20 at RT for 1hr. The membrane was washed at least 3x with TBS-0.1% Tween20. Detection and quantification of protein was done with the Licor Odyssey CLx instrument (Licor, Homburg, Germany).

Transfection

All cell lines were transfected with 3µL Lipofectamine 2000 (1734997, Invitrogen, Carlsbad, CA, USA) and 1µg plasmid DNA construct in a 12 well plate, according to manufacturer's instructions. Prior to transfection the cells' confluency ranged between 70-90%. 24 hrs post-transfection, medium was replaced with DMEM-H/10% FBS complemented with 2mM Glutamine. 48 hrs post-transfection, fluorescence images were taken by fluorescent microscopy and EVGR (RFP) and LVGR (GFP) expression quantified by flow cytometry.

Flow cytometer-based quantification

Forty-eight-hour post-transfection cells were washed once with 1mL PBS/2.5 mM EDTA (BioConcept, Switzerland) and then trypsinized with trypsin (without phenol red) /0.25mg/mL EDTA (Lonza, MD, USA). The cells were suspended in 1ml DMEM (without phenol red) containing 20mM of HEPES, 2mM Glutamine, 1% FCS and 1% Penicillin/Streptomycin (P/S) (Thermo Fischer Scientific, Switzerland) and transferred to 5 ml polystyrene round-bottom tubes (BD, Franklin Lakes, NJ, USA). Prior to each measurement, DAPI (D8417, Sigma-Aldrich, St. Louis, MO, USA) was added as a dead-cell marker to a final concentration of 1 ng/ml. Fortessa Cytometer (BD, Franklin Lakes, NJ, USA) was used for the flow cytometry measurements. RFP, Excitation was 561 nm (Yellow-Green Laser), Emission was 586/15 nm; GFP, Ex. 488 nm (Blue Laser), Em. 530/30 nm and DAPI, Ex. 405 nm (Violet Laser), Em. 450/50 nm. The weighted mean fluorescence intensities (MFI) for EVGR (red) and LVGR (green) expression were calculated by inserting the cell number (N), mean fluorescence (I) of the quadrants Q1 (red cells only), Q2 (red and green cells), and Q4 (green cells only) into the respective formulas: $\mathrm{MFI(red)} = \frac{(N_{Q1}*I_{Q1})+(N_{Q2}*I_{Q2})}{(N_{Q1}+N_{Q2}+N_{Q4})}$; $\mathrm{MFI(green)} = \frac{(N_{Q2}*I_{Q2})+(N_{Q4}*I_{Q4})}{(N_{Q1}+N_{Q2}+N_{Q4})}$. EVGR expression of all HPyV constructs was normalized to the *EVGR* of BKPyVww set as 1 while LVGR expression of all HPyV constructs was normalized to the *LVGR* of BKPyVww set as 1.

Plasmid copy number quantification

Plasmid copy numbers were quantified after total DNA extraction from cells suspended in 200ul PBS. A QIAamp DNA blood kit (Qiagen Hombrechtikon, Switzerland) was used for the extraction according to the manufactural instructions. Quantitative PCR protocol for detecting dsRed2, primers and probes shown below was used.

The following primer and probe was used for the qPCR

	Sequence (5' – 3')	Length/%GC	Tm [°C]
Primer			
DsRed2_f	GGT GGA GTT CAA GTC CAT CTA CAT G	25 bp/48	57.1
DsRed2_r	CCA CGA TGG TGT AGT CCT CGT T	22 bp/54.5	58.7
Probe			
DsRed2_FAM	FAM- CGG CTA CTA CTA CGT GGA CGC CAA GCT-TAMRA	27 bp/59.3	65.1

Amplicon: 121bp

Primers and probe concentrations used for qPCR

2X Mastermix	**10** µM	**Per Reaction**	**Per x10**	**Final Conc. nM**
qPCR Mix		12.50 µL	125.00 µL	
Forward Primer	10	0.25 µL	2.50 µL	100
Reverse Primer	10	0.25 µL	2.50 µL	100
Probe	10	0.25 µL	2.50 µL	100
Template	DNA	5.00 µL	-	
Water		6.75 µL	67.50 µL	
Total volume		25 µL	-	

In silico analysis of BKPyV ww NCCR

The MatInspector software tool (Genomatix, Munich, Germany)(324) was used to search for l potential LTag binding sites within the BKPyV ww NCCR. For identifying potential LTag binding sites the BKPyVww NCCR was uploaded into the software and the user-defined IUPAC string using the LTag binding site consensus sequence, GRGGC (where R = A or G) allowing no mismatch.

Statistics

The Mean and standard deviations were calculated from three independent transfections using GraphPad Prism software (version 6 for Mac OS).

FIGURE LEGEND

Fig. 1: **Presentation of BKPyV ww NCCR with putative LTag binding sequence motifs.** The LTag binding motifs are numbered from left to right (1-5) according to their respective orientations. (*R= reverse and F= forward*), binding *site A* contains 2 motifs (*R1 and R2*), *site B* (*F1 and F2*), *site C* (*R3 and R4*) and site D (*R5*). The number of the bp of each site is depicted. Lastly, the EVGR and LVGR start codon (ATG) are shown on the reverse and *forward strands*, respectively.

Fig. 2: **Differential effect of LTag binding site mutants in HEK293 cells.** (A) MFI of EVGR expression normalized to EVGR MFI BKPyVww set as 1. (B) MFI of LVGR expression normalized to EVGR MFI BKPyVww set as 1. NCCR driven EVGR and LVGR expression measured by FACS and MFI calculated as described in materials and methods. Error bars represent standard deviation of three independent experiments.

Fig. 3: Western blot of LTag expression in HEK293, HEK293T and HEK293TT cell lines.

Fig. 4: **Downregulation of EVGR and upregulation of LVGR expression in abundant amounts of LTag.** (A) EVGR expression fold in HEK293T and HEK293TT cells normalized to corresponding EVGR expression in HEK394 cells set as 1. Based on EVGR expression patterns, 3 groups were depicted (groups wwE, A and D). (B) LVGR expression fold in HEK293T and HEK293TT cells normalized to corresponding LVGR expression in HEK394 cells set as 1. Based on LVGR expression patterns, 2 groups were depicted (groups wwE, AD). Error bars indicate standard deviation of three independent experiments.

Fig. 5: **A model depicting the effect of LTag-binding sites mutageneses. (**A) In HEK293 cells containing no LTag, host transcription factors (orange ovals) interact with the BKPyV ww NCCRs to maintain the bidirectional EVGR (red arrow) and LVGR (green arrow) expression balance. (B) In the presence of small amount of LTag in HEK293T cells (few yellow triangles), EVGR and LVGR expression is increased via an unknown mechanism, (LTag binding sites, A = inverted V in light blue, B and C = inverted V in thick blue and D = inverted U in pink), since the mutant with all LTag binding sites mutated, displayed higher levels of EVGR expression. This may reflect inhibition of EVGR and LVGR negative regulatory factors by LTag (pink oval in contact with yellow triangle). (C) In the presence of large amounts of LTag in HEK293TT cells, LVGR expression increase even more for all mutants except in mutants containing mutant *A* and *D* mutations. Indicating the addition increase is mediated via *A* and *D sites*. Whereas, EVGR expression decreases compared to the levels in HEK293T cells for all mutants except those harboring mutant A indicating that the reduced EVGR expression involves site A. *B* and *C sites* showed little or no influence of EVGR expression in HEK293 cells expressing LTag. The expression pattern is similar to that of the ww NCCR. These sites may be involved in viral replication as described for a homologous *site II* in SV40.

Supplementary data

Table S1: Binding sites that are mutated in each of the mutants.

Table S2: TF binding sites that are gained or lost by mutations in LTag binding sides.

FIGURES

Fig. 1

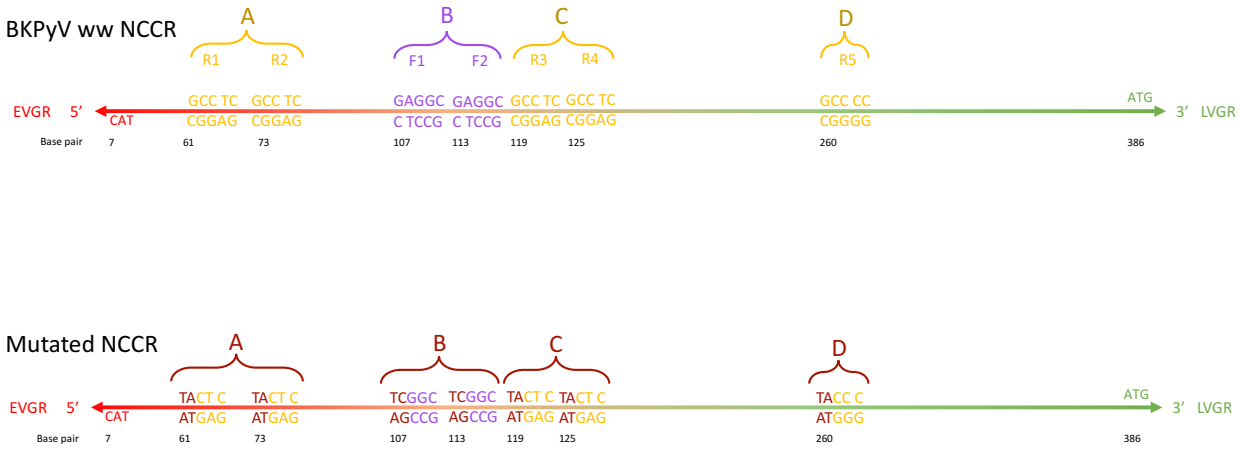


Fig. 2

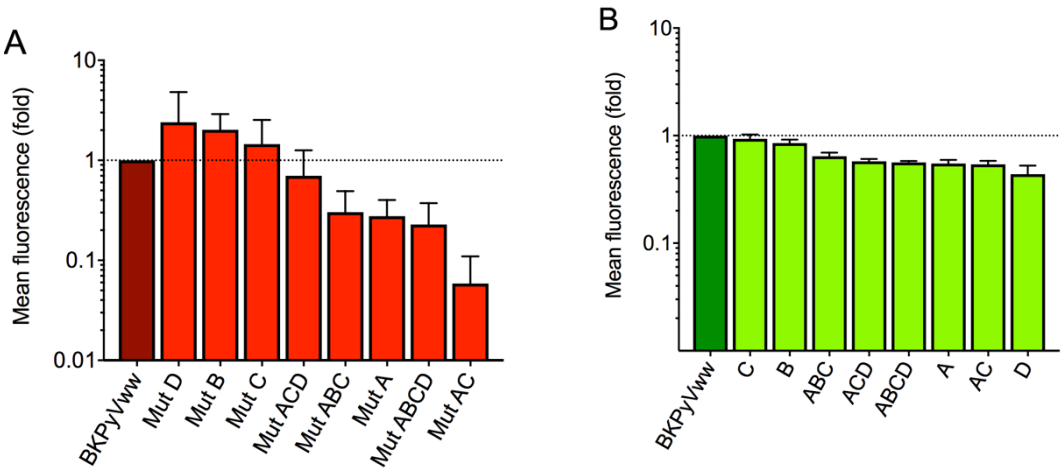


Fig. 3

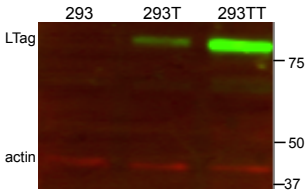

293
293T
293TT
LTag
actin
75
50
37

Fig. 4

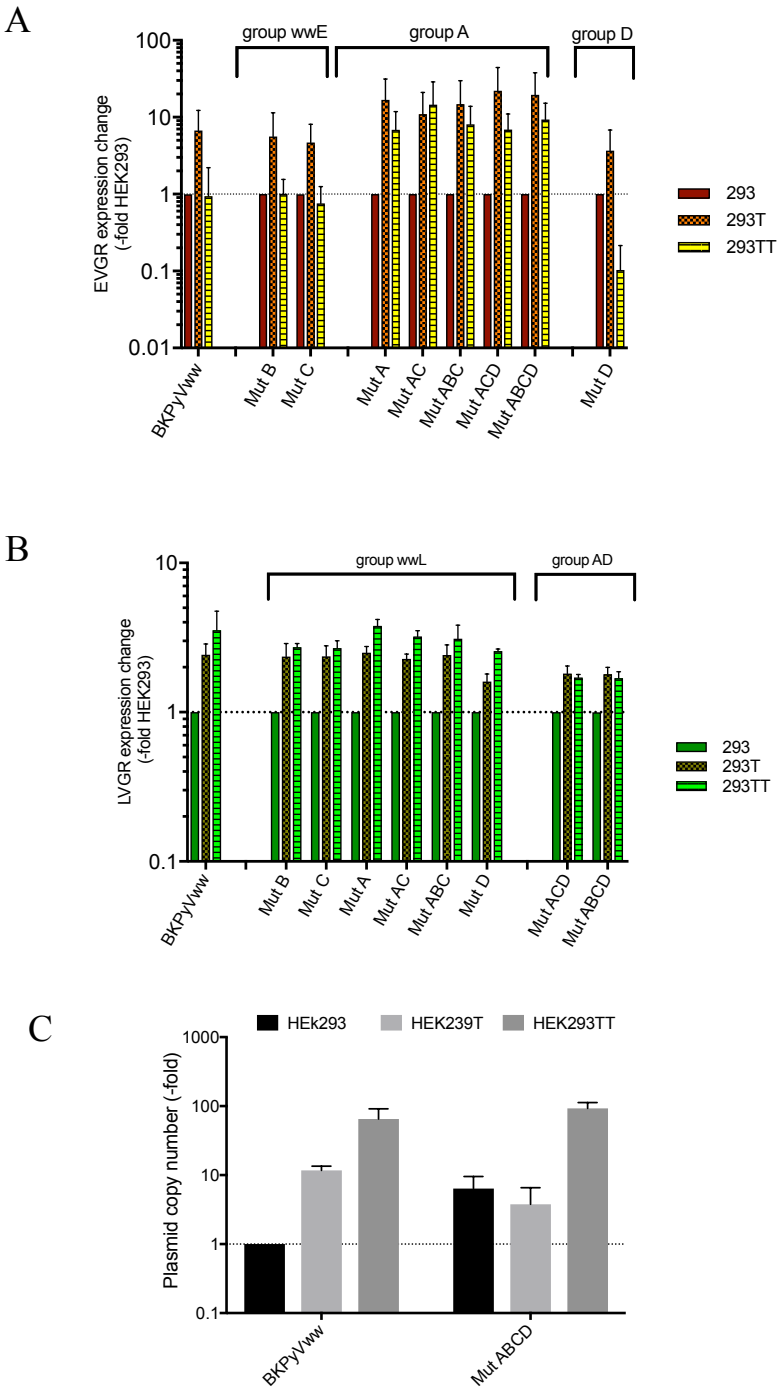

A
group wwE
group A
group D
EVGR expression change
(-fold HEK293)
100
10
1
0.1
0.01
BKPyVww
Mut B
Mut C
Mut A
Mut AC
Mut ABC
Mut ACD
Mut ABCD
Mut D
293
293T
293TT
B
group wwL
group AD
LVGR expression change
(-fold HEK293)
10
1
0.1
BKPyVww
Mut B
Mut C
Mut A
Mut AC
Mut ABC
Mut D
Mut ACD
Mut ABCD
293
293T
293TT
C
HEk293
HEK239T
HEK293TT
Plasmid copy number (-fold)
1000
100
10
1
0.1
BKPyVww
Mut ABCD

Fig. 5

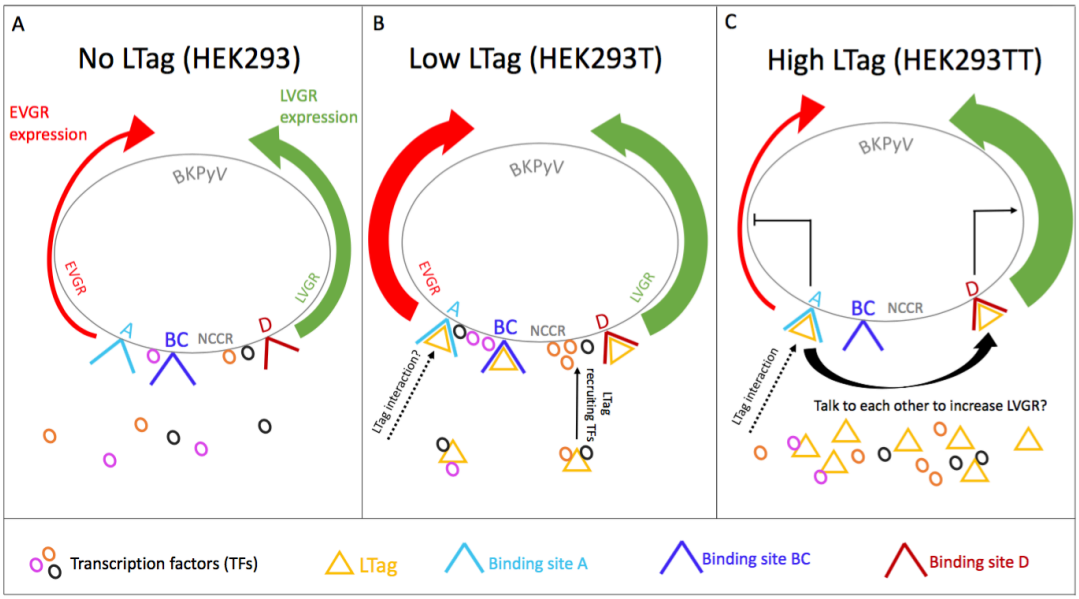

A
No LTag (HEK293)
EVGR expression
LVGR expression
BKPyV
EVGR
LVGR
A
BC
NCCR
D
B
Low LTag (HEK293T)
BKPyV
EVGR
LVGR
A
BC
NCCR
D
LTag interaction?
recruiting TFs
LTag
C
High LTag (HEK293TT)
BKPyV
A
BC
NCCR
D
LTag interaction
Talk to each other to increase LVGR?
Transcription factors (TFs)
LTag
Binding site A
Binding site BC
Binding site D

Table S1

Mutation	TF	Gain (+) / loss (-) of TF site in mutant	TF detail	Tissues in which TFs are found	Core sim.	Matrix sim.
A	CIZ	-	CAS interating zinc finger protein	Skeleton	1	0.976
	NKX31	+	NK3 homeobox 1	Cardiovascular System/ Central Nervous System/ Digestive System/ Embryonic Structures/Endocrine System/ Heart/ Islets of Langerhans/ Lung/ Muscles/ Myocardium/ Nervous System/ Pancreas/ Prostate/ Respiratory System/ Spinal Cord/ Thyroid Gland/ Urogenital System	1	0.817
	RFX3	-	Regulatory factor X 3	Antibody-Producing Cells/ Blood Cells/ Endocrine System/ Immune System/ Leukocytes/ Lymphocytes/ Testis/ Urogenital System	1	0.762
	ZFX	-	X-linked zinc finger protein	Embryonic Structures/ Endocrine System/ Germ Cells/ Ovary/ Testis/ Urogenital System	1	0.971
B	KLF2	-	Krueppel-like factor 2 (lung) (LKLF)	Blood Cells/ Bone Marrow Cells/ Embryonic Structures/ Erythrocytes/ Hematopoietic System	1	0.986
	PURα	-	Purine-rich element binding protein A	Brain/ Central Nervous System/ Muscle, Smooth/ Muscles/ Nervous System/ Neuroglia/ Neurons	1	0.97
	ZNF652	-	Zinc finger protein 652		1	0.76
C	ZFX	-	X-linked zinc finger protein	Embryonic Structures/ Endocrine System/ Germ Cells/ Ovary/ Testis/ Urogenital System	1	0.974
D	ZBED4	-	Zinc finger, BED-type		1	0.96
	E2F4	**-**	**E2F transcription factor 4**	**ubiquitous**	**1**	**0.975**
	LRRFIP1	-	Leucine rich repeat (in FLII) interacting protein 1		0.842	0.821
	GCM1	+	Glial cells missing homolog 1	Embryonic Structures/ Endocrine System/ Immune System/ Parathyroid Glands/ Thymus Gland	1	0.834
	INSM1	-	Zinc finger protein insulinoma-associated 1 (IA-1) functions as a transcriptional repressor	Digestive System/ Endocrine System/ Liver/ Pancreas	1	0.918
	KLF7	-	Krueppel-like factor 7	Blood Cells/ Bone Marrow Cells/ Embryonic Structures/ Erythrocytes/ Hematopoietic System	1	0.987
	MAZR	-	MYC-associated zinc finger protein	Blood Cells/ Immune System/ Leukocytes	1	0.898
	SP1	**-**	**Stimulating protein 1**	**ubiquitous**	**1**	**0.987**
	ZBTB7	-	Zinc finger and BTB domain containing 7A	Blood Cells/ Immune System/ Leukocytes/ Lymphocytes/ Thymus Gland	1	0.927

Table S2

Mutants	Mutated binding sites
A	R1, R2
B	F1, F2
C	R3, R4
D	R5
AC	R1, R2, R3, R4
ABC	R1, R2, R3, R4, F1, F2
ACD	R1, R2, R3, R4, R5
ABCD	R1, R2, R3, R4, R5, F1, F2

Co-authored publications during the course of this PhD thesis.

1. Tobias Bethge, **Elvis T. Ajuh,** Hans H. Hirsch* (2016) *Imperfect Symmetry of Sp1 and Core Promoter Sequences Regulates Early and Late Viral Gene Expression of the Bi-Directional BK Polyomavirus Non-Coding Control Region* J. Virol 90: 10101.
   Contribution: Performed transfection, western blot, contributed in writing the manuscript and approved the final copy.

2. *Donor-derived urothelial cancer after kidney transplantation associated with a BK polyomavirus with increased oncogenic potential*
   Contribution: Carried out experiments in molecular characterizing the BKPyV from the urothelial cancer and contributed in writing the manuscript.

5.4 Imperfect symmetry of Sp1 and core promoter elements regulates early and late viral gene expression of the bi-directional BK polyomavirus non-coding control region

Specificity protein 1 (Sp1) has a crucial role in the transcription of both cellular and viral genes (343)**.** All cells of the human body ubiquitously express this transcription factor (344)**.** Sp1 binds to GC-box motifs found in SV40 and BKPyV NCCR and its binding is vital for the regulation of EVGR and LVGR expression (340). This GC-box motif is found on the NCCR of all HPyVs, suggesting a vital role for this sequence in the growing family of HPyVs. Our group reported a detailed point mutants of the BKPyVww NCCR defective in 28 different TFBS without changing the overall BKPyV ww-NCCR length and architecture. A hierarchy of TFBS was identified, whereby 3 phenotypic groups were detected: group 1 TFBS mutations revealed activation of EVGR expression in expense of LVGR expression and strong replication rates similar to what was reported for BKPyV rr-NCCR variants from patients. Group 2 TFBS mutants displayed an intermediate phenotype with regards to EVGR expression and replication capacity, while group 3 showed reduced EVGR and LVGR expression and allowed modest or no viral replication (340).

Among all the TFBS, Sp1 sites revealed a prominent role in BKPyV NCCR-driven expression. BKPyV NCCR has at least 3 Sp1 binding sites; two of these sites (Sp1-2 and Sp1-4) appeared to have underlying opposite effects in the expression and replication of BKPyV. Mutagenesis of the Sp1-4 site proximal to the LVGR demonstrated a strong EVGR expression, whereas inactivation of Sp1-2 site proximal to the EVGR overrode both EVGR and LVGR expression. This effect of *sp1-2* mutant was also recorded in group 1 mutants including *sp1-4* mutant, indicating a vital regulatory role of SP1 in the NCCR as a whole (340).

The role of these two Sp1 sites and their relationship to core promoter elements in the BKPyV ww-NCCR was examined in greater detail. This role of Sp1 sites in BKPyV could be extrapolated to other HPyVs, especially those with similar Sp1 sites and core promoter elements (CPEs) architecture. We examined the bi-directional balance of EVGR and LVGR expression and viral replication with respect to Sp1-2 and Sp1-4 sites strand orientation, affinity and number of Sp1 sites. In addition, the contribution of EVGR and LVGR CPEs in EVGR and LVGR expression was also investigated

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology

Imperfect Symmetry of Sp1 and Core Promoter Sequences Regulates Early and Late Virus Gene Expression of the Bidirectional BK Polyomavirus Noncoding Control Region

Tobias Bethge,[a]* Elvis Ajuh,[a] Hans H. Hirsch[a,b,c]

Transplantation & Clinical Virology, Department of Biomedicine (Haus Petersplatz), University of Basel, Basel, Switzerland[a]; Division of Infection Diagnostics, Department of Biomedicine (Haus Petersplatz), University of Basel, Basel, Switzerland[b]; Infectious Diseases & Hospital Epidemiology, University Hospital Basel, Basel, Switzerland[c]

ABSTRACT

Rearrangements or point mutations in the noncoding control region (NCCR) of BK polyomavirus (BKPyV) have been associated with higher viral loads and more pronounced organ pathology in immunocompromised patients. The respective alterations affect a multitude of transcription factor binding sites (TFBS) but consistently cause increased expression of the early viral gene region (EVGR) at the expense of late viral gene region (LVGR) expression. By mutating TFBS, we identified three phenotypic groups leading to strong, intermediate, or impaired EVGR expression and corresponding BKPyV replication. Unexpectedly, Sp1 TFBS mutants either activated or inhibited EVGR expression when located proximal to the LVGR (*sp1-4*) or the EVGR (*sp1-2*), respectively. We now demonstrate that the bidirectional balance of EVGR and LVGR expression is dependent on affinity, strand orientation, and the number of Sp1 sites. Swapping the LVGR-proximal high-affinity *SP1-4* with the EVGR-proximal low-affinity *SP1-2* in site strand flipping or inserting an additional *SP1-2* site caused a rearranged NCCR phenotype of increased EVGR expression and faster BKPyV replication. The 5′ rapid amplification of cDNA ends revealed an imperfect symmetry between the EVGR- and LVGR-proximal parts of the NCCR, consisting of *TATA* and *TATA*-like elements, initiator elements, and downstream promoter elements. Mutation or deletion of the archetypal LVGR promoter, which is found in activated NCCR variants, abrogated LVGR expression, which could be restored by providing large T antigen (LTag) in *trans*. Thus, whereas Sp1 sites control the initial EVGR-LVGR expression balance, LTag expression can override inactivation of the LVGR promoter and acts as a key driver of LVGR expression independently of the Sp1 sites and core promoter elements.

IMPORTANCE

***Polyomaviridae* currently comprise more than 70 members, including 13 human polyomaviruses (PyVs), all of which share a bidirectional genome organization mediated by the NCCR, which determines species and host cell specificity, persistence, replication, and virulence. Here, we demonstrate that the BKPyV NCCR is fine-tuned by an imperfect symmetry of core promoter elements centered around *TATA* and *TATA*-like sequences close to the EVGR and LVGR, respectively, which are governed by the directionality and affinity of two Sp1 sites. The data indicated that the BKPyV NCCR is poised toward EVGR expression, which can be readily unlatched by a simple switch affecting Sp1 binding. The resulting LTag, which is the major EVGR protein, drives viral genome replication, renders subsequent LVGR expression independently of archetypal promoter elements, and can overcome enhancer/promoter mutations and deletions. The data are pivotal for understanding how human PyV NCCRs mediate secondary host cell specificity, reactivation, and virulence in their natural hosts.**

The *Polyomaviridae* comprise more than 70 polyomavirus (PyV) members, including at least 13 human PyV (HPyV) species (1). HPyV seroprevalence data obtained using the major capsid protein Vp1 as pentamers or virus-like particles indicate that HPyVs infect 50% to 90% of the general human population without known specific signs or symptoms of disease (2–7). As information about the virology and pathology of the 13 human PyVs is only emerging (8, 9), we focused on BK polyomavirus (BKPyV) as a model (10). After primary infection in early childhood, BKPyV persists latently in the renourinary tract with periods of low-level shedding into urine (4, 11, 12). Immunosuppression, which is commonly administered posttransplantation, has been shown to relax the adaptive antiviral immune control (13–16) and to permit high-level replication with median urine BKPyV loads of over 7 $\log_{10}$ copies/ml that become apparent as decoy cell shedding in urine (15–20). Progression to BKPyV disease is most consistently encountered as nephropathy in 1 to 15% of kidney transplant recipients (21–23) or as hemorrhagic cystitis in 5 to 20% of allogeneic hematopoietic stem cell transplant recipients (24–26).

BKPyV disease involves viral genotypes indistinguishable from those of archetype BKPyV strains shed asymptomatically in the urine of healthy individuals (27–29), arguing that impaired adaptive immune control is necessary and sufficient. However, immu-

Received 23 May 2016 **Accepted** 22 August 2016

Accepted manuscript posted online 31 August 2016

Citation Bethge T, Ajuh E, Hirsch HH. 2016. Imperfect symmetry of Sp1 and core promoter sequences regulates early and late virus gene expression of the bidirectional BK polyomavirus noncoding control region. J Virol 90:10083–10101. doi:10.1128/JVI.01008-16.

Editor: L. Banks, International Centre for Genetic Engineering and Biotechnology

Address correspondence to Hans H. Hirsch, hans.hirsch@unibas.ch.

* Present address: Tobias Bethge, Kantonsspital Aarau, Aarau, Switzerland.

nosuppressive drug-specific effects (20, 30) and viral determinants (31, 32) may also contribute to the evolution, kinetics, and severity of disease. At the core of what distinguishes archetype BKPyV strains with a low replicative capacity from the highly replicative strains found in kidney transplant patients are changes in the noncoding control region (NCCR) of the BKPyV genome (33). Akin to all PyVs, the BKPyV NCCR consists of an approximately 400-bp-long sequence which is sandwiched between the early viral gene region (EVGR) and the late viral gene region (LVGR), altogether yielding the double-stranded BKPyV DNA genome of 5,100 bp (34, 35). The NCCR contains *cis*-acting elements coordinating the timing and consecutive steps of EVGR expression, viral genome replication, and LVGR expression (33, 36).

The archetype BKPyV (*ww*) found in the urine of healthy immunocompetent individuals (29, 37, 38) is defined by a linear *ww*-NCCR sequence architecture arbitrarily denoted O_{143}, P_{68}, Q_{39}, R_{63}, and S_{63} (where the subscript numbers represent the number of nucleotides), which conveys rather slow replication in primary human renal tubular epithelial and urothelial cells (31, 39). In contrast, rearranged NCCRs (*rr*-NCCRs) associated with high blood viral loads and severe renal allograft pathology carry partial duplications and/or deletions of the archetype *ww*-NCCR sequence and give rise to virus strains with elevated levels of EVGR expression, faster viral replication, and more pronounced cytopathic effects in cell culture (31, 32, 40).

Although practically all emerging *rr*-NCCR BKPyV variants appear as unique sequences, we noted some common themes. These consist of an unaltered EVGR-proximal region covering the O_{143} block with the origin of replication and the central P_{68} block. Insertion strains carry complete or partial duplications of the P_{68} block, while deletion strains are characterized by the absence of the LVGR-proximal Q_{39} and R_{63} blocks and, rarely, the S_{63} block (31, 32, 40, 41).

In a detailed point mutational analysis in which more than 28 different transcription factor binding sites (TFBS) were inactivated without altering the overall archetype *ww*-NCCR architecture and length, we identified a hierarchy of TFBS resulting in three phenotypic groups: group 1 TFBS mutations resulted in EVGR activation at the expense of LVGR expression and fast replication rates, similar to the findings for clinical *rr*-NCCR variants. Group 2 TFBS mutations showed an intermediate phenotype with respect to EVGR expression and replication rates, whereas group 3 TFBS mutations reduced both EVGR and LVGR expression and permitted little or no viral replication (32).

Among the various factors, Sp1 appeared to have an outstanding role, since the BKPyV NCCR contained at least three Sp1 sites, two of which appeared to have fundamentally different effects on BKPyV expression and replication. Whereas inactivation of the *SP1-4* site proximal to the LVGR resulted in strong EVGR expression, inactivation of the *SP1-2* site proximal to the EVGR abrogated both EVGR and LVGR expression. This effect of the inactivating *sp1-2* mutant also overrode group 1 activation, including the one seen with the *sp1-4* mutant, suggesting a key role of Sp1 in the functionality of the NCCR as a whole.

In the present study, we investigated the role of these two key Sp1 sites and their relationship to core promoter elements in greater detail. Our results demonstrate that the BKPyV NCCR is fine-tuned by an imperfect symmetry of two core promoters with *TATA* and *TATA*-like elements at the EVGR and the LVGR, respectively. The bidirectional EVGR and LVGR expression is governed by the host cell transcription factor Sp1 and involves both DNA strand directionality and binding affinity. The data indicate that the BKPyV NCCR is poised toward EVGR expression, which can be readily discharged by a simple switch in Sp1 binding. The large T antigen (LTag), which is the major EVGR protein product, then drives genome replication and subsequent LVGR expression independently of Sp1 and TATA-like elements, thereby being able to overcome a range of deletions seen in the LVGR promoter of natural BKPyV variants. We discuss the implications of our results as a model for the novel human PyVs and the control of bidirectional gene expression.

MATERIALS AND METHODS

Cell culture. Human primary renal proximal tubule epithelial cells (RPTECs; PCS-400-010; ATCC, Manassas, VA, USA) were grown in epithelial cell medium (EpiCM; catalog number 4101; ScienCell Research Laboratory, Carlsbad, CA, USA) supplemented with epithelial cell growth supplement (EpiCGS; catalog number 4152; ScienCell Research Laboratory, Carlsbad, CA, USA) and 2% fetal bovine serum (FBS; catalog number 0010; ScienCell Research Laboratory, Carlsbad, CA, USA). HEK293 cells (CRL1573; ATCC, Manassas, VA, USA), HEK293T cells (CRL3216; ATCC, Manassas, VA, USA), and HEK293TT cells (catalog number 0508000; NCI-Frederick Repository Service, Frederick, MD, USA) were propagated in Dulbecco's modified Eagle's medium (DMEM), high-glucose formulation (DMEM-H; catalog number D5671; Sigma-Aldrich, St. Louis, MO, USA), containing 8% FBS (catalog number S0113; Biochrome AG, Berlin, Germany). Medium for cultivation of HEK293TT cells was supplemented with 400 μg/ml hygromycin B (Calbiochem, San Diego, CA, USA). COS-7 cells (CRL1651; ATCC, Manassas, VA, USA) were grown in DMEM-H, containing 5% FBS. All cultures were supplemented with 2 mM L-glutamine (catalog number K0302; Biochrome AG, Berlin, Germany).

FACS-based, bidirectional reporter assay. The bidirectional reporter construct used to measure expression of the early (red fluorescent protein [RFP]) and late (green fluorescent protein [GFP]) virus gene regions has been described before (31). Furthermore, we described a fluorescence-activated cell sorter (FACS)-based quantification method for this reporter construct (32). While the use of flow cytometry for the reporter assay enables a convenient means of visualization of expression, at the same time, accurate and reproducible quantification is achieved. In particular, the gating for the same number of transfected cells from each sample as a method of normalization overcomes the problem of different transfection efficiencies between different plasmid preparations, mutants, or cell lines. HEK293, HEK293T, or HEK293TT cells were seeded in 12-well plates and transfected at 70 to 80% confluence with the Lipofectamine 2000 reagent (catalog number 11668-019; Invitrogen, Carlsbad, CA) at a ratio of 3:1 (3 μl reagent, 1 μg plasmid DNA) in Opti-MEM medium (Gibco, Grand Island, NY, USA) according to the manufacturer's instructions. Medium was replaced with DMEM-H–10% FBS on the next morning. At 48 h posttransfection, the cells were rinsed once with phosphate-buffered saline (PBS)–2.5 mM EDTA and then detached, suspended, and transferred to 5-ml polystyrene round-bottom FACS tubes (BD, Franklin Lakes, NJ, USA) with DMEM (without phenol red) and 1% fetal calf serum (FCS). Directly before each measurement, DAPI (4′,6-diamidino-2-phenylindole; catalog number D8417; Sigma-Aldrich, St. Louis, MO, USA) was added (final concentration, 1 ng/ml) as a dead cell marker, and the cells were resuspended. FACS measurements were carried out on a Fortessa cytometer (BD, Franklin Lakes, NJ, USA). For GFP, excitation was at 488 nm (blue laser) and emission was at 530/30 nm; for RFP, excitation was at 561 nm (yellow-green laser) and emission was at 586/15 nm; and for DAPI, excitation was at 405 nm (violet laser) emission was at 450/50 nm. In order to calculate the weighted mean fluorescence intensity (MFI) for red (early) and green (late) expression, the cell number (*N*) and mean fluorescence (*I*) of quadrants Q1 (red cells), Q2 (red and green cells), and

Q4 (green cells) were inserted into the following formulas: MFI (red) = $[(N_{Q1} \cdot I_{Q1}) + (N_{Q2} \cdot I_{Q2})]/(N_{Q1} + N_{Q2} + N_{Q4})$ and MFI (green) = $[(N_{Q2} \cdot I_{Q2}) + (N_{Q4} \cdot I_{Q4})]/(N_{Q1} + N_{Q2} + N_{Q4})$.

The early expression of all constructs was normalized to that of the Dunlop strain (set as 100%), and late expression was normalized to that of the archetype *ww(1.4)* strain (set as 100%). The mean values and standard deviations were calculated with GraphPad Prism software (version 6 for Mac OS).

EMSA. Electrophoretic mobility shift assay (EMSA) analysis of TFBS was performed as described previously (42). Briefly, 10 μg of nuclear extract from RPTECs was mixed with binding buffer (20 mM PIPES [piperazine-*N*,*N*′-bis(2-ethanesulfonic acid)], pH 6.8; 50 mM NaCl; 1 mM dithiothreitol [DTT]; 0.25 mg/ml bovine serum albumin; 100 μM $ZnSO_4$; 0.05% NP-40; 4% Ficoll) and approximately 5 fmol of ^{32}P-labeled, duplexed oligonucleotide in a final volume of 20 μl. The oligonucleotides with the *SP1-2* and *SP1-4* binding sites were described before (32). Nuclear extracts were prepared as described by Schreiber et al. (43) by scraping cells with ice-cold PBS off 10-cm dishes and collecting them in Eppendorf tubes. After centrifugation they were swollen in hypotonic buffer (10 mM HEPES, pH 7.9; 10 mM KCl; 0.1 mM EDTA; 2.5 mM DTT) and lysed by addition of NP-40 (final concentration, 0.5%). After centrifugation of the nuclei, proteins were extracted in nuclear extract buffer (20 mM HEPES, pH 7.9; 25% glycerol; 400 mM NaCl; 1 mM EDTA; 2.5 mM DTT). Competing, unlabeled oligonucleotides were used as described in the legend to Fig. 2. After incubating the mixture on ice for 30 min, samples were separated on a native 4% polyacrylamide gel with 0.25× TBE (Tris-borate-EDTA) as the running buffer. The detection of β decay was carried out with a Fujifilm FLA-7000 image plate reader.

ChIP. Chromatin immunoprecipitation (ChIP) was performed according to the manufacturer's instructions (Magnify chromatin immunoprecipitation system; Life Technologies/Thermo Fisher Science). Briefly, HEK293TT cells were transfected in 10-cm dishes with the BKPyV *ww* reporter construct and incubated for 48 h. The chromatin was cross-linked directly in the dish for 10 min with 10% formaldehyde. The cross-linking was quenched with 0.5 M glycine solution. Cells were harvested with ice-cold PBS, collected by centrifugation, and lysed. For the chromatin fragmentation, lysates were treated with micrococcal nuclease (20 min at 37°C; NEB) and sonicated with 30 cycles of 30-s pulses and 30 s of cooling in between each pulse (Sonicator Q700; QSonica). The ChIP antibodies were Sp1 D4C3 (ChIP-validated rabbit monoclonal antibody; Cell Signaling) (44) and a rabbit control antibody included with the kit. For PCR analysis by HotStar *Taq* (Qiagen) and a quantitative PCR (qPCR) SYBR green assay (PowerUp SYBR green; Thermo Fisher Scientific), the following primers were validated and used: primer NCCR-f (TCAGAAAAAGCCTCCACACC) and primer NCCR-r (CTTTGGCCAGTTTCCACTTC). qPCR data were analyzed by the $\Delta\Delta C_T$ threshold cycle (C_T) method as described before using *ww*-NCCR input and background control ChIP for normalization.

5′ RACE analysis of early and late transcripts. To characterize the 5′ ends of viral transcripts and, hence, to identify the transcription start sites, we performed 5′ rapid amplification of cDNA ends (5′-RACE) procedures according to the manufacturer's instructions (5′ RACE system for rapid amplification of cDNA ends; Invitrogen/Thermo Fisher Scientific). Briefly, we isolated total RNA from RPTECs (RNeasy minikit; Qiagen) which were infected with either archetype virus or the Dunlop strain and reverse transcribed the mRNA by using a mixture of random hexamers and oligo(dT) primers (SuperScript II; Invitrogen/Thermo Fisher Scientific). The resulting cDNAs of viral origin were specifically amplified by using the following gene-specific primers (GSP): sTag-GSP1 (5′-AAGGGCAGTGCACAGAAG-3′), sTag/LTag (5′-AATCAGGCTGATGAGCTACC-3′), LTag-GSP1 (5′-TCTTCAGGCTCTTCTGG-3′), LTag-GSP2 (5′-TGGTGTKGAGTGTTGAGA-3′), agno-GSP1 (5′-TACAGCAGGTAAAGCAGTG-3′), agno-GSP2 (5′-TCCCGTCTACACTGTCTT-3′), VP2-GSP1 (5′-TGATAGAGGCCTACAGTG-3′), VP1-GSP1 (5′-AAGGCCTCTCCACTGTTG-3′), and VP1-GSP2 (5′-CCAGCACTCAACTGGATAAG-3′).

***In silico* analysis of the BKPyV NCCR.** To search for potential transcription factor binding sites within the BKPyV NCCR and to evaluate the effect of mutations on the binding sites, we used the MatInspector software tool (Genomatix, Munich, Germany). Additionally, to analyze the presence of core promoter elements, we used the search tool ElemeNT (http://lifefaculty.biu.ac.il/gershon-tamar/index.php/resources) (45).

Transfection of BKPyV genome into COS-7 cells and infection of RPTECs. Transfection of BKPyV genomes into cells was initiated by cutting BKPyV plasmid DNA with BamHI and religating the diluted DNA as described previously (46). Transfection of religated BKPyV genomic DNA into COS-7 cells was performed when the cells were at 90 to 95% confluence in 6-well plates using Lipofectamine 2000 (catalog number 11668-019; Invitrogen, Carlsbad, CA, USA) at a reagent/DNA ratio of 3:1 according to the manufacturer's instructions. At 6 h after transfection, the medium was replaced with DMEM-H containing 5% FCS. At 7 days post-transfection, COS-7 cell supernatants were harvested and the number of DNase-protected genomic equivalents (GEq) per milliliter was quantified by quantitative reverse transcription-PCR (see the next paragraph). For the infection of RPTECs with COS-7 cell supernatants, cells were seeded at 75,000 cells per well of a 24-well plate in 0.5 ml of supplemented EpiCM. After 24 h, when the cells were at a confluence of approximately 50%, RPTECs were exposed to 200 μl of the corresponding virus preparations adjusted to 2×10^7 GEq for each construct at 37°C for 2 h, followed by removal, washing, and replacement with EpiCM supplemented with 0.5% FCS.

BKPyV viral loads by real-time PCR. BKPyV loads were quantified after DNA extraction from 200-μl cell culture supernatants with a QIAamp DNA blood minikit (Qiagen, Hombrechtikon, Switzerland). The real-time PCR protocol for detection of BKPyV DNA samples targets the BKPyV large T antigen-coding sequence and has been described before (47).

Western blotting. Lysis of HEK293, HEK293T, and HEK293TT cells was done in NP-40 buffer (50 mM Tris-HCl, pH 8.0; 150 mM NaCl; 1% NP-40; 1× protease inhibitors [Roche]). Equal volumes of cell lysates, adjusted to 10 μg total protein, were separated by SDS-PAGE and electrotransferred onto a 0.45-μm-pore-size polyvinylidene difluoride membrane (catalog number IPFL00010; Millipore/Merck, Darmstadt, Germany). The membrane was blocked with Odyssey blocking buffer (catalog number 927-40000; LI-COR, Lincoln, NE, USA) diluted 1:2 in Tris-buffered saline (TBS). The membrane was then incubated with the following primary antibodies: polyclonal rabbit anti-LTag (1:10,000; C. H. Rinaldo, University of North Norway, Tromsø, Norway) and monoclonal mouse antiactin (1:5,000; Abcam, Cambridge, England) in Odyssey blocking buffer. After washing 5 times with TBS–0.1% Tween 20, the membrane was incubated in the following secondary antibodies: goat anti-rabbit immunoglobulin antibody conjugated with IRDye 800CW (1:10,000; catalog number 926-32211; LI-COR) and donkey anti-mouse immunoglobulin antibody conjugated with Alexa Fluor 680 (1:15,000; catalog number A10038; Invitrogen). Protein detection and quantification were done with a LI-COR Odyssey CLx system.

RESULTS

Swapping of Sp1 sites within the NCCR causes rearrangement-like activation of EVGR expression. To measure the early and late promoter activity of the BKPyV NCCR, we transfected a bidirectional reporter construct into human embryonic kidney HEK293 cells, which allows monitoring of EVGR expression by the red fluorescence of DsRed and monitoring of LVGR expression by the green fluorescence of enhanced green fluorescent protein (31, 32). Using this reporter assay, the strain Dunlop NCCR, as a representative of *rr*-NCCRs bearing duplications and deletions, showed strong EVGR expression (red) but only a low level of LVGR expression (green) (Fig. 1A). The Dunlop NCCR contains an intact

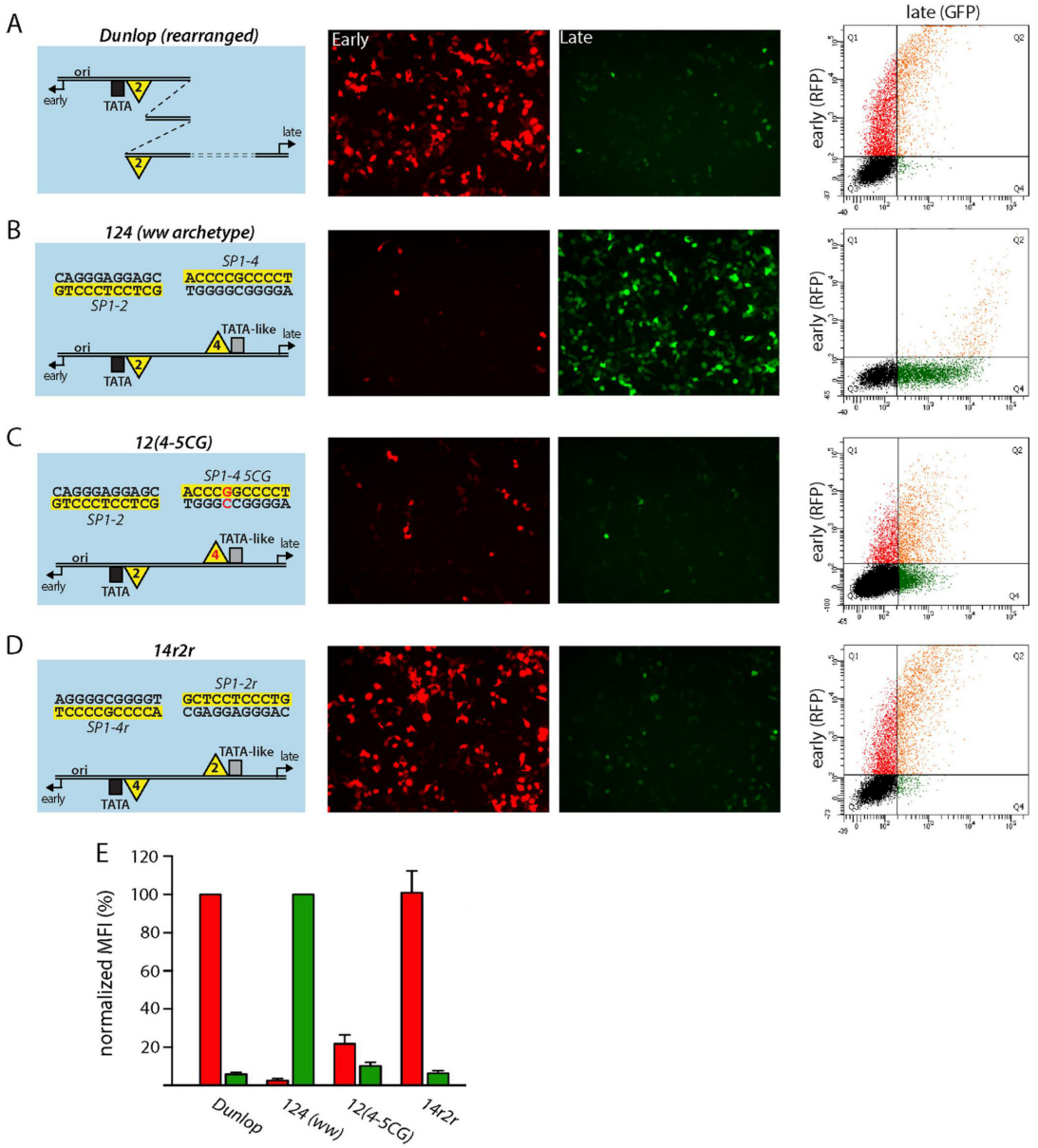


FIG 1 Swapping of *SP1-2* and *SP1-4* in the BKPyV NCCR affects early and late promoter activity similarly to major rearrangements. (A to D) (Left) Schematic representation of NCCRs tested with the FACS-based, bidirectional reporter assay. All NCCRs were cloned into the previously described reporter vector pHRG, in which RFP serves as a marker for early expression and GFP serves as a marker for late expression (31). The Gardner isolate-derived laboratory strain Dunlop is characterized by triplication of the P-block sequences and deletion of Q- and R-block sequences. The BKPyV archetype NCCR *124* is used as a reference. Horizontal black lines, double-stranded DNA of the NCCR; black box, conserved TATA box; gray box, TATA-like element; yellow triangles labeled 2, *SP1-2* Sp1 transcription factor binding site; yellow triangles labeled 4, *SP1-4* Sp1 transcription factor binding site; black number 4, Sp1-binding site with archetype sequence; red number 4, Sp1-binding site carrying one or more mutations; yellow highlight, the respective Sp1 upper-strand sequence with respect to the direction of early and late transcription (black arrows). (Middle two columns) The indicated NCCR reporter constructs were transfected into HEK293 cells, and fluorescence images for RFP (early) and GFP (late) expression were taken at 2 days posttransfection. (Right) Representative flow cytometry measurements of the indicated NCCRs. *x* axis, GFP fluorescence; *y* axis, RFP fluorescence. For each measurement, we gated for 5,000 transfected (fluorescent) cells in quadrants Q1, Q2, and Q4. (E) For quantification, the normalized MFI (in percent) was calculated as described in Materials and Methods. Early expression was normalized to that of strain Dunlop (red MFI = 100%), and late expression was normalized to that of NCCR *124* (*ww* archetype) (green MFI = 100%). Quantification results are from at least 3 independent replicates.

SP1-2 site upstream of the TATA box but bears two incomplete duplications of the P block, adding one *SP1-2* site at a more distant upstream location, while LVGR proximal sequences, including *SP1-4*, are deleted (Fig. 1A, left, with the yellow triangles labeled 2 indicating *SP1-2*). In contrast, the archetype BKPyV *ww* strain has all three Sp1 sites in the archetype position (*ww*-NCCR *124*) and showed a low level of EVGR expression (red) but strong LVGR expression (green) (Fig. 1B, left, *ww* archetype NCCR *124*, with the yellow triangles labeled 2 and 4 indicating *SP1-2* and *SP1-4*, respectively). As reported previously (32), point mutations, like the naturally occurring *SP1-4* variant 5CG (which has a C-to-G transition at position 5) from a patient with BKPyV-associated hemorrhagic cystitis (48), are sufficient to cause an activation of EVGR expression compared to the archetype *ww*-NCCR, at the expense of LVGR expression [NCCR *12*(*4*-5CG) in Fig. 1C and E]. Given the striking effects of rather subtle Sp1 site mutations with respect to EVGR and LVGR expression (32), we investigated whether or not the different Sp1 sites might have different sequence-encoded properties. We noted that *SP1-2* and *SP1-4* have inverse orientations relative to one another, which is possible because Sp1 sites are composed of nonpalindromic sequences. In a first approach, the early *SP1-2* site and the late *SP1-4* site were simply swapped and placed in reverse orientation to maintain the archetypal strand orientation, yielding NCCR *14r2r* (where the letters *r* indicate inverted *SP1-4* and *SP1-2* sequences, respectively) (Fig. 1D and E; the C-rich sequence of the EVGR-proximal Sp1 site and the C-rich sequence of the LVGR-proximal Sp1 site are highlighted in yellow). Compared to the archetype, NCCR *14r2r* caused a strong increase of EVGR expression and decreased LVGR, similar to the phenotype produced by the *rr*-NCCR of the Dunlop strain (Fig. 1A, D, and E). Quantification using the mean fluorescence intensities (MFI) demonstrated that the increase in EVGR expression corresponded to 80 to 100% of the Dunlop NCCR signals (Fig. 1E). At the same time, the level of LVGR expression was reduced to about 10% of the archetype level. We concluded that maintenance of the Sp1 sites but simple swapping of *SP1-2* and *SP1-4* within the archetypal NCCR was sufficient to invert the bidirectional expression pattern of the archetype promoter/enhancer system.

The LVGR *SP1-4* has a higher *in vitro* Sp1-binding affinity than the EVGR *SP1-2*. Although Sp1 has been shown to bind to both *SP1-2* and *SP1-4 in vitro* (32), there are differences in the respective sequence compositions. While *SP1-4* represents a perfect Sp1 consensus sequence, *SP1-2* carries deviations by A/T base pairs (Fig. 1B to D, sequences indicated in yellow). To examine whether the sequence differences were associated with different binding affinities, an EMSA-based competition assay was employed. ^{32}P-radiolabeled, duplexed oligonucleotides carrying *SP1-2* were incubated with nuclear extract from primary human renal proximal tubule epithelial cells (RPTECs) and mixed with increasing concentrations of unlabeled oligonucleotides containing *SP1-2* or *SP1-4* (Fig. 2A). As shown, competition with the unlabeled *SP1-4* oligonucleotides decreased the EMSA signal more readily than competition with the unlabeled *SP1-2* oligonucleotides: while 10 fmol of the *SP1-4* oligonucleotide was enough to reduce the *SP1-2* signal to about 30% of the signal for the archetype, self-competition required approximately 30 times higher competitor concentrations (Fig. 2B). This finding indicates that the *SP1-4* binding site has a higher affinity for Sp1 binding than *SP1-2*, as expected from comparison of their sequences with the perfect Sp1 consensus sequence. Together, the combined data suggest that the affinity of Sp1 binding to the intact local *SP1*-2 and *SP1-4* sites determined the higher activity of the LVGR promoter in the archetype *ww*-NCCR. To complement the *in vitro* data, a chromatin immunoprecipitation (ChIP) assay was performed in *ww*-NCCR reporter construct-transfected cells (Fig. 2C and D). ChIP with the Sp1 antibody yielded an enrichment of about 1,600-fold compared to that achieved with a control antibody, demonstrating the strong binding of Sp1 to the BKPyV NCCR *in vivo* (Fig. 2E).

Strand inversion of Sp1 sites disrupts the archetype EVGR-LVGR expression balance. To examine the potential contributions of the nonpalindromic Sp1 sites to the directionality of gene expression, *SP1-2* and *SP1-4* were flipped in place, yielding the NCCRs *124r*, *12r4*, and *12r4r*, as well as swapped with respect to their original archetype position (NCCR *142*) (Fig. 3A to D). All NCCR inversion mutants showed impaired LVGR expression with a reduction in the level of expression to about 20% of the archetype level. Of note, all of these constructs except *142* still carried an intact high-affinity *SP1-4* site in the LVGR promoter region, including NCCR *12r4*, in which *SP1-4* was even maintained in its original strand orientation (Fig. 3B and E). At the same time, all inversion mutants showed a moderate level of EVGR activation compared to the archetype, whereby the NCCR *12r4* had the strongest expression in the range of 25% of the level seen for the Dunlop *rr*-NCCR. These observations indicate that the relative Sp1 site orientation was critical for regulating bidirectional expression, in addition to affinity. Importantly, the Sp1 sites were able to convey an orientation-dependent directionality that seemed to be functionally as important as the local Sp1-binding affinity or deletion of *SP1-4*. Interestingly, NCCR *142* yielded a high EVGR expression level similar to that of Dunlop and the initially tested *14r2r* mutant (Fig. 3D and E and 1D and E). We concluded that, when placing the high-affinity *SP1-4* site in the proximity of the EVGR, its orientation relative to the remaining NCCR did not matter anymore. This notion was further corroborated by another set of mutants, where the low-affinity *SP1-2* was replaced with the high-affinity *SP1-4*, such that the EVGR and LVGR proximal positions contained Sp1 sites with equal affinities in all possible combinations of site orientations (Fig. 3F to J, NCCRs *14r4*, *14r4r*, *144*, and *144r*). All of the resulting mutants showed an equally high level of EVGR expression in the range of 80% of that of strain Dunlop, coupled to a profound loss of LVGR expression. Thus, a high-affinity Sp1 site in the proximity of the EVGR provided a strong activation, which in no case permitted rescue of the archetype expression pattern by an Sp1 site in an LVGR of similar sequence, affinity, or orientation (Fig. 3J). Conversely, the data further support the notion that the bidirectional fine-tuning of NCCR expression depends on the modulation of Sp1 affinity at the EVGR site.

Early promoter activity correlates with the Sp1-binding affinity of the early Sp1 site. To examine the role of Sp1 affinity in the critical EVGR position, the mutant *14r4* was used for further study by introducing point mutations that had been identified in patient variants (48) and previously shown by us to lower the Sp1-binding affinity (32). By slightly lowering the affinity of the EVGR-proximal Sp1 site with the mutant *1*(*4r*-10CT)*4* (where 10CT indicates the exchange of C to T in position 10), the expression of EVGR decreased to about half the level seen for mutant *14r4* (or 30 to 50% of that for the Dunlop reference strain) (Fig. 4B

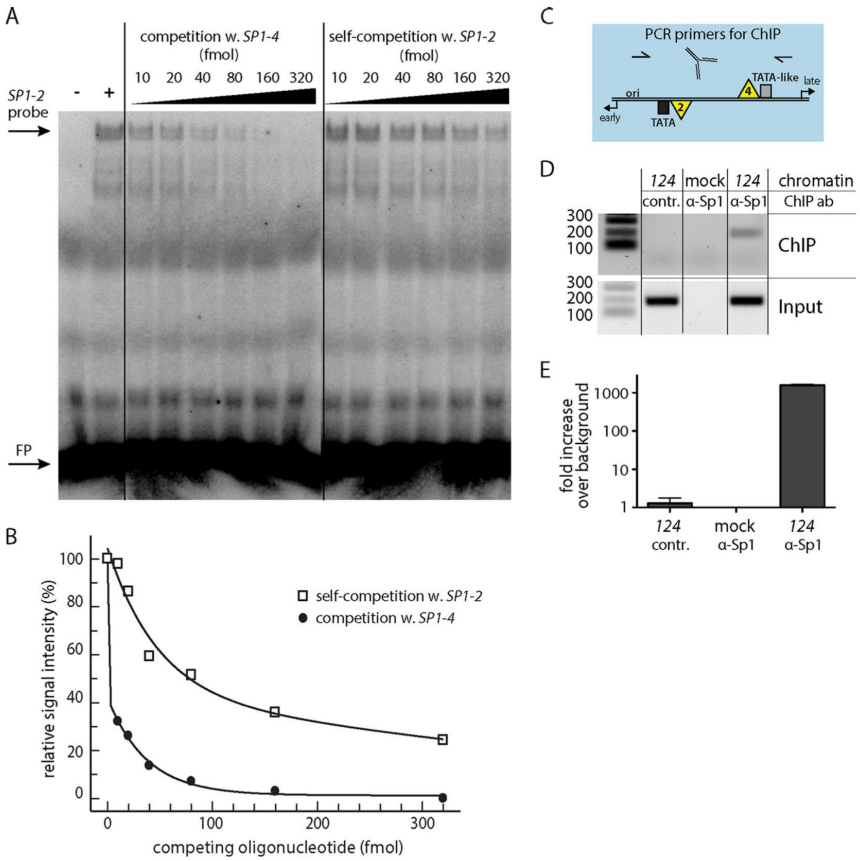


FIG 2 *SP1-2* and *SP1-4* differ in Sp1 affinity in EMSA competition assays. (A) Representative EMSA result after competition of the radiolabeled *SP1-2* probe with (w.) unlabeled *SP1-4* or unlabeled *SP1-2* (self-competition) at the indicated concentrations. Lane −, binding mix without nuclear extract; lane +, binding mix without competition; FP, free probe. (B) Plot of quantitative readout of EMSA signals. *x* axis, concentration of competing oligonucleotide; *y* axis, signal strength relative to that of the uncompeted *SP1-2* signal, set to 100%. (C) Scheme of NCCR with primer binding sites for ChIP assay. (D) Semiquantitative analysis of ChIP assay results by agarose gel electrophoresis. Mock, untransfected HEK293TT cells; contr., ChIP with rabbit negative-control antibody (ab) supplied with the kit. Numbers on the left are denote DNA fragment sizes (in base pairs). (E) Quantification of ChIP assay result by qPCR with SYBR green. Calculation was carried out by $\Delta\Delta C_T$ method (with normalization to the *ww*-NCCR input and the background control).

and F). The intermediate-affinity mutant *1(4r*-5CG*)4* allowed approximately 15% of the level of *14r4* EVGR expression (10% of that of Dunlop), while the double mutant showed only 5% of the level of expression by Dunlop (a >75% reduction compared to that for the *14r4* mutant) (Fig. 4C, D, and F). In contrast to the gradual decrease in the level of EVGR expression, the late promoter activity remained at the same low level (about 20% of the archetype level of expression). Hence, just lowering the affinity of the EVGR-proximal Sp1 site was not sufficient to restore a high level of LVGR expression. The data suggest that, in addition to affinity, particularities of the *SP1-2* sequence might play a role for achieving a high level of LVGR activity.

To confirm the mechanistic relevance of these findings, the NCCR mutants were placed into a BKPyV Dunlop backbone and analyzed in a viral replication assay. To minimize potential effects of EVGR expression and, hence, LTag levels, infectious supernatants were generated in COS-7 cells constitutively expressing simian virus 40 (SV40) LTag, as reported previously (32), and were used to infect RPTECs. Supernatants from RPTECs were then taken at 1, 3, 6, and 9 days postinfection (dpi), and viral loads were quantified by qPCR (Fig. 4G). As expected, the Dunlop strain had the highest replication rate of the tested strains, reaching supernatant viral loads of more than 10^{10} genome equivalents (GEq)/ml at 6 dpi, which corresponded to a more than 200,000-fold increase compared to the input (1 dpi). Importantly, the recombinant virus strains *14r4* and *1(4r*-10CT*)4* were functional and showed accelerated replication compared to that of the archetype virus, reaching levels of replication 1,000-fold that of the archetype virus at 6 dpi and about 40,000-fold that of the archetype virus at 9 dpi, respec-

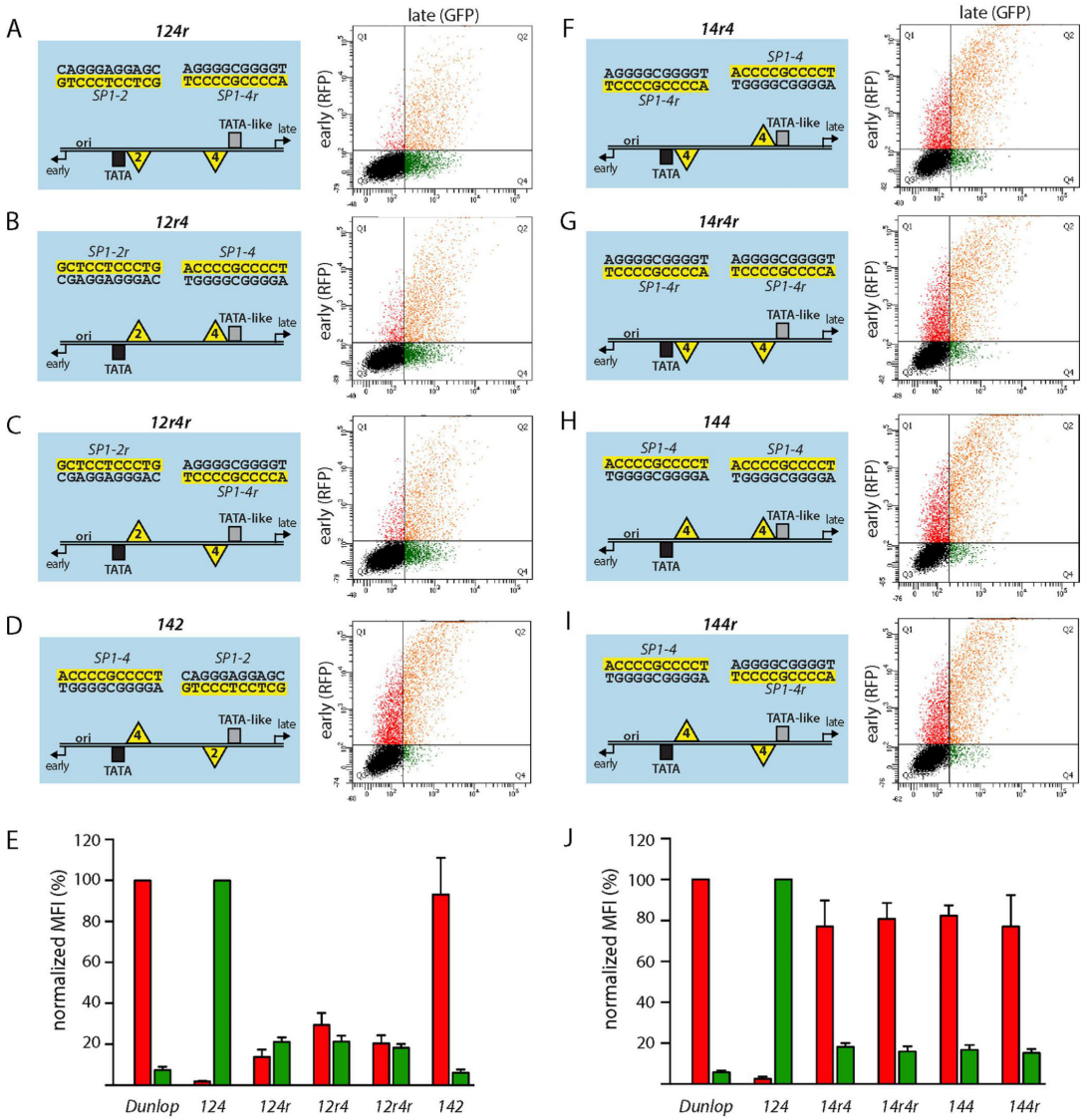


FIG 3 Position, orientation, and affinity of Sp1 sites determine the directionality of EVGR and LVGR expression. (A to D) (Left) Schematic representation of NCCRs tested with the FACS-based, bidirectional reporter assay in HEK293 cells. All mutant NCCRs have the length and architecture of the BKPyV archetype NCCR *124* but carry inverted Sp1 sites, as indicated with the letter *r* (e.g., NCCR *124r* has an inverted *SP1-4* sequence). (Right) Representative flow cytometry measurements are shown next to the schematic representations, which are described in the legend to Fig. 1. (E) Quantification of the constructs listed in panels A to D using the normalized MFI from at least 3 independent replicates, as described in the legend to Fig. 1. (F to I) Schematic representation of NCCRs, where the archetype *SP1-2* was replaced with the sequence of the high-affinity *SP1-4*. (J) Quantification of the constructs listed in panels F to I using the normalized MFI from at least 3 independent replicates, as described in the legend to Fig. 1.

tively. However, the weakened affinity in the *1*(*4r*-10CT)*4* strain seen in the reporter assay did not translate into measurable differences in replication kinetics, suggesting that the EVGR-proximal Sp1 affinity was still strong enough to drive EVGR expression in the context of viral infection in cell culture. In contrast, the viruses carrying the stronger mutations *1*(*4r*-5CG)*4* and *1*(*4r*-5CG/10CT)*4* did not seem to allow rapid viral replication (Fig. 4G). Thus, the data indicate that the Sp1

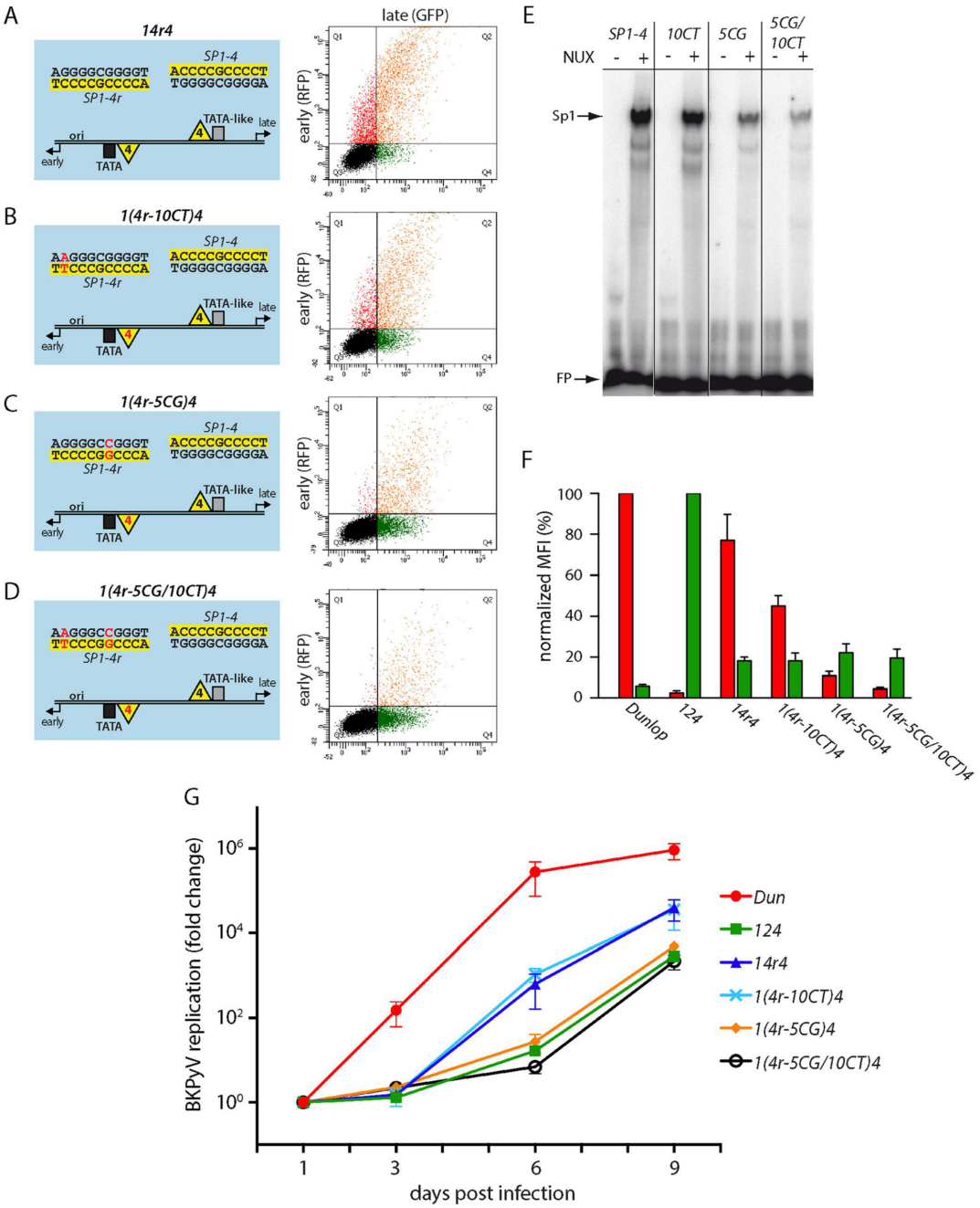


FIG 4 Lowering of the Sp1 affinity in the EVGR-proximal position of NCCR *14r4* reduces EVGR expression and viral replication. (A to D) Schematic representation (left) and representative FACS plot (right) of the indicated NCCRs tested in the bidirectional reporter assay in HEK293 cells. The numbering of the point mutants is relative to the first position of the *SP1-4* site in the archetype orientation [e.g., NCCR *1*(*4r*-10CT)*4* contains the exchange of C to T in position 10, while the late *SP1-4* site is unaltered]. (E) EMSA analysis of binding affinity in the absence and presence of nuclear extracts (NUX) confirms that the inserted point mutations lower the *in vitro* affinity of the respective variants. (F) Quantification of the constructs listed in panels A to D using the normalized MFI from at least 3 independent replicates, as described in the legend to Fig. 1. (G) Mutant BKPyV replication in primary human RPTECs using infectious supernatants generated by transfection of COS-7 cells. The cell culture supernatants were taken at 1, 3, 6, and 9 days postinfection (*x* axis), and DNase-protected viral loads were quantified by qPCR. Results were normalized to the amount of input virus (day 1) and are shown as the fold change (*y* axis). Dun, strain Dunlop.

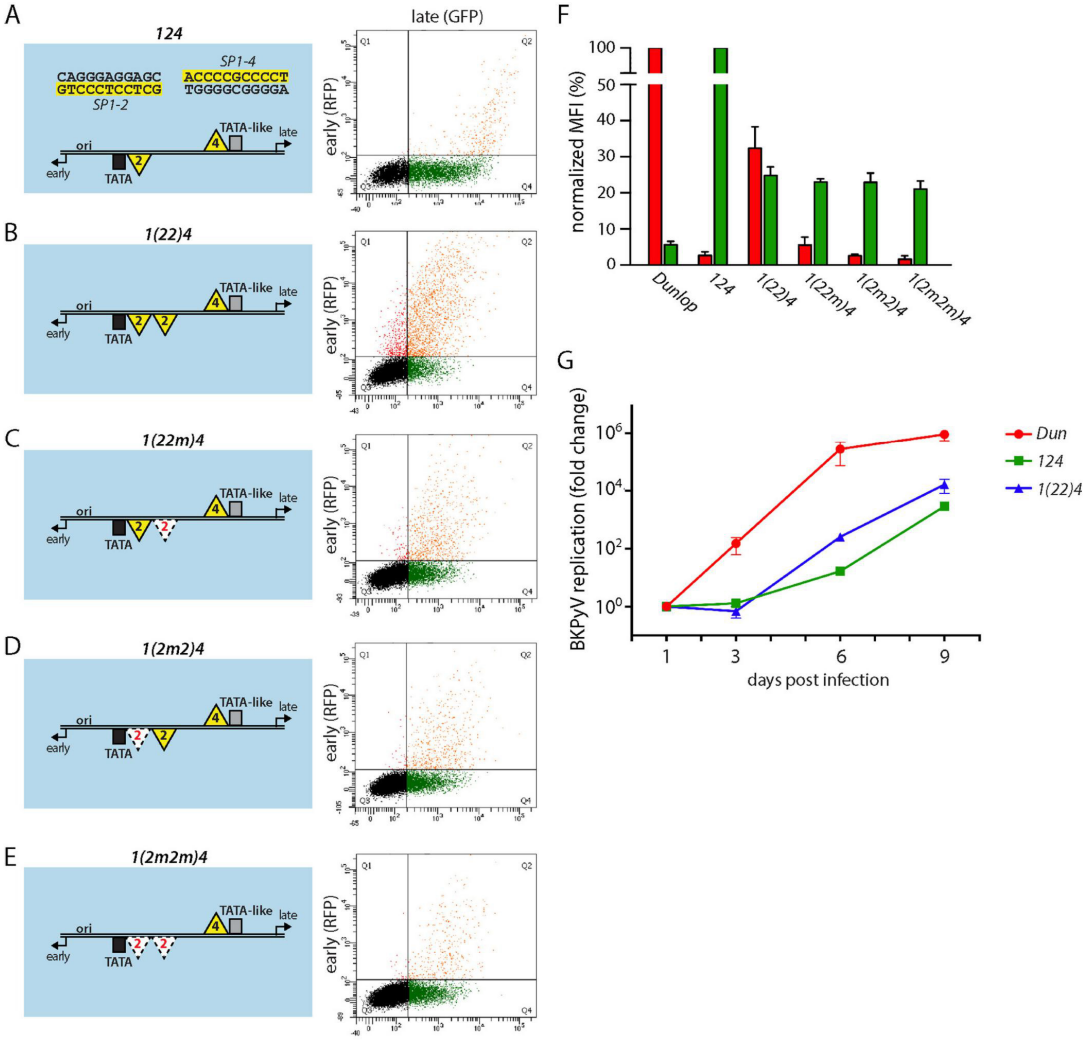


FIG 5 Addition of a second low-affinity *SP1-2* site increases the level of EVGR expression. (A to E) The indicated mutant NCCRs inserted into the bidirectional reporter constructs (left) were transfected into HEK293 cells, and representative FACS plots are shown (right). (F) Quantification of the constructs listed in panels A to E using the normalized MFI from at least 3 independent replicates, as described in the legend to Fig. 1. Note that the *y* axis was divided for better visibility of the low range. (G) Mutant BKPyV replication in primary human RPTECs using infectious supernatants generated by transfection of COS-7 cells. Cell culture supernatant was taken at 1, 3, 6, and 9 days postinfection (*x* axis), and DNase-protected viral loads were quantified by qPCR. Results were normalized to the amount of input virus (day 1) and are shown as the fold change (*y* axis).

affinity of the EVGR-proximal site was also reflected in the level of BKPyV replication, although the infection assay did not achieve the same nuanced resolution observed by the flow cytometry assay interrogating the bidirectional reporter constructs.

EVGR expression can be enhanced by low-affinity Sp1 site duplication. Given the frequently observed duplications in the vicinity of the EVGR of rearranged NCCRs, including one of the NCCRs of the Dunlop strain, we examined expression in another set of NCCR mutants which carried a second *SP1-2* site in the EVGR-proximal position, NCCR *1(22)4* (Fig. 5A and B). This additional binding site was sufficient to increase the level of EVGR expression 15-fold compared to that for the archetype, reaching about 30% of the level for Dunlop (Fig. 5F). To exclude a potential distance or length effect, we also examined inactivated Sp1 sites yielding the NCCRs *1(22m)4*, *1(2m2)4*, and *1(2m2m)4*, in which

the proximal *SP1-2* site, the distal *SP1-2* site, or both *SP1-2* sites were mutated to sequences that no longer bound Sp1 (Fig. 5C, D, and E). These additional NCCR mutants showed reduced EVGR expression compared to the *1(22)4* mutant, in line with the notion that Sp1 binding and not any inserted sequence mediated an elevated level of EVGR expression. Together, these findings confirm that *SP1-2* duplications found in parts of the P block of some patient isolates (31) as well as the Dunlop strain contribute to activation of the EVGR phenotype. Interestingly, all the experimentally designed insertions of the low-affinity *SP1-2* diminished LVGR expression to a similar degree, again suggesting that some additional factors in the NCCR sequence might play a role. In the viral replication assay, recombinant BKPyV bearing NCCR *1(22)4* propagated slightly faster than the archetype NCCR *124* virus, which is in line with the results of the bidirectional reporter assay (Fig. 5G).

Characterizing the early and late core promoter elements of the archetype NCCR. Given the strong effects of Sp1 sites on EVGR and LVGR expression through position, orientation, and affinity, we addressed potential cooperating components in their close proximity. Analysis of the NCCR sequence in the vicinity of the *SP1-2* and the *SP1-4* sites identified a canonical *TATA* box upstream of the EVGR start codon but only potential *TATA*-like elements upstream of the LVGR start codon which deviated from a perfect *TATA* box by several mismatches (49). The actual transcriptional start site (TSS) for the BKPyV NCCR EVGR was derived by 5′ rapid amplification of cDNA ends (5′-RACE) analysis (Fig. 6A), which also revealed a sequence consistent with loci of initiator elements (*Inr*) (Fig. 6A). The positions of the Inr elements allowed the deduction of other core promoter elements (CPEs), which were located within the canonical distances upstream and downstream of the identified TSSs. The web-based search tool ElemeNT (45) confirmed the position of the early core promoter, which is located close to the O-block-P-block boundary and which is defined by the conserved *TATA* box and by a prominent Inr 31 bp downstream (Fig. 6A). In addition, a downstream promoter element (DPE) was identified at the typical distance of 28 to 32 nucleotides downstream from the Inr TSS. When the LVGR promoter region of the archetype NCCR was analyzed, a similar composition with a dominant Inr and potential DPEs was found (Fig. 6B), but in contrast to the EVGR promoter with its conserved TATA box, two potential *TATA*-like elements were located upstream of the Inr/major TSS (Fig. 6B, plot below the sequence). Of note, this major TSS was in line with a previous mapping of the BKPyV late promoter by a primer extension assay (50). One of the *TATA*-like elements had two mismatches compared to the sequence of a perfect *TATA* box but showed the canonical distance of 30 nucleotides to the major TSS (*TATA*-like element 1; Fig. 6B), while the other had three mismatches and was located further upstream (*TATA*-like element 2). An additional potential CPE was identified as a transcription factor IIB (TFIIB) recognition element upstream (*BREu*), which is typically located directly upstream of a *TATA* box. Thereby, the promoter elements of EVGR and LVGR presented an imperfect symmetry of a classical *TATA* and a *TATA*-like promoter. Interestingly, in the case of the LVGR promoter, the *BREu* was adjacent to *TATA*-like element 2 and consisted of a GC-rich consensus sequence which partially overlapped *SP1-4* (Fig. 6B and D). This overlap of the LVGR *BREu* and high-affinity *SP1-4* raised intriguing questions about the functional implications with respect to the regulation of LVGR expression through high-affinity Sp1 binding. It was therefore of interest to first examine deletion-bearing *rr*-NCCRs. As shown in Fig. 6C, the Dunlop *rr*-NCCR is characterized by a deletion of the Q_{35} and R_{63} blocks, which removes the *TATA*-like elements as well as *SP1-4*, and the dominant Inr, thereby disrupting the aforementioned imperfect symmetry of the late and the early sides of the archetype NCCR. The 5′-RACE analysis of the strain Dunlop NCCR revealed an overall dispersion of TSS but a slight overrepresentation of TSS at two potential Inr elements located in a canonical distance to the DPEs of 28 to 33 nucleotides (Fig. 6C, plot below the sequence). The results indicate that the loss of the LVGR CPEs could be partially compensated for by multiple initiator elements, which were found to be located throughout the triplicate P-block sequences and which acted in conjunction with the conserved DPEs of the S block. Together, the data indicate that the basal expression pattern of the archetype BKPyV NCCR involved an imperfect symmetry centering around a *TATA* box and *TATA*-like elements, which are modulated by juxtaposed Sp1-binding sites. Disruption of this basal symmetry by LVGR-proximal deletions was compensated for by the next-best fitting of analogous sequences.

The *TATA* box is essential for both early and late expression. To better understand the contribution of the respective core promoter elements for EVGR and LVGR expression, we designed a set of constructs in which we either removed defined sequence elements or changed them to obtain optimized consensus sequences but, again, without altering the overall length of the archetype NCCR. When we mutated the EVGR *TATA* box (Fig. 7A, *TATA* box-out, where "out" indicates that the CPEs were removed) or replaced it with a *TATA*-like element similar to the one at the LVGR (Fig. 7B, *TATA*-like-in early, where "in" indicates that the CPEs were inserted), not only was EVGR expression almost completely abolished but also LVGR expression was reduced to 23% of that of the archetype NCCR (Fig. 7K). These observations demonstrate the central importance of the *TATA* box not only for EVGR but also for the NCCR as a whole. In contrast, mutating the LVGR *TATA*-like elements individually or together (Fig. 7C to E, *TATA*-like-out late 1, *TATA*-like-out late 2, and *TATA*-like-out late 1 + 2) caused an increase in the level of EVGR expression to 10 to 15% of that of the archetype, while the level of LVGR expression was reduced to 20 to 25% of that of the archetype (Fig. 7K). Thus, effective LVGR expression from the late core promoter seemed to be equally dependent on both TATA-like elements, although the *TATA*-like element 1 sequence better matched a perfect *TATA* box and was situated within the canonical distances to the major TSS (Fig. 6B). When we mutated *SP1-4*, in addition to TATA-like element 1 (Fig. 7F, *TATA*-like-out late 1/*sp1-4*), LVGR expression went down by another 10%, suggesting a cooperative effect of the *TATA*-like element and the Sp1-binding site on the late promoter activity (Fig. 7K). Interestingly, changing the *TATA*-like elements individually or together to perfect TATA boxes did not increase the level of LVGR expression compared to that of the archetype NCCR (Fig. 7G to I, *TATA* box-in late 1, *TATA* box-in late 2, and *TATA* box-in late 1 + 2). Instead, the distance to the major Inr seemed to gain importance when a perfect *TATA* box sequence was inserted into the position of *TATA*-like element 1, as the level of LVGR expression then reached 35% of that of the archetype, while that of the corresponding construct with the *TATA* box-in late 2 reached only 23% of that of the archetype (Fig. 7K). Hence, the insertion of consensus *TATA* boxes in both positions diminished the LVGR promoter activity compared to that of the

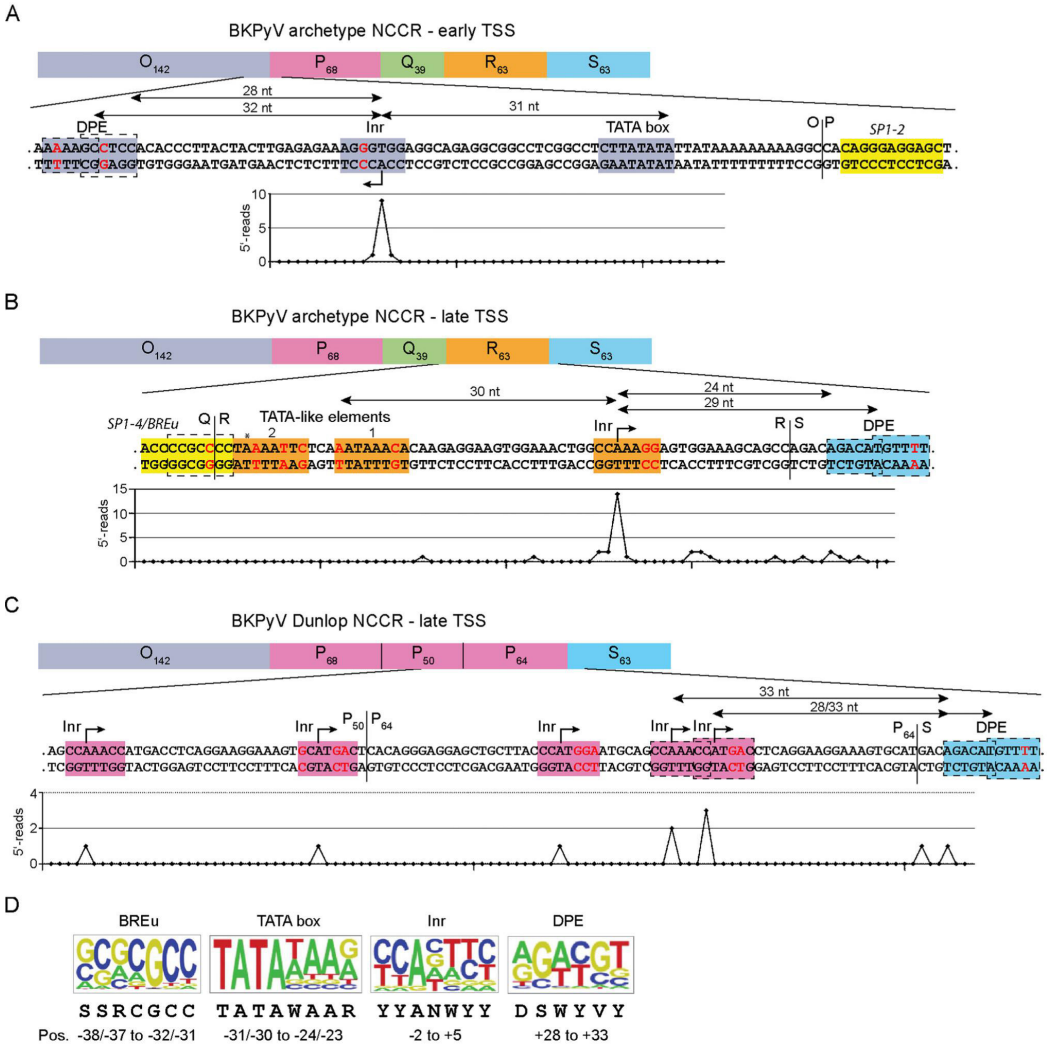


FIG 6 Experimental and computational core promoter analysis indicates the imperfect symmetry of the BKPyV archetype NCCR. The arbitrary sequences of the O, P, Q, R, and S blocks of the NCCR are shown on top. Selected nucleotide sequences enlarged with core promoter elements (CPE) are highlighted in the color of the respective sequence block, except for Sp1 sites, which are highlighted in yellow. Inr, initiator element; DPE, downstream promoter element (sequence in the dashed box highlighted in blue); *BREu*, TFIIB recognition element upstream, overlapping (dashed box) with *SP1-4* highlighted in yellow; red nucleotides, mismatches compared to the CPE consensus sequence; vertical black lines, sequence block boundaries; horizontal black lines, distance (in nucleotides [nt]) from CPEs; dashed boxes, overlapping CPEs. The plot below the sequence shows 5′-RACE data, used to determine the transcription start site (TSS), with each point corresponding to the base pair position in the sequence presented above and the *y* axis showing the number of 5′ reads for the respective position. (A) Analysis of early core promoter located at the O block-P block boundary. (B) Analysis of the archetype late core promoter in the Q, R, and S blocks. *, position with the known polymorphism A/G, with G introducing an additional mismatch into the *TATA*-like element. (C) Analysis of Dunlop late core promoter with deletions removing the Q and R blocks. (D) Relevant CPEs with sequence logos, the consensus sequence in the IUPAC ambiguity code, and the typical position (Pos.) relative to the TSS (position +1) (45). R, A/G; W, A/T; S, C/G; Y, C/T; V, A/C/G; N, any nucleotide.

archetype, but the canonical positioning of the *TATA* box could partlycompensate for this adverse effect.The single insertions showed an activation of EVGR expression of about 5-fold compared to that of the archetype (11% of that of the Dunlop strain), and the double mutant *TATA* box-in late 1 + 2 showed EVGR expression in the range of 5% of that of Dunlop.

As mentioned above, the high-affinity *SP1-4* overlaps a potential *BREu*. Both share a GC-rich consensus sequence, and the

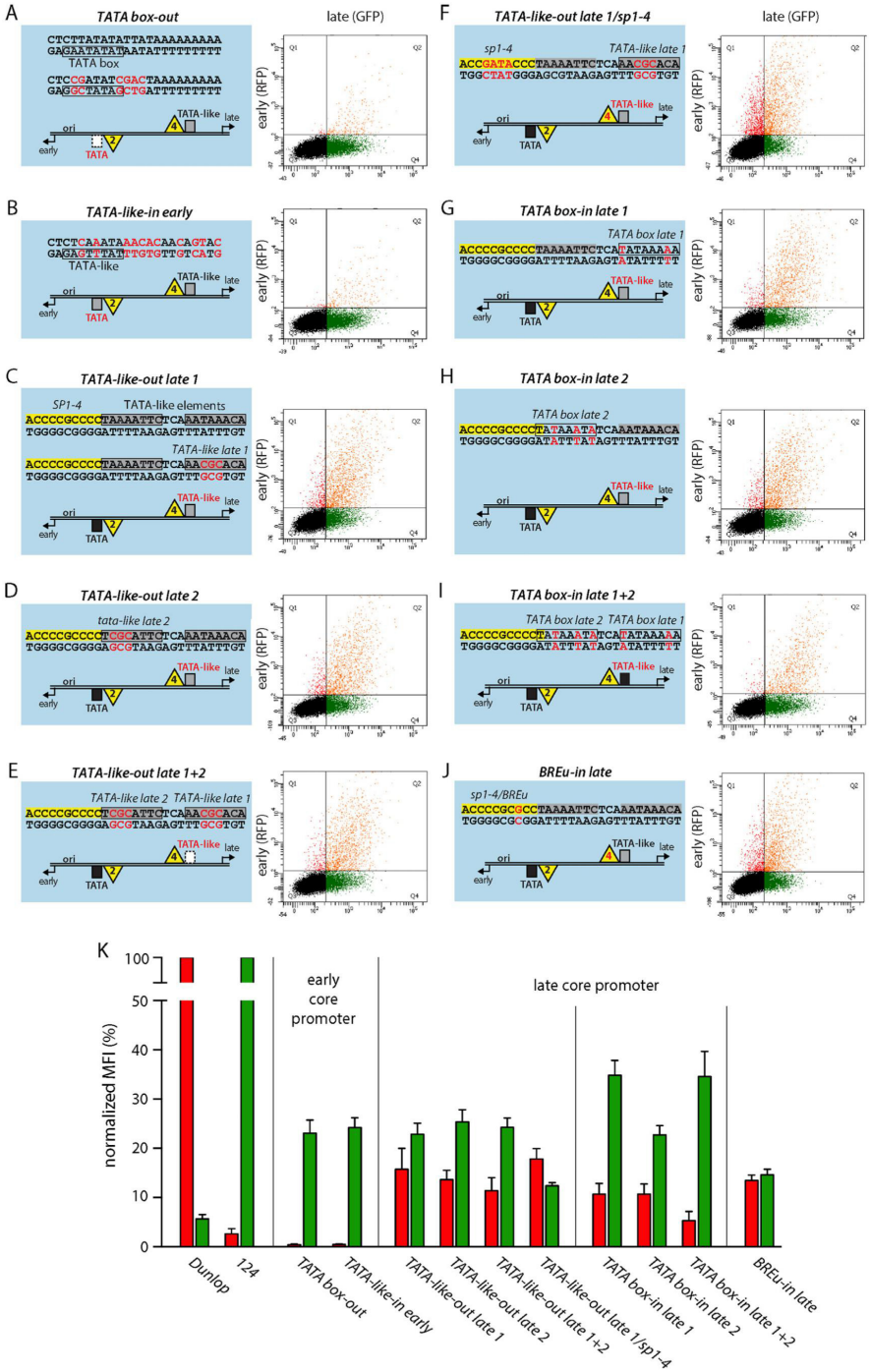


FIG 7 Mutating the EVGR TATA box or the LVGR *TATA*-like elements increases early expression. (A to J) (Left) The indicated mutant NCCRs inserted into the bidirectional reporter constructs were transfected into HEK293 cells. Red letters, mutations compared to the archetype sequence. CPEs were either removed (indicated "out") or inserted (indicated "in"). (Right) Representative FACS plots. (K) Quantification of the constructs listed in panels A to J using the normalized MFI from at least 3 independent replicates, as described in the legend to Fig. 1.

SP1-4 site carries only one mismatch compared to the sequence of a perfect *BREu* site (Fig. 6D). Therefore, a point mutation was introduced to create a perfect consensus *BREu* site (Fig. 7J, BREu-in late), which at the same time inactivated Sp1 binding, as shown for the previously analyzed *SP1-4* point mutations (Fig. 1). The results show that introduction of a perfect BREu consensus site did not increase the level of LVGR expression but, rather, reduced it to a level of about 15% of that of the archetype. At the same time, EVGR expression was increased only moderately to about 14% of that of Dunlop or 7-fold of that of the archetype. This finding indicates that the *SP1-4* site is more important than the potential BREu for driving LVGR expression.

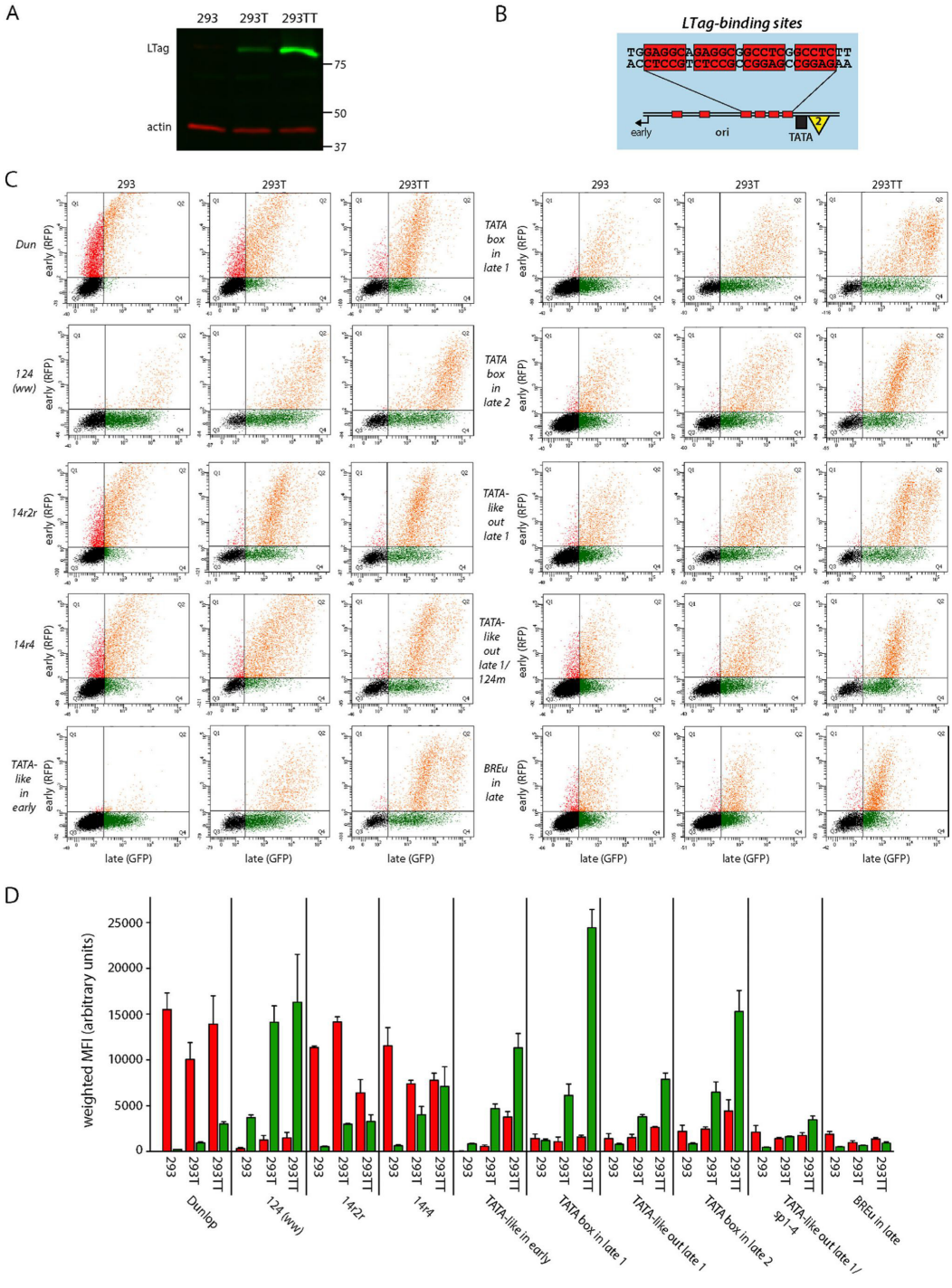
LTag can drive LVGR expression independently of archetypal late CPEs. LTag is a major product of EVGR expression and an important regulator of the viral life cycle with respect to viral DNA replication and LVGR expression. We therefore compared the effects of critical NCCR mutants in HEK293, HEK293T, and HEK293TT cells, which produce no, small, or large amounts of SV40 LTag in *trans*, respectively (Fig. 8A). One essential function of LTag for the viral life cycle is to bind to the viral origin of replication (in the O block) and to facilitate, together with cellular factors, the replication of the viral genome (Fig. 8B). Interestingly, for almost all NCCR mutants tested, LTag was able to boost LVGR expression in a concentration-dependent manner. Despite variations in the absolute fluorescence intensities (Fig. 8C and D), most NCCR reporter constructs showed similar relative increases in the levels of LVGR expression in HEK293T and HEK293TT cells compared to the levels in their HEK293 cell counterparts (indicated by a dashed line, set to 1 for early and late fluorescence and each construct in HEK293 cells; Fig. 8E). The Dunlop strain, for example, which is characterized by a very low level of late expression in HEK293 cells, expressed 4 times more LVGR in HEK293T cells and 13 times more LVGR in HEK293TT cells than in HEK293 cells. At the same time, the level of EVGR expression of Dunlop stayed at the same level in the three cell lines. Since the LTag-mediated increase of LVGR expression was also observed with the Dunlop NCCR, which carries a large deletion of the archetypal late promoter elements (Fig. 6C), we concluded that LTag rendered LVGR expression independent of the expression of most archetypal late CPEs. This is in line with the finding that the *TATA*-like elements are dispensable, as observed with the construct *TATA*-like-out late 1, which showed an 8-fold increase in HEK293T cells and an 18-fold increase in HEK293TT cells compared with that in HEK293 cells. On the other hand, the highest absolute intensity (24,400 arbitrary units; Fig. 8D) and the highest relative increase (20-fold; Fig. 8E) in HEK293TT cells were recorded with the *TATA* box-in late 1 construct, which carried a perfect *TATA* box in a canonical distance to the major late Inr (Fig. 7G). We concluded that although the activating effect of LTag on LVGR expression could occur without the archetypal late promoter, there still seemed to be a benefit from the presence of a perfect *TATA* box and the high-affinity *SP1-4* site. The latter becomes visible when the levels of expression in the constructs NCCR *14r2r* and *14r4*, both of which carry the high-affinity *SP1-4* site in the early promoter but differ in the affinity of the late Sp1 site, are compared. While swapping of NCCR *14r2r* increased the level of LVGR expression by only 6-fold in HEK293TT cells, LVGR expression of NCCR *14r4* was boosted by 11-fold. However, LVGR expression of the archetype NCCR *124* was already high in HEK293 cells, which explains why the absolute and relative increases in HEK293T and HEK293TT cells were only moderate (4-fold). Intriguingly, LTag activation of LVGR expression also worked in the absence of the otherwise essential and conserved *TATA* box when it was compared in the EVGR mutants *TATA* box-out (not shown) and *TATA*-like-in early (Fig. 8C to E). The only exception to this LTag-mediated increase in LVGR expression was in the mutant NCCR *BREu*-in late, where a point mutation in the high-affinity *SP1-4* site created a perfect BREu site. Here, the increase in the level of LVGR expression was only 1.3-fold (HEK293T cells) and 1.8-fold (HEK293TT cells). Since both the *SP1-4* mutant (not shown) and the double mutant *TATA*-like-out late 1/*sp1-4* (Fig. 8E) were activated by LTag, we can only speculate at this point that the binding of TFIIB to the late promoter might pose an obstacle to LTag-mediated LVGR expression. Finally, we observed that EVGR expression remained similar for most constructs in this comparative transfection study, suggesting that the previously reported negative feedback of LTag on EVGR expression might be offset by increasing reporter construct replication. Notable exceptions to this observation were two constructs which had very low levels of EVGR expression in HEK293 cells, namely, the archetype NCCR *124* and the *TATA* box mutant *TATA*-like-in early. Here, few changes in early expression occurred, and these were probably attributable to plasmid copy number changes or an only subtle stimulation of the early promoter by LTag, which resulted in large relative increases in the expression levels.

DISCUSSION

The NCCR determines within only 400 bp essential functions of PyV biology, including viral persistence and the appropriate timing of virus gene expression, genome replication, and virion assembly. Whereas other similarly sized viruses, like members of the *Papillomaviridae* or proviral copies of the *Retroviridae*, use unidirectional promoter organization, PyVs have effectively evolved a complex bidirectional expression modality to sequentially role out a strand- and orientation-specific set of virus gene products from their NCCRs. Notably, these steps of the PyV life cycle are host specific and, within a given host, restricted to certain cells, arguing that cellular differentiation and activation states must be readily sensed and interpreted by PyV-specific NCCRs.

As the archetype NCCR of BKPyV is stably maintained in the general population, we were intrigued by the rapid real-time emergence of BKPyV variants with rearranged NCCRs in immunosuppressed kidney transplant patients and their functional link to increased EVGR expression, accelerated viral replication, higher blood viral loads, and more advanced pathology (31). Given the diverse array of affected TFBS in the *rr*-NCCRs, we recently introduced defined point mutations that excluded the potential effects of overall length and architecture on NCCR function and identified three phenotypic groups of TFBS (32): group 1 and group 2 mutations caused a strong and an intermediate level of activation of EVGR expression and viral replication, respectively, similar to that caused by natural BKPyV NCCR variants associated with disease, whereas group 3 mutations permitted only reduced or no NCCR function (32). Importantly, Sp1 emerged as a key regulatory factor, being able to mediate either extreme of the group 1 mutant (e.g., *SP1-4*) or the group 3 mutant (e.g., *SP1-2*) phenotype.

In the present study, we demonstrate that Sp1 can exert these differential effects as a result of the position, orientation, and affinity of two Sp1 TFBS located on the LVGR- and the EVGR-proximal sides of the archetype NCCR. Swapping of the low-af-


A
293
293T
293TT
LTag
actin
75
50
37
B
LTag-binding sites
TGGAGGCAGAGGCGGCCTCGGCCTCTT
ACCTCCGTCTCCGCCGGAGCCGGAGAA
early
ori
TATA
C
293
293T
293TT
Dun
124 (ww)
14r2r
14r4
TATA-like in early
TATA box in late 1
TATA box in late 2
TATA-like out late 1
TATA-like out late 1/ 124m
BREu in late
early (RFP)
late (GFP)
D
weighted MFI (arbitrary units)
25000
20000
15000
10000
5000
0
Dunlop
124 (ww)
14r2r
14r4
TATA-like in early
TATA box in late 1
TATA-like out late 1
TATA box in late 2
TATA-like out late 1/ sp1-4
BREu in late

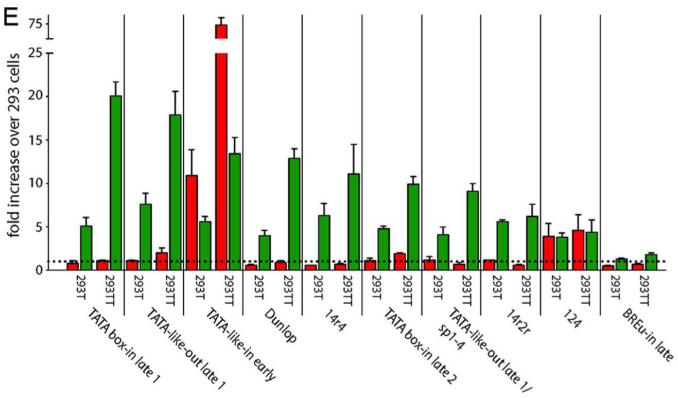


FIG 8 LTag expression in *trans* increases the level of LVGR expression in a concentration-dependent manner and can drive late expression independently of core promoter elements. (A) Western blot showing the different LTag expression levels by HEK293, HEK293T, and HEK293TT cells and the level of actin expression as a loading control. Numbers on the right are molecular masses (in kilodaltons). (B) Schematic view of LTag-binding sites (sequences with a red background) in the O block of the BKPyV NCCR. (C) Representative flow cytometry of HEK293, HEK293T, and HEK293TT cells transfected with the indicated BKPyV NCCR bidirectional reporter constructs. (D) For quantification of EVGR and LVGR expression in the transfected cells, the normalized MFI (in percent) was calculated as described in Materials and Methods. The level of EVGR expression was normalized to that of strain Dunlop (red MFI = 100%), and the level of LVGR expression was normalized to that of NCCR *124* (*ww* archetype) (green MFI = 100%). Quantification was based on at least 3 independent replicates. (E) Relative change in the levels of EVGR and LVGR expression by the indicated reporter constructs in HEK293T and HEK293TT cells using HEK293 cells as a reference (for which the level of expression was set equal to 1 and is indicated by a horizontal dashed line).

finity *SP1-2* site and high-affinity *SP1-4* site (e.g., NCCR *14r2r*), introducing affinity-lowering mutations into the high-affinity *SP1-4* site [e.g., NCCR *12*(*4*-5CG)], flipping the Sp1 site orientation (e.g., NCCR *124r*), and creating various permutations and combinations thereof perturbed the EVGR-LVGR balance and in most cases led to a significant increase in the level of EVGR expression. In a key experiment, we made use of the construct NCCR *14r4*, which carries the same high-affinity *SP1-4* sequences in the respective EVGR- and LVGR-proximal positions, followed by introducing affinity-lowering mutations that caused a stepwise decline in the level of the activated EVGR phenotype. In another approach, the low-affinity *SP1-2* site was duplicated, showing that two low-affinity Sp1 sites were better than one or none, overall supporting the role of local Sp1-binding affinity. Importantly, the mutant NCCRs were functional in recombinant BKPyVs, yielding elevated viral replication compared to that of archetype BKPyV in RPTECs, as predicted. Thus, the nonpalindromic nature of the Sp1-binding sequence and the deviations from its consensus sequence affected the directionality of expression and regulated the delicate EVGR-LVGR balance of the BKPyV archetype NCCR.

We also obtained evidence that the NCCR balance of the Sp1 sites acts in conjunction with core promoter elements (CPE), showing an imperfect rotation symmetry of a developmental promoter on the EVGR and a housekeeping promoter on the LVGR side of the BKPyV archetype NCCR. In general, most eukaryotic promoters can be assigned to (i) developmental or regulated promoters, which are characterized by a focused transcriptional start, that need to be activated by a stimulus and contain either a *TATA* box or a DPE and (ii) housekeeping or tissue-restricted promoters that are constitutively active and show multiple, dispersed transcriptional start sites, that are *BREu* enriched, typically without a *TATA* box, and that contain a CpG island (51, 52). Indeed, the EVGR side carries downstream of the Sp1 site a perfect *TATA* box and one major *Inr*/TSS in a canonical distance of 31 bp, followed by a consensus downstream promoter element (DPE) at a distance of 28 bp (53, 54). Thus, our results indicate that the EVGR promoter of the BKPyV archetype NCCR fulfills the criteria of a regulated promoter with a perfect *TATA* box, a focused TSS, and a low level of basal activity. The LVGR promoter is less easily classified, since it bears the high-affinity *SP1-4* partially overlapping a potential *BREu*, followed by two *TATA*-like elements, a major *Inr* besides some dispersed TSS, and potential DPEs. We demonstrate here that the archetype BKPyV LVGR promoter shows a mixed focused-dispersed activity, which changes to a fully dispersed type of activity in the BKPyV Dunlop NCCR due to the deletion of the archetypal late promoter, extending the initial observation in hybrid NCCRs (50). The LVGR promoter is reminiscent of the previously described *TATA*-less housekeeping promoters described in the model of *Saccharomyces cerevisiae*, which essentially contain *TATA*-like elements (49). It is currently unclear whether *TATA*-like elements bind *TATA*-binding protein (TBP). One hypothesis is that they do but, compared to the results obtained with perfect *TATA* boxes, that binding leads to the assembly of altered preinitiation complexes (PIC), which—at least in yeast—have been shown to be Taf1/TFIID enriched (49). Alternatively, these kinds of sequences might be binding sites for paralogs of TBP, like TRF2, thereby leading to differential PIC formation and transcriptional activity. Accordingly, the imperfect symmetry of Sp1 sites, *TATA* elements, *Inr*, and DPEs in the BKPyV NCCR seems to allow the maintenance of the archetypal focus on the LVGR side while repressing EVGR expression. At the same time, this organization permits an effective bidirectional system that is highly poised to shift the transcriptional balance from LVGR to EVGR expression upon differential stimulation and mutation (32). A summary model of our findings is depicted in Fig. 9. It should be noted that a second layer of regulation that involves BKPyV microRNAs

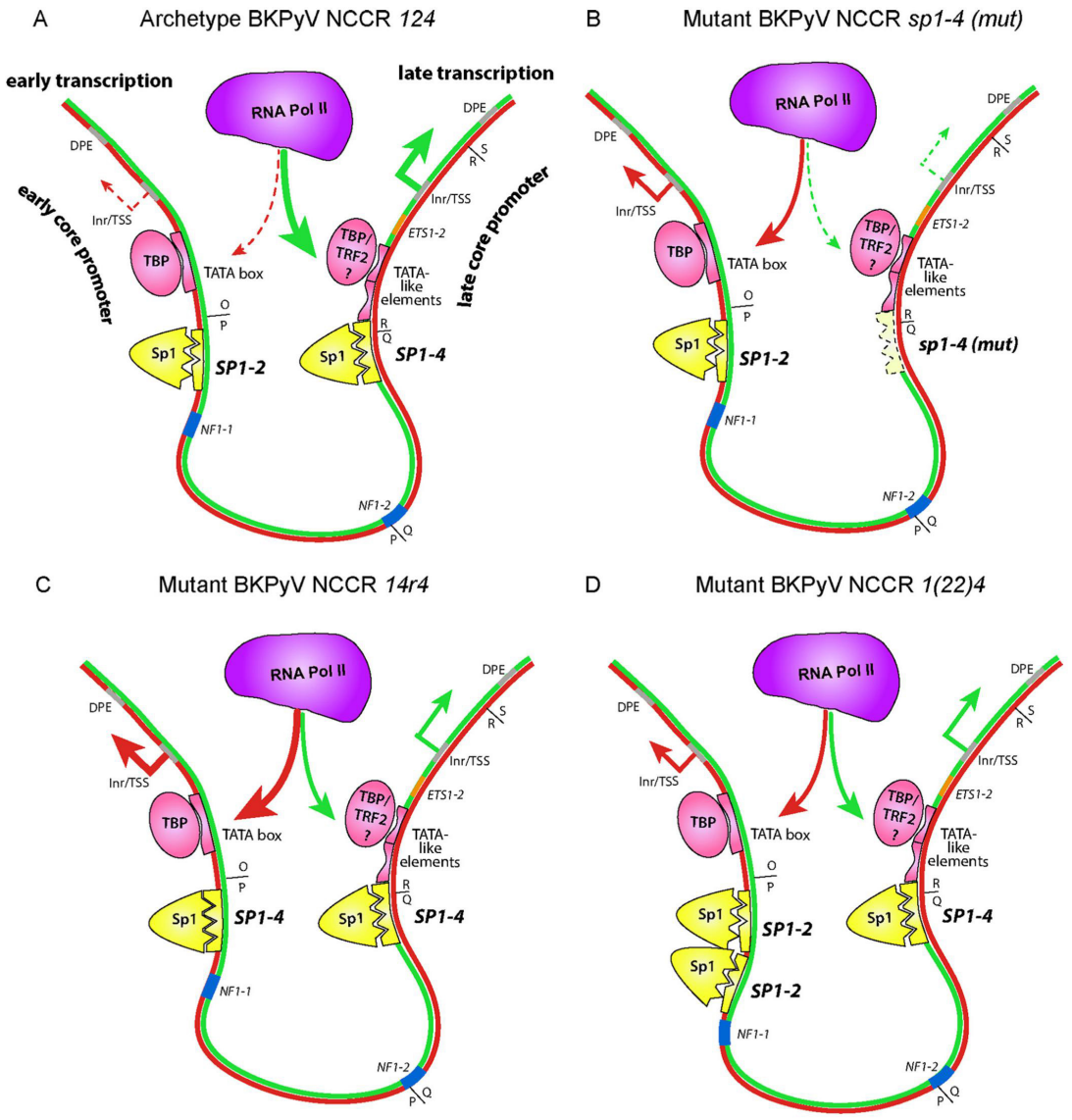


FIG 9 Sp1, TATA, and downstream promoter element (DPE) core promoter elements forming an imperfect symmetry underlying bidirectional EVGR and LVGR expression of the BKPyV NCCR. (A) BKPyV archetype NCCR *124* predominantly drives LVGR expression as a result of the high-affinity *SP1-4* (the three triangular indentations represent the zinc finger binding) and TATA-like elements, whereas EVGR expression is regulated by the low-affinity *SP1-2* site (two triangular indentations) and the *TATA* box. The Sp1 protein is depicted in yellow, with the three small triangles symbolizing zinc fingers. Pink ovals depict TBP (*TATA* box-binding protein) or TRF2 (TBP-related factor 2), with the latter potentially being recruited by *TATA*-like elements. A low level of early expression is due to less efficient recruitment of RNA polymerase II (Pol II) to the early promoter (red dashed arrows), and a high level of late transcription is due to recruitment of a major fraction of RNA polymerase II to the late promoter (fat green arrows). Inr, initiator element. For simplicity, other components, such as TFIIB and the TFIID complex or transcription factors like NF1, are not depicted (see reference 32). (B) Mutant BKPyV NCCR *124* carrying an *SP1-4* sequence with a point mutation (mut) that abrogated Sp1 affinity and thereby the recruitment of polymerase II to the late promoter, leading to a decrease in the level of LVGR expression (green dashed arrows). The low-affinity *SP1-2* now permits preferential recruitment of RNA polymerase II to the early promoter, causing increased levels of EVGR expression (red arrows). (C) Mutant BKPyV NCCR *14r4* in which the low-affinity *SP1-2* sequence was replaced with a high-affinity *SP1-4* sequence in the EVGR promoter. An increased fraction of RNA polymerase II is recruited to the early tandem of the *TATA* box and *SP1-4*, giving rise to high levels of EVGR expression (fat red arrows), while the unaltered late promoter is disfavored (green arrows). (D) Mutant BKPyV NCCR *1(22)4* carrying a duplication of the low-affinity *SP1-2* site in the early promoter increases RNA polymerase II recruitment and EVGR expression (red arrows), while the unaltered late promoter is disfavored (green arrows).

(miRNAs) also governs the effective control of EVGR expression. These miRNAs are transcribed in the late direction and target the EVGR transcripts, thus mediating repression of LTag expression (55). Thereby, persistent archetype BKPyV infection maintains a double-stitched mechanism to ensure tight control of LTag expression below the radar of the immune system (55, 56). In turn, activating rearrangements not only switch on EVGR expression but also overcome miRNA suppression, an alteration that requires the absence of a functional immune response (31).

The late DPEs located at the start of the S block are maintained in most patient isolates (31, 32), thus keeping the DPEs in place in *rr*-NCCR BKPyV variants lacking archetype late promoter sequences. Interestingly, these LVGR promoter deletions significantly increased EVGR expression at the expense of LVGR expression, but the associated increases in viral replication and progeny still depend on the corresponding expression of the LVGR structural gene products, e.g., Vp1, Vp2, and Vp3. We approached this issue by also providing two levels of SV40 LTag in *trans* in the HEK293 cell derivatives HEK293T and HEK293TT when transfecting selected NCCR mutants, thereby complementing the results of recombinant BKPyV infection in primary human RPTECs. The HEK293, HEK293T, and HEK293TT cell transfection experiments indicated that LTag expression in *trans* can overcome the adverse effects of deleting *SP1-4* or the TATA-like elements in the LVGR promoter. These results are in line with those of a previous study reporting that the SV40 LTag is able to rescue a crippled promoter function which was lost due to the lack of a TBP-associated factor (TAF), TAF(II)250 (57). TAFs and TBP have been shown to form the TFIID complex and target the core promoter region by binding to DPEs, Inr, and TATA elements (52, 58). Our data suggest that LTag is able to activate the LVGR promoter by executing a TAF-like function within the TFIID complex, helping to recognize the remaining CPEs, namely, the DPEs of the S_{63} block and the dispersed *Inr* in the P_{68} block of the NCCR. As mentioned earlier, few mutants also showed an increase in the level of EVGR expression, suggesting a small contribution of LTag to early promoter activity and/or an increase in episomal plasmid copy number via *ori*-binding functions. The NCCR *BREu*-in late mutant was a notable exception with respect to LVGR activation, as it showed little change in its level of expression in HEK293T and HEK293TT cells, suggesting a possible negative interference with LTag-mediated activation. This mutation abrogated Sp1 binding, while it introduced a perfect consensus *BREu*, thus indicating that the *SP1-4* site is more important than the potential *BREu* for driving LVGR expression.

This point mutational dissection of the bidirectional BKPyV NCCR may have implications for other HPyVs which show an architecture and an organization similar to those of BKPyV but which show a different TFBS composition and tissue specificity (59). In fact, NCCR rearrangements are best known from JC polyomavirus (JCPyV) variants from patient suffering from progressive multifocal leukoencephalopathy (46, 60–62). Following duplication of the early promoter and deletion of the late promoter sequences, the JCPyV *rr*-NCCRs demonstrate activation of EVGR at the expense of LVGR similar to those of BKPyV (46). These alterations also broaden the host cell range, which has particular relevance for neurotropism and pathology (34, 63, 64). Given the power of the bidirectional reporter construct to predict PyV NCCR expression and viral replication, our approach may help to close the current knowledge gaps in viral host cell specificity and replication of the dozen novel human PyVs.

In conclusion, this study provides a novel view of the BKPyV NCCR by revealing an imperfect symmetry of the EVGR and LVGR promoter characteristics that are functionally balanced by the position, orientation, and affinity of Sp1 and core promoter elements. As there is currently no antiviral treatment available for BKPyV or the closely related JCPyV, this extended mechanistic understanding of the BKPyV NCCR might be utilized to design new therapeutic strategies targeting viral persistence and reactivation. Inhibition of EVGR promoter activation or the interaction of LTag with the LVGR promoter might open up the possibility of blocking viral infection while maintaining sufficient immunosuppression in transplant recipients.

ACKNOWLEDGMENTS

We thank Rainer Gosert, Julia Manzetti, Gunhild Unterstab, and Marion Wernli of the research group Transplantation & Clinical Virology in Basel, Switzerland, as well as Christine Hanssen Rinaldo from the University Hospital North Norway/Arctic University in Tromsø, Norway, for support and helpful discussions.

FUNDING INFORMATION

This work was funded by an appointment grant of the University of Basel to Hans H. Hirsch. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

REFERENCES

1. **Polyomaviridae Study Group of the International Committee on Taxonomy of Viruses, Calvignac-Spencer S, Feltkamp MC, Daugherty MD, Moens U, Ramqvist T, Johne R, Ehlers B.** 2016. A taxonomy update for the family Polyomaviridae. Arch Virol **161:**1739–1750. http://dx.doi.org/10.1007/s00705-016-2794-y.
2. **DeCaprio JA, Garcea RL.** 2013. A cornucopia of human polyomaviruses. Nat Rev Microbiol **11:**264–276. http://dx.doi.org/10.1038/nrmicro2992.
3. **Rinaldo CH, Hirsch HH.** 2013. The human polyomaviruses: from orphans and mutants to patchwork family. APMIS **121:**681–684. http://dx.doi.org/10.1111/apm.12125.
4. **Kardas P, Leboeuf C, Hirsch HH.** 2015. Optimizing JC and BK polyomavirus IgG testing for seroepidemiology and patient counseling. J Clin Virol **71:**28–33. http://dx.doi.org/10.1016/j.jcv.2015.07.305.
5. **Kean JM, Rao S, Wang M, Garcea RL.** 2009. Seroepidemiology of human polyomaviruses. PLoS Pathog **5:**e1000363. http://dx.doi.org/10.1371/journal.ppat.1000363.
6. **Gossai A, Waterboer T, Nelson HH, Michel A, Willhauck-Fleckenstein M, Farzan SF, Hoen AG, Christensen BC, Kelsey KT, Marsit CJ, Pawlita M, Karagas MR.** 2016. Seroepidemiology of human polyomaviruses in a US population. Am J Epidemiol **183:**61–69. http://dx.doi.org/10.1093/aje/kwv155.
7. **Nicol JT, Robinot R, Carpentier A, Carandina G, Mazzoni E, Tognon M, Touze A, Coursaget P.** 2013. Age-specific seroprevalences of Merkel cell polyomavirus, human polyomaviruses 6, 7, and 9, and trichodysplasia spinulosa-associated polyomavirus. Clin Vaccine Immunol **20:**363–368. http://dx.doi.org/10.1128/CVI.00438-12.
8. **van der Meijden E, Kazem S, Dargel CA, van Vuren N, Hensbergen PJ, Feltkamp MC.** 2015. Characterization of T antigens, including middle T and alternative T, expressed by the human polyomavirus associated with trichodysplasia spinulosa. J Virol **89:**9427–9439. http://dx.doi.org/10.1128/JVI.00911-15.
9. **Tsang SH, Wang R, Nakamaru-Ogiso E, Knight SA, Buck CB, You J.** 2016. The oncogenic small tumor antigen of Merkel cell polyomavirus is an iron-sulfur cluster protein that enhances viral DNA replication. J Virol **90:**1544–1556. http://dx.doi.org/10.1128/JVI.02121-15.
10. **Tsai B, Qian M.** 2010. Cellular entry of polyomaviruses. Curr Top Microbiol Immunol **343:**177–194. http://dx.doi.org/10.1007/82_2010_38.
11. **Gardner SD, Field AM, Coleman DV, Hulme B.** 1971. New human

papovavirus (B.K.) isolated from urine after renal transplantation. Lancet **i:**1253–1257.
12. **Hirsch HH, Steiger J.** 2003. Polyomavirus BK. Lancet Infect Dis **3:**611–623. http://dx.doi.org/10.1016/S1473-3099(03)00770-9.
13. **Egli A, Kohli S, Dickenmann M, Hirsch HH.** 2009. Inhibition of polyomavirus BK-specific T-cell responses by immunosuppressive drugs. Transplantation **88:**1161–1168. http://dx.doi.org/10.1097/TP.0b013e3181bca422.
14. **Binggeli S, Egli A, Schaub S, Binet I, Mayr M, Steiger J, Hirsch HH.** 2007. Polyomavirus BK-specific cellular immune response to VP1 and large T-antigen in kidney transplant recipients. Am J Transplant **7:**1131–1139. http://dx.doi.org/10.1111/j.1600-6143.2007.01754.x.
15. **Hirsch HH, Knowles W, Dickenmann M, Passweg J, Klimkait T, Mihatsch MJ, Steiger J.** 2002. Prospective study of polyomavirus type BK replication and nephropathy in renal-transplant recipients. N Engl J Med **347:**488–496. http://dx.doi.org/10.1056/NEJMoa020439.
16. **Brennan DC, Agha I, Bohl DL, Schnitzler MA, Hardinger KL, Lockwood M, Torrence S, Schuessler R, Roby T, Gaudreault-Keener M, Storch GA.** 2005. Incidence of BK with tacrolimus versus cyclosporine and impact of preemptive immunosuppression reduction. Am J Transplant **5:**582–594. http://dx.doi.org/10.1111/j.1600-6143.2005.00742.x.
17. **Randhawa P, Ho A, Shapiro R, Vats A, Swalsky P, Finkelstein S, Uhrmacher J, Weck K.** 2004. Correlates of quantitative measurement of BK polyomavirus (BKV) DNA with clinical course of BKV infection in renal transplant patients. J Clin Microbiol **42:**1176–1180. http://dx.doi.org/10.1128/JCM.42.3.1176-1180.2004.
18. **Ginevri F, Azzi A, Hirsch HH, Basso S, Fontana I, Cioni M, Bodaghi S, Salotti V, Rinieri A, Botti G, Perfumo F, Locatelli F, Comoli P.** 2007. Prospective monitoring of polyomavirus BK replication and impact of pre-emptive intervention in pediatric kidney recipients. Am J Transplant **7:**2727–2735. http://dx.doi.org/10.1111/j.1600-6143.2007.01984.x.
19. **Funk GA, Gosert R, Comoli P, Ginevri F, Hirsch HH.** 2008. Polyomavirus BK replication dynamics in vivo and in silico to predict cytopathology and viral clearance in kidney transplants. Am J Transplant **8:**2368–2377. http://dx.doi.org/10.1111/j.1600-6143.2008.02402.x.
20. **Hirsch HH, Vincenti F, Friman S, Tuncer M, Citterio F, Wiecek A, Scheuermann EH, Klinger M, Russ G, Pescovitz MD, Prestele H.** 2013. Polyomavirus BK replication in de novo kidney transplant patients receiving tacrolimus or cyclosporine: a prospective, randomized, multicenter study. Am J Transplant **13:**136–145. http://dx.doi.org/10.1111/j.1600-6143.2012.04320.x.
21. **Randhawa PS, Finkelstein S, Scantlebury V, Shapiro R, Vivas C, Jordan M, Picken MM, Demetris AJ.** 1999. Human polyoma virus-associated interstitial nephritis in the allograft kidney. Transplantation **67:**103–109. http://dx.doi.org/10.1097/00007890-199901150-00018.
22. **Binet I, Nickeleit V, Hirsch HH, Prince O, Dalquen P, Gudat F, Mihatsch MJ, Thiel G.** 1999. Polyomavirus disease under new immunosuppressive drugs: a cause of renal graft dysfunction and graft loss. Transplantation **67:**918–922. http://dx.doi.org/10.1097/00007890-199903270-00022.
23. **Hirsch HH, Randhawa P.** 2013. BK polyomavirus in solid organ transplantation. Am J Transplant **13**(Suppl 4)**:**S179–S188. http://dx.doi.org/10.1111/ajt.12110.
24. **Arthur RR, Shah KV, Baust SJ, Santos GW, Saral R.** 1986. Association of BK viruria with hemorrhagic cystitis in recipients of bone marrow transplants. N Engl J Med **315:**230–234. http://dx.doi.org/10.1056/NEJM198607243150405.
25. **Cesaro S, Dalianis T, Rinaldo CH, Koskenvuo M, Einsele H, Hirsch HH.** 2016. ECIL 6—guidelines for the prevention, diagnosis, and treatment of BK polyomavirus disease in stem cell transplant patients. ECIL 6 Meet, Sophia Antipolis, France, 11–12 September 2015. https://www.ebmt.org/Contents/Resources/Library/ECIL/Documents/2015%20ECIL6/ECIL6-BK-virus-08-12-2015-Cesaro-S-et-al%20final.pdf.
26. **Hirsch HH, Pergam SA.** 2016. Human adenovirus, polyomavirus, and parvovirus infections in patients undergoing hematopoietic stem-cell transplantation, p 1090–1104. *In* Forman SJ, Negrin RS, Antin H, Appelbaum FR (ed), Thomas' hematopoietic cell transplantation, 5th ed. John Wiley & Sons Ltd., Chichester, United Kingdom. http://www.wiley.comgoappelbaumTransplantation.
27. **Randhawa P, Zygmunt D, Shapiro R, Vats A, Weck K, Swalsky P, Finkelstein S.** 2003. Viral regulatory region sequence variations in kidney tissue obtained from patients with BK virus nephropathy. Kidney Int **64:**743–747. http://dx.doi.org/10.1046/j.1523-1755.2003.00132.x.
28. **Bressollette-Bodin C, Coste-Burel M, Hourmant M, Sebille V, Andre-Garnier E, Imbert-Marcille BM.** 2005. A prospective longitudinal study of BK virus infection in 104 renal transplant recipients. Am J Transplant **5:**1926–1933. http://dx.doi.org/10.1111/j.1600-6143.2005.00934.x.
29. **Egli A, Infanti L, Dumoulin A, Buser A, Samaridis J, Stebler C, Gosert R, Hirsch HH.** 2009. Prevalence of polyomavirus BK and JC infection and replication in 400 healthy blood donors. J Infect Dis **199:**837–846. http://dx.doi.org/10.1086/597126.
30. **Hirsch HH, Yakhontova K, Lu M, Manzetti J.** 2016. BK polyomavirus replication in renal tubular epithelial cells is inhibited by sirolimus, but activated by tacrolimus through a pathway involving FKBP-12. Am J Transplant **16:**821–832. http://dx.doi.org/10.1111/ajt.13541.
31. **Gosert R, Rinaldo CH, Funk GA, Egli A, Ramos E, Drachenberg CB, Hirsch HH.** 2008. Polyomavirus BK with rearranged noncoding control region emerge in vivo in renal transplant patients and increase viral replication and cytopathology. J Exp Med **205:**841–852. http://dx.doi.org/10.1084/jem.20072097.
32. **Bethge T, Hachemi HA, Manzetti J, Gosert R, Schaffner W, Hirsch HH.** 2015. Sp1 sites in the noncoding control region of BK polyomavirus are key regulators of bidirectional viral early and late gene expression. J Virol **89:**3396–3411. http://dx.doi.org/10.1128/JVI.03625-14.
33. **Imperiale MJ, Jiang M.** 2015. What DNA viral genomic rearrangements tell us about persistence. J Virol **89:**1948–1950. http://dx.doi.org/10.1128/JVI.01227-14.
34. **DeCaprio JA, Imperiale MJ, Major EO.** 2013. Polyomaviruses, p 1633–1661. In Howley PM, Cohen JI, Griffin DE, Lamb RA, Martin MA, Racaniello VR, Roizman B (ed), Fields virology, 6th ed. Lippincott Williams & Wilkins, Philadelphia, PA.
35. **Henriksen S, Tylden GD, Dumoulin A, Sharma BN, Hirsch HH, Rinaldo CH.** 2014. The human fetal glial cell line SVG p12 contains infectious BK polyomavirus. J Virol **88:**7556–7568. http://dx.doi.org/10.1128/JVI.00696-14.
36. **Bernhoff E, Gutteberg TJ, Sandvik K, Hirsch HH, Rinaldo CH.** 2008. Cidofovir inhibits polyomavirus BK replication in human renal tubular cells downstream of viral early gene expression. Am J Transplant **8:**1413–1422. http://dx.doi.org/10.1111/j.1600-6143.2008.02269.x.
37. **Markowitz RB, Eaton BA, Kubik MF, Latorra D, McGregor JA, Dynan WS.** 1991. BK virus and JC virus shed during pregnancy have predominantly archetypal regulatory regions. J Virol **65:**4515–4519.
38. **Negrini M, Sabbioni S, Arthur RR, Castagnoli A, Barbanti-Brodano G.** 1991. Prevalence of the archetypal regulatory region and sequence polymorphisms in nonpassaged BK virus variants. J Virol **65:**5092–5095.
39. **Li R, Sharma BN, Linder S, Gutteberg TJ, Hirsch HH, Rinaldo CH.** 2013. Characteristics of polyomavirus BK (BKPyV) infection in primary human urothelial cells. Virology **440:**41–50. http://dx.doi.org/10.1016/j.virol.2013.01.024.
40. **Olsen GH, Hirsch HH, Rinaldo CH.** 2009. Functional analysis of polyomavirus BK non-coding control region quasispecies from kidney transplant recipients. J Med Virol **81:**1959–1967. http://dx.doi.org/10.1002/jmv.21605.
41. **Olsen GH, Andresen PA, Hilmarsen HT, Bjorang O, Scott H, Midtvedt K, Rinaldo CH.** 2006. Genetic variability in BK virus regulatory regions in urine and kidney biopsies from renal-transplant patients. J Med Virol **78:**384–393. http://dx.doi.org/10.1002/jmv.20551.
42. **Schreiber E, Matthias P, Muller MM, Schaffner W.** 1988. Identification of a novel lymphoid specific octamer binding protein (OTF-2B) by proteolytic clipping bandshift assay (PCBA). EMBO J **7:**4221–4229.
43. **Schreiber E, Matthias P, Muller MM, Schaffner W.** 1989. Rapid detection of octamer binding proteins with 'mini-extracts', prepared from a small number of cells. Nucleic Acids Res **17:**6419. http://dx.doi.org/10.1093/nar/17.15.6419.
44. **Wang S, Wu S, Meng Q, Li X, Zhang J, Chen R, Wang M.** 2016. FAS rs2234767 and rs1800682 polymorphisms jointly contributed to risk of colorectal cancer by affecting SP1/STAT1 complex recruitment to chromatin. Sci Rep **6:**19229. http://dx.doi.org/10.1038/srep19229.
45. **Sloutskin A, Danino YM, Orenstein Y, Zehavi Y, Doniger T, Shamir R, Juven-Gershon T.** 2015. ElemeNT: a computational tool for detecting core promoter elements. Transcription **6:**41–50. http://dx.doi.org/10.1080/21541264.2015.1067286.
46. **Gosert R, Kardas P, Major EO, Hirsch HH.** 2010. Rearranged JC virus noncoding control regions found in progressive multifocal leukoencephalopathy patient samples increase virus early gene expression and replication rate. J Virol **84:**10448–10456. http://dx.doi.org/10.1128/JVI.00614-10.
47. **Dumoulin A, Hirsch HH.** 2011. Reevaluating and optimizing polyoma-

virus BK and JC real-time PCR assays to detect rare sequence polymorphisms. J Clin Microbiol **49:**1382–1388. http://dx.doi.org/10.1128/JCM.02008-10.

48. **Priftakis P, Bogdanovic G, Kalantari M, Dalianis T.** 2001. Overrepresentation of point mutations in the Sp1 site of the non-coding control region of BK virus in bone marrow transplanted patients with haemorrhagic cystitis. J Clin Virol **21:**1–7. http://dx.doi.org/10.1016/S1386-6532(00)00171-2.
49. **Rhee HS, Pugh BF.** 2012. Genome-wide structure and organization of eukaryotic pre-initiation complexes. Nature **483:**295–301. http://dx.doi.org/10.1038/nature10799.
50. **Kraus RJ, Shadley L, Mertz JE.** 2001. Nuclear factor 1 family members mediate repression of the BK virus late promoter. Virology **287:**89–104. http://dx.doi.org/10.1006/viro.2001.1024.
51. **Kadonaga JT.** 2012. Perspectives on the RNA polymerase II core promoter. Wiley Interdiscip Rev Dev Biol **1:**40–51. http://dx.doi.org/10.1002/wdev.21.
52. **Decker KB, Hinton DM.** 2013. Transcription regulation at the core: similarities among bacterial, archaeal, and eukaryotic RNA polymerases. Annu Rev Microbiol **67:**113–139. http://dx.doi.org/10.1146/annurev-micro-092412-155756.
53. **Kutach AK, Kadonaga JT.** 2000. The downstream promoter element DPE appears to be as widely used as the TATA box in Drosophila core promoters. Mol Cell Biol **20:**4754–4764. http://dx.doi.org/10.1128/MCB.20.13.4754-4764.2000.
54. **Ohler U, Liao GC, Niemann H, Rubin GM.** 2002. Computational analysis of core promoters in the Drosophila genome. Genome Biol **3:**RESEARCH0087.
55. **Broekema NM, Imperiale MJ.** 2013. miRNA regulation of BK polyomavirus replication during early infection. Proc Natl Acad Sci U S A **110:**8200–8205. http://dx.doi.org/10.1073/pnas.1301907110.
56. **Cioni M, Leboeuf C, Comoli P, Ginevri F, Hirsch HH.** 2016. Characterization of immunodominant BK polyomavirus 9mer epitope T cell responses. Am J Transplant **16:**1193–1206. http://dx.doi.org/10.1111/ajt.13598.
57. **Yu Y, Alwine JC.** 2002. Human cytomegalovirus major immediate-early proteins and simian virus 40 large T antigen can inhibit apoptosis through activation of the phosphatidylinositide 3′-OH kinase pathway and the cellular kinase Akt. J Virol **76:**3731–3738. http://dx.doi.org/10.1128/JVI.76.8.3731-3738.2002.
58. **Goodrich JA, Tjian R.** 2010. Unexpected roles for core promoter recognition factors in cell-type-specific transcription and gene regulation. Nat Rev Genet **11:**549–558. http://dx.doi.org/10.1038/ni0710-549.
59. **Moens U, Van Ghelue M, Ludvigsen M, Korup-Schulz S, Ehlers B.** 2015. Early and late promoters of BK polyomavirus, Merkel cell polyomavirus, Trichodysplasia spinulosa-associated polyomavirus and human polyomavirus 12 are among the strongest of all known human polyomaviruses in 10 different cell lines. J Gen Virol **96:**2293–2303. http://dx.doi.org/10.1099/vir.0.000181.
60. **Miyamura T, Furuno A, Yoshiike K.** 1985. DNA rearrangement in the control region for early transcription in a human polyomavirus JC host range mutant capable of growing in human embryonic kidney cells. J Virol **54:**750–756.
61. **Martin JD, King DM, Slauch JM, Frisque RJ.** 1985. Differences in regulatory sequences of naturally occurring JC virus variants. J Virol **53:**306–311.
62. **Loeber G, Dorries K.** 1988. DNA rearrangements in organ-specific variants of polyomavirus JC strain GS. J Virol **62:**1730–1735.
63. **Marshall LJ, Moore LD, Mirsky MM, Major EO.** 2012. JC virus promoter/enhancers contain TATA box-associated Spi-B-binding sites that support early viral gene expression in primary astrocytes. J Gen Virol **93:**651–661. http://dx.doi.org/10.1099/vir.0.035832-0.
64. **Tan CS, Ellis LC, Wuthrich C, Ngo L, Broge TA, Jr, Saint-Aubyn J, Miller JS, Koralnik IJ.** 2010. JC virus latency in the brain and extraneural organs of patients with and without progressive multifocal leukoencephalopathy. J Virol **84:**9200–9209. http://dx.doi.org/10.1128/JVI.00609-10.

We demonstrated that the affinity, position and strand orientation of Sp1-2 and Sp1-4 proximal to the EVGR and LVGR sites exerted different effects on BKPyVww NCCR-driven EVGR and LVGR expression. Swapping the high-affinity Sp1-4 site and the low-affinity Sp1-2 site (e.g., NCCR *14r2r*), lowering the Sp1-4 affinity by introducing affinity-lowering mutations (e.g., NCCR *12*(*4*-5CG)), flipping the orientation of the Sp1 sites (e.g., NCCR *124r*) and various combinations of permutations upset the EVGR and LVGR balance and led to a significant EVGR expression in most cases. A step-wise lowering-affinity mutations caused a decline in EVGR expression in the NCCR *14r4* containing the high-affinity Sp1-4 site in respective EVGR and LVGR proximal positions. Duplicating the lower-affinity Sp1-2 site increased EVGR expression and elevated viral replication compared to BKPyV ww, suggesting that two lower-affinity sites were better than one or none. The non-palindromic nature of Sp1 sites and alteration from the consensus organization affected the BKPyVww NCCR-driven EVGR-LVGR balance.

There was evidence, that the Sp1 regulatory balance of EVGR and LVGR acted in conjunction with CPEs showing an imperfect symmetry of Sp1, EVGR and LVGR core promoters with TATA and TATA-like elements. The EVGR of BKPyVww NCCR fulfil the criteria of a regulatory promoter whereas LVGR promoter cannot be easily classified, but it is similar to a TATA-less housekeeping promoter described in *Saccharomyces cerevisiae*, which actually contains TATA-like elements.

Most rr-NCCR variants from patients lacking the late promoter sequences maintained the late DPEs. Deletion of the LVGR promoter increased EVGR expression at the expense of LVGR expression. However, progeny production relied on LVGR gene products. Transfection experiments of Sp1-4 and TATA-like elements deletion in the LVGR promoter in cell lines expressing different amounts of SV40 LTag showed, that LTag can overcome the effect of the deletions by increasing LVGR expression. The results corroborate previous studies showing that SV40 LTag was able to rescue a non-functional promoter lost due to lack of TBP-associated factor (TAF). Since TAFs and TBP have been shown to form the TFIID complex and bind the core promoter elements (DPEs, Inr, and TATA elements), LTag may activate LVGR promoter by exerting a TAF-like function with TFIID complex to recognize the remaining CPEs, the DPEs and the dispersed Inr.

The dissection of these different point mutations in BKPyV ww NCCR may also have implications for other HPyVs, especially those with similar NCCR architecture. Rearranged JCPyV NCCR variants from PML from patients showed EVGR promoter duplication and LVGR promoter deletion resulting in increased EVGR expression in the expense of LVGR expression similar to BKPyV.
This study revealed an imperfect symmetry of the BKPyV EVGR and LVGR promoter consisting of TATA and TATA-like elements, CPEs which function in conjunction with Sp1 binding sites to regulate the bi-directional balance of EVGR and LVGR expression. This knowledge can be used to design new therapeutic approaches against BKPyV infection, especially in transplant patients.

Manuscript under review.

5.5 Donor-derived urothelial cancer after kidney transplantation associated with a BK polyomavirus with increased oncogenic potential

Potentially oncoproteins, Tags are encoded by all HPyVs. These viruses can cause cancer by interrupting cell proliferation control. The Tags interact with tumor suppressors p53 and pRB and drive the cell into S-phase, when components of the host DNA synthesis machinery are expressed to support viral DNA replication. A key concept in HPyV transformation has been that of permissiveness. Once HPyV enters a permissive cell and EVGR, DNA replication and LVGR happen, the outcome is productive infection. On the other hand, when HPyV infects a cell that cannot allow viral genome replication and LVGR expression, but supports EVGR expression, the outcome is cellular transformation (17).
BKPyV has been reported in 3 single cases of urothelial and renal tubular malignancies in the context of kidney transplantation. Additionally, in our study, BKPyV was found associated with urothelial cancer of the bladder (UC) in a kidney transplant recipient. The BKPyV genome was amplified by PCR from the extracted DNA of the tumor using outward primers from the Vp1 gene at the EcoR1 site and was sequenced. A full-length BKPyV genome (5116 bp) containing 17 bp deletion in

the p-block of the NCCR (P41-P57) and a shorter genomic fragment (3270 bp) lacking the NCCR, an N-terminal part of Vp2 and LTag was revealed. The viral genome was integrated into the cellular genome, and abundant BKPyV/SV40 LTag was detected in the tumor.

For functional analysis, the UC BKPyV NCCR was cloned into the bi-directional reporter vector phRG, transfected in HEK293 cells and the EVGR and LVGR expression compared to that of BKPyV-DUN and BKPyV (ww)-NCCR reporter constructs. Recombinant viruses were made from two clones of the UC BKPyV NCCR variant and the replication was compared to that of BKPyVww, BKPyV-DUN and Sp1-4 BKPyV NCCR mutant described above in **section 5.3.**

Donor-derived, metastatic urothelial cancer after kidney transplantation associated with a BK polyomavirus with oncogenic potential
- Oncogenic BK virus and donor kidney derived urothelial cancer

David C. Müller[1,2#], Maarit Rämö[1,2,#], Klaudia Naegele[3], Sebastian Ribi[2], Christian Wetterauer[1], Valeria Perrina[2], Luca Quagliata[2], Tatjana Vlajnic[2], Christian Ruiz[2], Beate Balitzki[4], Rainer Grobholz[5], Rainer Gosert[3], **Elvis T. Ajuh[6]**, Hans H. Hirsch[3,6,‡], Lukas Bubendorf[2,‡], Cyrill A. Rentsch[1, ‡]

1 Department of Urology, University Hospital Basel, Switzerland
2 Institute for Pathology, University Hospital Basel, Switzerland
3 Division Infection Diagnostics, Department Biomedicine, University of Basel
4 Institute of Forensic Medicine, University of Basel, Switzerland
5 Institute for Pathology, Kantonsspital Aarau, Switzerland
6 Transplantation & Clinical Virology, Department Biomedicine, University of Basel, Switzerland
[#,‡] equal contribution

Corresponding author:
Cyrill A. Rentsch, MD-PhD
University Hospital Basel
Spitalstrasse 21
4031 Basel, Switzerland
e-mail: crentsch@uhbs.ch
phone: +41 (0) 61 558 71 22

There are no conflicts of interest. The array data was submitted to the Gene Expression Omnibus repository. The accession number is GSE90778. Reviewers may gain access using the following link: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=cxovygacbhuzbyd&acc=GSE90778.

Word count for main text: 1991

Abstract

BK polyomavirus (BKPyV) increases the risk of developing urothelial carcinoma (UC) in immunosuppressed patients. Here, we performed a comprehensive genomic analysis of a BKPyV-associated, metachronous, multifocal and finally metastatic micropapillary UC in a 51 years-old kidney transplant (KTx) recipient. Dissecting cancer heterogeneity by sorting technologies prior to array CGH followed by short tandem repeat analysis revealed that the metastatic UC originated from the donor organ, a kidney from a 4-year-old boy at 10 years after transplantation. The top 50 cancer associated genes displayed no key driver mutations as assessed by next generation sequencing (NGS) of the cancer genome. Combining NGS of the viral genome and BKPyV-specific amplification provided evidence for full-length episomal and chromosomally integrated BKPyV (break point chr:1), both of which contained a unique and so far non-reported deletion in the non-coding control region (NCCR) of the BKPyV genome. This deletion resulted in impaired progression into the late viral life cycle and cytolysis. Consequently, urothelial cells are exposed to high levels of the early and transforming BKPyV proteins LTag and sTag from both, the chromosomally integrated and the episomal viral genome. To the best of our knowledge, this report links for the first time BKPyV-NCCR rearrangement with its transforming capacity in UC. A closer look at mutations BKPyV and the oncogenic implications is warranted in immunosuppressed patients.

Keywords

Urothelial, cancer, transplantation, kidney, bladder, polyomavirus, mutation, genome

Brief definitive report

Introduction

Several studies report an increased incidence of malignancies after solid organ transplantation. This was not explainable by the transmission of donor malignancies alone. Instead, recent findings indicate a role of infectious diseases in tumourigenesis [1, 2]. In particular, an association of BK polyomavirus (BKPyV) infection and the risk for urothelial cancer (UC) was suspected [3]. BKPyV genome contains three different functional areas, the early viral genome region (EVGR), the late viral genome region (LVGR) and the non-coding control region (NCCR). EVGR-proteins such as the large tumour-antigen (LTag) and the small tumour-antigen (sTag) have been identified as oncoproteins [4].

We report on a BKPyV-associated, multifocal, metastatic and micropapillary UC in a kidney transplant recipient. At the age of 42 years, he received a kidney from a 4-year-old male donor. Eight years post-transplantation (ptx), BKPyV/SV40 large T-antigen (LTag) positive UC of the bladder was diagnosed. A first recurrence appeared nine years ptx. Ten years ptx he was diagnosed with a BKPyV/SV40 LTag-positive, multifocal micropapillary and muscle-invasive UC with involvement of the graft kidney pelvis, the bladder wall and a single pelvic lymph node metastasis (Figure 1A, Figure S1A, Table S1) as evidenced after surgical removal of the organs.

Despite multiple recurrences and progression to metastatic UC disease, the patient has shown no evidence of UC during the 4 years of follow-up after surgery and termination of immunosppressive therapy. Noteworthy, the patient had not received any systemic treatment other than termination of the immunosuppressive therapy.

Material and Methods

In the course of routine diagnostics morphologic histopathological evaluation was performed. Immunohistochemistry identified the presence of polyomavirus using an antibody raised against the simian virus 40 (SV40) LTag (MRQ-4 mouse clone, Ventana Medical Systems, Inc., Tucson, AZ) on the Ventana BenchMark XT platform. Fresh-frozen tissue obtained from surgical specimen was used for further analyses. We applied DNA content based cell nuclei sorting prior to high-resolution genomic analysis by arrayCGH in order to discover the evolutionary history of this individual cancer [5].

To investigate the genealogy of the cancer tissues, we performed STR analysis. The investigated samples were the patient's innate kidney as a reference for recipient's DNA and healthy allograft-kidney tissue as a donor's DNA reference.

To further analyze the relatedness of the different tumour sites and to search for possible tumour-causing mutations we performed targeted NGS. The IonTorrent-platform by Life Technologies (Carlsbad, CA, USA) was used for sequencing. We performed targeted sequencing using the commercially available IonAmpliseq™ Cancer Hotspot Panel v2. 10 ng input of extracted genomic DNA was used for library preparation.

Extracted DNA from tumour and unaffected tissue were tested by a quantitative real-time PCR for BKPyV genome loads [6] and a human host cell gene (acetylase) was quantified for normalization per total number of diploid cells [7]. BKPyV-specific long-range PCR was performed using outward primers partially overlapping in EcoR1 site present in the VP1 gene [8] giving rise to a full-length BKPyV genome of 5116 bp and shorter derivative of 3500 bp for sequencing. The entire BKPyV genome was sequenced using different primer combinations either directly from the extracted tumour DNA following nested PCR or using the gel-purified long-range PCR products. The sequences were analyzed using codon code aligner (Codon Code Corporation, Centerville, MA, USA) and the FASTA files were depicted using ApE [9]. To detect potential chromosomal integration, shallow whole genome NGS was performed searching for BKPyV genome sequences using the Illumina NextSeq (Illumina, CA, USA) sequencing platform (10x average read depth). The libraries were prepared using the Illumina Nextera XT DNA Library Prep kit according to protocol. Reads were mapped against the hs37d5 reference genome as well as the BKPyV reference

sequence (NCBI acc. no. AB211371). Integration was called if read ends alignment to both reference sequences. Additionally, to sequence the BKPyV genome with higher coverage, targeted amplicon sequencing was performed on the Illumina MiniSeq with previously amplified amplicons of 5116bp and 3270bp and libraries prepared with Nextera XT DNA Library Prep kit. The data was analyzed with the CLC Genomics Workbench (Qiagen Bioinformatics, Denmark) and the Integrative Genomics Viewer (IGV, Broad Institute, USA). Primers spanning the integration breakpoints were designed based on detection of paired reads from whole genome sequencing aligning to the BKPyV reference sequence and chromosome 1 in the human genome (GRCh37/hg19). Using these primers, the integration of BKPyV into chromosome 1 was confirmed and the integrated parts of the genome sequenced by Sanger sequencing.

For a functional analysis, the NCCR of the BKPyV-UC genome was amplified, and inserted into the bidirectional reporter gene plasmid pHRG1, and sequenced as described previously [10, 11]. The respective NCCR pHRG1 reporter constructs were transfected into HEK293 cells and then quantified by flow-cytometry for EVGR expression (red fluorescence) and LVGR expression (green fluorescence) as described previously [11]. Recombinant full-length BKPyV genomes were generated by inserting the NCCRs derived from the Dunlop, WW archetype; archetype point mutant (*sp1-4*) and two BKPyV UC-derived NCCR clones CA1-2 and CA1-5 into the Dunlop genome as described previously [10, 11]. The recombinant BKPyV genomes were cut out from the plasmids and re-circularized prior to transfection of COS7 cells. Infectious supernatants from the transfected COS7 cells were harvested after 7 to 10 days for determination of the BKPyV load by real-time PCR [6]. Equivalent amounts were used for infection of human primary proximal renal tubular epithelial cells (hRPTECs) as described [10, 11]. The hRPTEC culture supernatants were harvested at 1, 3, 5, 7, and 9 days post-infection and BKPyV loads released into the culture supernatants were compared using real-time PCR [10, 11].

We obtained ethical approval from the Ethical Committee Nordwest- und Zentralschweiz in Basel, Switzerland (EKNZ 2014-313).

Results and Discussion

ArrayCGH revealed that the UC (both, the metastasis and the bladder localization) most likely originated from the donor kidney pelvis (Figure 1B, Figure S1B). Adding short-tandem-repeat (STR) analyses confirmed that the UC originated from the graft kidney pelvis of an in summary 14 year-old donor organ. Surprisingly, we did not find any somatic key-driver point mutation by targeted NGS in 50 cancer genes (Table S2). Molecular characterization of the BKPyV strain in the 10y ptx cancer specimens revealed both integrated and episomal BKPyV genomes (Figure 2). The presence of the integrated BKPyV genome was detected only after performing whole genome NGS and identified the breakpoints in chr1:16055650_PLEKHM2 and the capsid coding sequences of VP1 and Vp2/VP3, thereby deleting 1300 bp between VP1 and VP2/3 (Figure 2; supplement data). In addition, episomal BKPyV genomes consisting of a full-length and a smaller deletion episome was detected. Importantly, all three genomes carried the identical 17bp deletion in the non-coding control region (NCCR) of BKPyV (Figure 2; ΔATGACCTCAGGAAGGAA). This deletion was unique and affected NCCR sequences, which are rarely affected in other natural variants (Figure 3B), and contains a multitude of predicted transcription factor binding sites [10, 11]. Functional analyses in a bidirectional reporter assay revealed that this NCCR-variant is sufficient to significantly activate EVGR expression and contributes to LTag expression (Figure 3B). However, the NCCR-deletion impairs progression into the late viral life cycle and thereby offsets efficient lytic replication (Figure 3C). The intact full-length viral episome carrying the identical NCCR-deletion as the truncated integrated genome argue that this NCCR-signature deletion occurred first, and was followed by viral genomeintegration, both of which contribute to LTag expression.

Notably, this patient had a favorable outcome despite both, unfavorable pathological stage and histopathological features. Postoperative termination of immunosuppression alone without any neoadjuvant or adjuvant systemic chemotherapy was sufficient to control the disease. This points towards a strong transforming oncogenic force of aberrant LTag expression from a unique BKPyV-NCCR deletion, which cannot be cleared immunologically unless immunosuppression is discontinued. Indeed, recent work by Leboeuf et al. has demonstrated an important contribution of LTag-directed CD8 T-cells to BKPyV clearance in KTX patients [12].

It came as a surprise that the UC originated in the pelvis of an in summary 14 year-old donor kidney rather than in a 52 year-old recipient's urinary bladder that had a much longer history of potential exposure to carcinogenic agents, such as the actual ongoing cigarette smoking, and the accumulation of oncogenic events during lifetime. It remains challenging to dissect the relative contribution of immunosuppression and ongoing BKPyV replication during this process. It is well known that chronic immune-suppression by itself favors the development of malignancy [1]. However, one would expect pre-neoplastic conditioning by exposure to carcinogens or hereditary susceptibility before such cells can be unleashed to give rise to established cancer after release of immune control.

There is growing evidence that reactivation of polyomavirus infection under immunosuppression might play an important role in the development of aggressive UC after renal transplantation [13]. A driving role of BKPyV infection is also endorsed by the low mutational load in cancer related genes as compared to common urothelial carcinomas known to be among the malignant human tumors with the highest mutation rate [14]. This is in line with findings of low mutational loads in Merkel cell carcinoma, which is also known to be associated with BKPyV infection [15]. Intriguingly, all of the tumour manifestations of our patient across time (at 8y, 9y and 10y ptx) were SV40 LTag-positive, and eventually turned into a micropapillary variant at the final, most advanced stage. The micropapillary variant of UC has previously been linked to SV40 LTag-positive aggressive UC developed under immunosuppression [13]. The known association between SV40-LTag positivity with aggressive micropapillary morphology strongly suggests that BKPyV is not an innocent bystander infection, but rather acts as a driving oncogenic force in such a setting [16]. Although high-level BKPyV replication occurs in 30% of all patients after renal transplantation, the vast majority of UCs diagnosed in patients is SV40 LTag-negative, whereas the aggressive, SV40 LTag-positive UC is rare. The factors that render BKPyV oncogenic in individual patients remain largely unknown, but consistently are associated with constitutive EVGR expression encoding LTag and sTag [4]. Viral integration into the human DNA has been proposed as an important feature of oncogenic polyomaviruses disrupting constitutive EVGR- from lytic LVGR expression by interrupting LVGR gene-sequences. This can occur by breakpoints in the viral DNA either followed by host DNA [17] or further copies of viral DNA [18]. Such integration is supposed to support an imbalance between increased LTag expression and counteracting regulatory viral

elements, resulting in sustained perturbation of p53 and pRb, both binding to LTag. Our analyses provide evidence of both, NCCR-rearrangement and genomic integration, although the relative contribution of either in the last metastasizing state cannot be assessed. Our functional studies challenge the view that integration of polyomaviruses is a *condition sine qua non* to tumourigenesis. The deletion in the NCCR P-segment in a non-integrated BKV may lead to the same imbalance as reported before by the interruption of the LVGR [17, 18]. The 17bp deletion in the here reported virus constitutively activated LTag, and sTag EVGR expression, while impairing LVGR and the replicative capacity of isogenic derivatives. Therefore, the lytic virus replication with accumulation of viral particles and miRNAs and subsequent cell death appeared to be significantly impaired. Importantly, all stages of the UC in our patient were SV40 LTag positive suggesting that expression of this virus-encoded transforming function was not dispensable, but most likely contributed to malignant progression. Given the relapse-free time of 4 years after diagnosis of this metastasizing UC also points to the role of regaining anti-tumour immunity following stop of immunosuppression, in line with recent immunological studies [12].

Taken together, the rare occurrence of aggressive SV40 LTag-positive UC can be regarded as a "biological accident" in the light of the high prevalence of high-level BKPyV replication in immunosuppressed patients e.g. after kidney transplantation [17]. However, given the high level of BKPyV replication and the huge number of viral generations in an immunosuppressed host, the odds for a rare clinical manifestation can be predicted to become a significant complication affecting the long-term perspective of otherwise successful kidney transplantation. Further studies are needed to explore the relative incidence and clinical importance of such alternative events.

Acknowledgements

This study was supported through a grant by the Krebsforschung Schweiz (KFS-3059-08-2012; PubMed Rentsch, Ruiz, Bubendorf). Further, we would like to thank Christine Bauer, Heike Püschel, Thomas Lorber and Darius Juskevicius for their support.

Statement of author contribution

Rämö, Nägele, Müller, Perrina, Vlajnic, Balitzki and **Ajuh** carried out the experiments.

Müller, Rämö, Nägele, Ribi, Quagliata, Ruiz, Rentsch, Balitzki and Gossert performed data analyses and interpretation.

Wetterauer and Grobholz acquired samples.

Müller, Rämö, Gosert and **Ajuh** wrote the manuscript.

Rentsch, Bubendorf, and Hirsch reviewed the manuscript.

Rentsch, Ruiz, Bubendorf and Hirsch designed the study.

Supplementary files

- Supplement, including supplementary text and tables:
  - Table S1: Summary of histopathological results.
  - Table S2: NGS-results of 10y ptx UC sites.
  - Table S3: Tabular summary of STR results.
- Figure S1: Overview over disease progression and cancer distribution 10y ptx.
- Figure S2: Sorting and arrayCGH profiles of cancer sites 8y and 9y ptx.
- Figure S3: Sorting profiles of UC sites 10y ptx.
- Figure S4: Detection of BKPyV integration site in chromosome 1 by NGS.
- Figure S5: Coverage of whole genome and targeted amplicon sequencing of BKPyV.

References

1. Vajdic CM, van Leeuwen MT. Cancer incidence and risk factors after solid organ transplantation. *Int J Cancer* 2009; **125**: 1747-1754.
2. Hirsch HH, Babel N, Comoli P*, et al.* European perspective on human polyomavirus infection, replication and disease in solid organ transplantation. *Clin Microbiol Infect* 2014; **20 Suppl 7**: 74-88.
3. Weinreb DB, Desman GT, Amolat-Apiado MJ*, et al.* Polyoma virus infection is a prominent risk factor for bladder carcinoma in immunocompetent individuals. *Diagn Cytopathol* 2006; **34**: 201-203.
4. Harris KF, Christensen JB, Imperiale MJ. BK virus large T antigen: interactions with the retinoblastoma family of tumor suppressor proteins and effects on cellular growth control. *J Virol* 1996; **70**: 2378-2386.
5. Ruiz C, Lenkiewicz E, Evers L*, et al.* Advancing a clinically relevant perspective of the clonal nature of cancer. *Proc Natl Acad Sci U S A* 2011; **108**: 12054-12059.
6. Dumoulin A, Hirsch HH. Reevaluating and optimizing polyomavirus BK and JC real-time PCR assays to detect rare sequence polymorphisms. *J Clin Microbiol* 2011; **49**: 1382-1388.
7. Drachenberg CB, Hirsch HH, Papadimitriou JC*, et al.* Polyomavirus BK versus JC replication and nephropathy in renal transplant recipients: a prospective evaluation. *Transplantation* 2007; **84**: 323-330.
8. Henriksen S, Tylden GD, Dumoulin A*, et al.* The human fetal glial cell line SVG p12 contains infectious BK polyomavirus. *J Virol* 2014; **88**: 7556-7568.
9. Davis MW. ApE - A plasmid Editor. http://biologylabs.utah.edu/jorgensen/wayned/ape/ 2016.
10. Gosert R, Rinaldo CH, Funk GA*, et al.* Polyomavirus BK with rearranged noncoding control region emerge in vivo in renal transplant patients and increase viral replication and cytopathology. *J Exp Med* 2008; **205**: 841-852.
11. Bethge T, Hachemi HA, Manzetti J*, et al.* Sp1 sites in the noncoding control region of BK polyomavirus are key regulators of bidirectional viral early and late gene expression. *J Virol* 2015; **89**: 3396-3411.
12. Leboeuf C, Wilk S, Achermann R*, et al.* BK Polyomavirus-Specific 9mer CD8 T Cell Responses Correlate With Clearance of BK Viremia in Kidney Transplant

Recipients: First Report From the Swiss Transplant Cohort Study. *Am J Transplant* 2017.

13. Yan L, Salama ME, Lanciault C*, et al.* Polyomavirus large T antigen is prevalent in urothelial carcinoma post-kidney transplant. *Hum Pathol* 2016; **48**: 122-131.
14. Lawrence MS, Stojanov P, Polak P*, et al.* Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 2013; **499**: 214-218.
15. Starrett GJ, Marcelus C, Cantalupo PG*, et al.* Merkel Cell Polyomavirus Exhibits Dominant Control of the Tumor Genome and Transcriptome in Virus-Associated Merkel Cell Carcinoma. *MBio* 2017; **8**.
16. Papadimitriou JC, Randhawa P, Rinaldo CH*, et al.* BK Polyomavirus Infection and Renourinary Tumorigenesis. *Am J Transplant* 2016; **16**: 398-406.
17. Kenan DJ, Mieczkowski PA, Burger-Calderon R*, et al.* The oncogenic potential of BK-polyomavirus is linked to viral integration into the human genome. *J Pathol* 2015; **237**: 379-389.
18. Kenan DJ, Mieczkowski PA, Latulippe E*, et al.* BK Polyomavirus Genomic Integration and Large T Antigen Expression: Evolving Paradigms in Human Oncogenesis. *Am J Transplant* 2016. 19. Holley T, Lenkiewicz E, Evers L*, et al.* Deep clonal profiling of formalin fixed paraffin embedded clinical samples. *PLoS One* 2012; **7**: e50586.
19. Holley T, Lenkiewicz E, Evers L*, et al.* Deep clonal profiling of formalin fixed paraffin embedded clinical samples. *PLoS One* 2012; **7**: e50586.
20. Drachenberg CB, Hirsch HH, Papadimitriou JC*, et al.* Polyomavirus BK versus JC replication and nephropathy in renal transplant recipients: a prospective evaluation. *Transplantation* 2007; **84**: 323-330.
21. Dumoulin A, Hirsch HH. Reevaluating and optimizing polyomavirus BK and JC real-time PCR assays to detect rare sequence polymorphisms. *J Clin Microbiol* 2011; **49**: 1382-1388.
22. Gosert R, Rinaldo CH, Funk GA*, et al.* Polyomavirus BK with rearranged noncoding control region emerge in vivo in renal transplant patients and increase viral replication and cytopathology. *J Exp Med* 2008; **205**: 841-852.

References 19 – 22 are cited in the supplementary text.

Figures

1 - Morphological and genomic characterization of cancer sites 10y ptx.

2 - BKPyV integration site

3 - Characterization of the BKPyV UC genome.

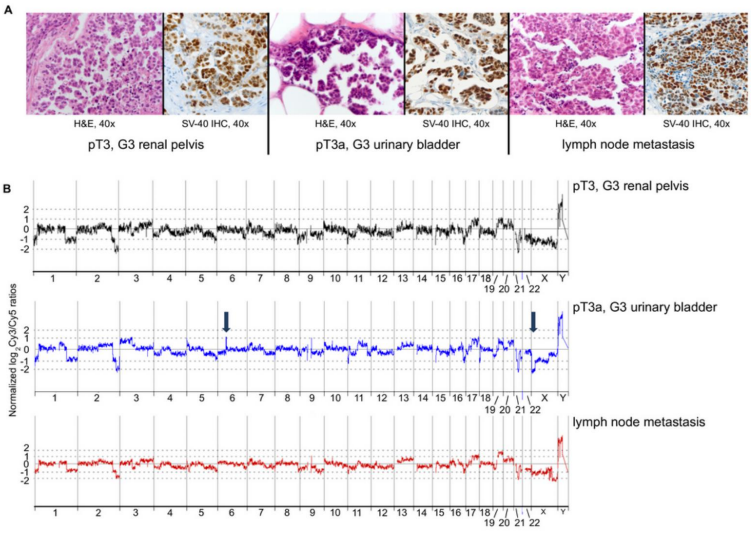


Figure 1: Morphological and genomic characterization of cancer sites 10y ptx. **A.** H&E-stainings and SV40-LTag IHC of micropapillary UC of the kidney transplant recipient at 40x magnification. The morphological picture indicates a close relationship of the different sites at 10y ptx and differ clearly form the previous tumour appearances (not shown). **B**. ArrayCGH profiles with copy number alterations of aneuploid cancer populations 10y ptx. Arrows indicate differences. X-axis: chromosome number; Y-axis: Normalized $\log_2$ Cy3/Cy5 ratios (relative count of DNA-probe signals). The 10y ptx samples appear highly related. In contrast, there was no evidence of clonal relationship to the previous tumour appearances (not shown).

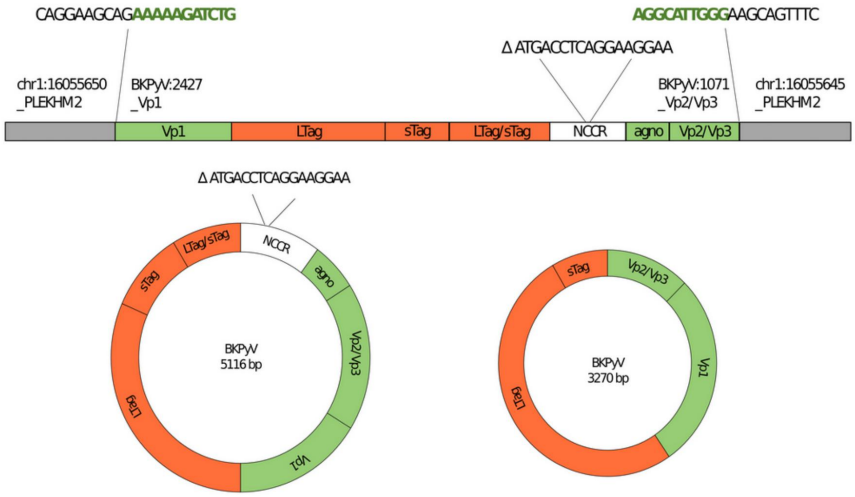


Figure 2: BKPyV integration site. Integrated and episomal BKPyV genome maps in urothelial cancer. The integrated genome and the breakpoints in chromosome 1 were located in the PLEKHM2 gene as identified by whole-genome NGS. The breakpoint sequences are shown above and colored in black (Chr 1) and green (BKPyV). In addition, a shorter and a larger episomal BKPyV genome of 5116 bp and3270 bp was identified, all confirmed by specific PCR.

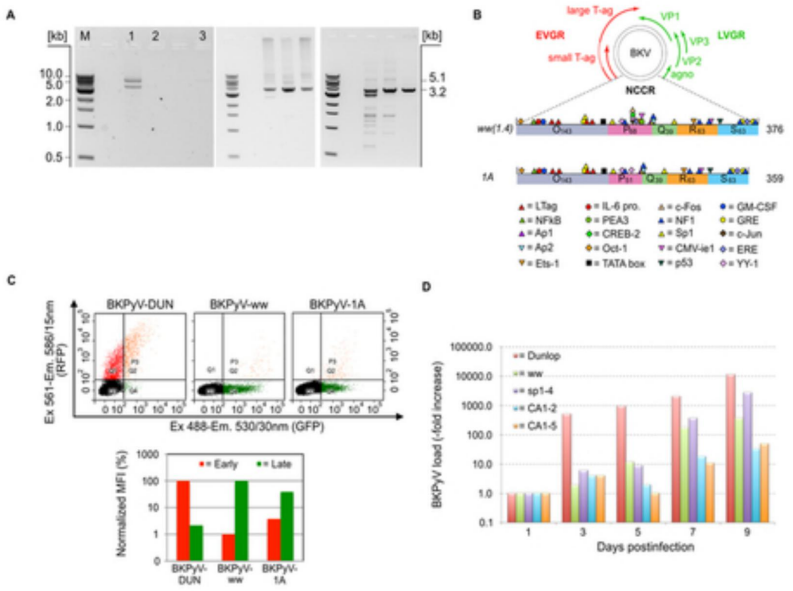


Figure 3: Characterization of the BKPyV UC genome. **A.** Amplification using outward primers overlapping in the EcoR1 site of the VP1 gene yields a full-length BKPyV DNA genome (5116 bp) bearing small deletion in the non-coding control region and a smaller sub-genomic fragment (3270 bp) lacking the amino terminal VP2, the NCCR, and the amino-terminal LTag region fragment. Lanes: M, DNA marker in kilobase pairs; 1, 50ng of UC tissue DNA; 2, negative template control; 3, weak BKPyV-positive urine as control Amplification chromosome 1 integrated BKPyV genome of 3786 bp by breakpoint-specific primers as confirmed by Sanger sequencing (supplement data S4). Lanes: M, DNA marker in kilobase pairs; 4-6, UC tissue DNA amplified by three different pairs of breakpoint-specific primers; 7-9, UC tissue DNA amplified by one pair of breakpoint-specific primers with a gradient PCR (61°C, 63°C and 65°C). **B**. Schematic comparison of the transcription factor binding sites in the BKPyV archetype NCCR (ww1.4) of 376 bp, and the BKPyV 1A NCCR (359 bp) detected in the urothelial cancer shows the deletion of 17 bp in the P-block P41-P57;EVGR, early viral gene region; LVGR, late viral gene region.**C**. NCCR-driven expression of EVGR (red fluorecence; dsRed) and LVGR (green; EGFP) in the bi-directional reporter vector pHRG1 following transfection into HEK293 cells (top panels: flow cytometry; bottom panels: normalized mean fluorescence intensity). **D**. Comparison of viral replication in Vero cells after transfection of the indicated BKPyV NCCR variants: Dunlop strain,

archetype WW(1.4); WW(sp1-4) mutant; and two independent recombinant clones bearing the cancer NCCR CA1-2 and CA1-5.

SUPPLEMEMNTARY INFORMATION

1 – Supplementary PATIENT AND METHODS

2 – Supplementary RESULTS

3 – Supplementary FIGURES

4 – Supplementary TABLES

1-Supplementary PATIENT AND METHODS

Patient

The recipient suffered from end stage renal disease due to Alport-syndrome. The donor's HLA-status was: blood type: A+; HLA-A: 1, 2; HLA-B: 8, 27; HLA-DR: 13, 7. In comparison, the recipient's status was: blood type A+; HLA-A: 24, 26; HLA-B:44, 50; HLA-DR: 13, 7. The immunosuppressive therapy consisted of tacrolimus, azathioprine, and prednisone for the first 6 months post-transplant (ptx) followed by tacrolimus and azathioprine. 3 months ptx a tissue biopsy was SV40-IHC negative. No further SV40-IHC staining on non-cancerous tissue was performed. 6 months ptx decoy cells were found in the patient's urine (11 /10 High Power Field 400x). Previous urine specimens did not show any decoy cells. In summary, there was no evidence for BKPyV-associated nephropathy (BKPyVAN). Thirteen years after transplantation, the patient is alive but requires continued renal replacement therapy by dialysis. PET/CT scans show no evidence of the urothelial cancer.

Cell nuclei extraction and DNA content-based cell sorting

For the purpose of dissecting the intratumour heterogeneity within one tissue-sample we performed DNA based cell sorting. Formalin fixed and paraffin embedded (FFPE) tumour specimens were analyzed morphologically with the Axioskop 2 plus microscope (Zeiss, Oberkochen, Germany). Biopsy samples were fixed in 10% buffered formalin and embedded in paraffin according to standard operating procedures. Cell nuclei extraction was done as previously described [19]. The extracted nuclei were separated by DNA content based on sorting using the Influx (Becton-Dickinson, San Jose, CA) cytometer with ultraviolet excitation. DNA content

and cell cycle were analyzed using the software program MultiCycle (Phoenix Flow Systems, San Diego, CA). Sorted nuclei were digested overnight in 180 µl of Incubation Buffer (Promega Corporation, Madison, WI) and 20 µl of Proteinase K (Promega Corporation, Madison, WI) at 56°C, 650 rpm for genomic DNA extraction.

DNA extraction

Total genomic DNA from sorted populations was extracted with MaxWell 16 FFPE Plus LEV DNA Purification Kit (Promega Corporation, Madison, WI) according to the kit protocol. For NGS purposes, DNA was also directly extracted from the matched FFPE tissues without sorting. Two to three 25 µm sections were placed into 2 ml microtubes. Samples were washed three times with 1 ml Xylene for 5 min to remove the remaining paraffin, following rehydration two times in 1 ml 100 % EtOH, which was then evaporated for 2 min in 26°C. 180 µl of Incubation Buffer (Promega Corporation, Madison, WI) and 20 µl of Proteinase K (Promega Corporation, Madison, WI) were added. Samples were digested overnight at 56°C, 650rpm and DNA was extracted with MaxWell 16 FFPE Plus LEV DNA Purification Kit (Promega Corporation, Madison, WI) according to the kit protocol. Because of impurities of DNA of the 2011 and 2012 samples were purified with Genomic DNA Clean and Concentrator Kit (Zymo Research, Irvine, CA) before using these samples for next-generation-sequencing (NGS).

Array comparative genomic hybridization (aCGH)

In order to investigate copy-number changes within the different tumour populations we performed aCGH. Genomic alterations betweensamples were evaluated with Agilent SurePrint G3 HMN CGH 4x180 K Oligo Microarray Kit (Agilent Technologies, Santa Clara, CA) according to the manufacturer's protocol. 40-100 ng DNA of each sorted population was subjected to the arrays and normal Female Human Genomic DNA (Promega Corporation, Madison, WI) was used as a reference DNA during aCGH. Reference DNA was digested with DNAse I. References and samples were labelled with Cy-3 dUTP and Cy-5 dUTP respectively, using a BioPrime Array Kit (Invitrogen, Carlsbad, CA). Filtering, hybridization and washing were carried out according to manufacturer's protocol (Agilent Technologies, Santa Clara, CA, USA).

Short tandem repeat analysis (STR)

DNA was purified using the Genial® DNA Kit according to the provider's protocol for isolation of total DNA. The DNA was eluted in 30 µl of low TE-buffer. All profiles were genotyped with the PowerPlex® ESI 17 Kit (Promega Corporation, Madison, WI). The kit is based on a five dye technology (blue: 6-FAM™, green: VIC®, yellow: NED™, red: PET®, orange: LIZ®) and contains 16 autosomal STR loci and AMEL (blue: D10S1248, vWA, D16S539, D2S1338; *green*: AMEL, D8SS1179, D21S11, D18S51; yellow: D22S1045; D19S433, TH01; FGA; red: D2S441, D3S1358, D1S1656, D12S391, SE33; orange: internal standard). Each run was performed on an Applied Biosystems Genetic Analyzer 3500 (Life Technologies, Carlsbad, CA, USA) with injection voltage 3 kV and injection time 10 seconds. Results were then analyzed with the GeneMapper IDX Software (Version 1.4). The peak threshold was 50 rfu (relative fluorescence units). In order to define the recipient's profile available FFPE tissue from the recipient's innate kidney was used. No non-transplanted donor tissue was available. To define the donor's STR-profile we took healthy FFPE-tissue in adequate distance from the tumour from the allograft. The extracted DNA from the recipient's innate kidney was diluted 1:100 for further analyses.

Next generation sequencing (NGS)

The panel used for targeted sequencing includes 207 amplicons covering mutations from 50 oncogenes and tumour suppressor genes. Chiploading was assisted by the IonChef-System (Life Technologies, Carlsbad, CA, USA). For Sequencing we used the LifeTech Personal Genome Machine (PGM) (Life Technologies, Carlsbad, CA, USA). Sequencing data was analyzed using the IonReporter™ software pipeline (Life Technologies, Carlsbad, CA, USA). Manual review was performed on called mutations in order to reduce the number of false positive calls.

2-Supplementary RESULTS

Histopathology

The morphological picture of the early tumours from 8 (pTa, high-grade) and 9 years ptx (pT1, high-grade) differs from the one of the three sites obtained 10 years ptx (pT3, G3). In the 8y- and 9y-ptx tissue samples, we saw an ordinary high-grade urothelial carcinoma. The 8y-ptx tumour was papillary while the 9y-ptx tissues also had a solid component. However, the 10y-ptx samples showed a micropapillar morphology,

clearly differing from the previous tumour appearance. The histopathological findings of all tumour manifestations are summarized in Table S1.

DNA content based cell sorting of tumour-specimen and aCGH

Fluorescent activated flow sorting of cell nuclei from tumour tissue based on DNA content indicated that all analyzed tumour samples were composed of a diploid and an aneuploid tumour cell population. We further investigated the aneuploid populations by arrayCGH. The aneuploid populations from 8y- and 9y-ptx displayed distinct genomic differences with each population harboring several private amplifications and deletions (Figure S2). Consequently, there was no evidence of clonal relationship among these tumours. Furthermore, both aneuploid populations from 8y- and 9y-ptx clearly differed from the three aneuploid 10y cancer samples suggesting that the 8y, 9y and 10y-ptx tumours each represent genomically independent cancer manifestations. Conversely, the aneuploid populations sorted from the pT3 renal pelvis, the pT3 bladder and the metastatic lymph node manifestations at 10y-ptx shared a common profile. The only exceptions were a private amplification at 6p12.3 and a private deletion at Xp22.33 - Xp22.11 in the bladder manifestation of the 10y-ptx tumour (see Figure S3 + Figure 1B). Regarding the copy number variations (CNV) of relevant tumour suppressor genes and oncogenes in the 10y ptx samples, we found low-level gains in *CTNNB1*, *ERBB2*, *GNAS* and *SRC*. ArrayCGH did not show any deletions of tumour suppressor genes. All microarray files have been deposited at the National Center for Biotechnology Information Gene Expression Omnibus (accession number GSE90778).

Short tandem repeat analysis (STR)

DNA-amplification of tumour samples from year 8 and 9 was not successful presumably due to poor DNA quality. However, bulk-tissue cuts of all 10y cancer samples as well as several healthy tissue samples fromthe donor and the recipient were suitable for STR analysis. We could define a distinct recipient's healthy profile in all 16 STR-systems. Due to contamination with recipient derived blood cells or ingrown stromal cells we expected the donor's healthy specimen to show a mixed profile. As such, main-profile (MP) and a side-profile (SP) were discriminated in the donor's healthy specimen. All detected alleles of the SP were identical with the recipient's profile. All 10y-ptx cancer specimens showed a mixed profile consisting of a donor

derived MP and a recipient derived SP. The detailed STR-results are depicted in the Table S3.

In summary, the STR analysis demonstrated that the 10y-ptx cancer manifestations identified in the allograft kidney pelvis (pT3), in the patient's own bladder (pT3a / CIS) and in a single pelvic lymph node metastasis, were all of donor-origin.

Next generation sequencing

Polymerase chain reaction (PCR) based DNA-amplification of tumour samples from 8y- and 9y-ptx was not successful (see above). However, NGS revealed 18 genomic nucleotide variants in all three 10y-ptx cancer samples. According to the University of California, Santa Cruz (UCSC) common single nucleotide polymorphism (SNP) database, 15 of them were reported SNPs. In none of the investigated 50 cancer genes somatic mutations were identified. The remaining three nucleotide variants are not yet classified SNPs or known somatic mutations by publicly available databases. Of these, one did not pass a threshold of 10% allele ratio and therefore has a high odd for being an amplification-artefact. The two other nucleotide variants were: *KDR* / InDel chr4:55962545 and *Akt1* / c.138C>A. Further details are shown in Table S2.

Characterization of the BKPyV UC

The BKPyV viral load in the 10y-ptx UC tissue samples was determined by real-time PCR as 1 BKPyV copy per diploid human cell using the human acetylase gene as normalization reference [20, 21] . No BKPyV DNA could be detected in healthy tissue from the same time point. To characterizethe BKPyV UC genome, two primers targeting the conserved EcoR1 site in the BKPyV VP1 capsid gene were used in an outward orientation. As shown in Figure 2A, two fragments corresponding to a full-length BKPyV genome of 5,116 bp and a shorter fragment of 3,270 bp were generated, whereas no products were obtained from unaffected tissue DNA. Sequencing of the larger fragment revealed a complete BKPyV genome bearing a total of 49 single nucleotide differences from the BKPyV archetype sequence in LTag, sTag, VP1 and VP2 (data not shown), with the exception of a 17 bp deletion in the P-block of the NCCR (Figure 2B). The shorter fragment represented a truncated BKPyV genome due to a 1.9 kb deletion from the small T-antigen coding region at bp position 4668 to the VP2 coding region at position 1398 bp, which removed the N-terminal coding region of LTag and sTag, and the NCCR, and thereby could not account for the LTag-positive

staining. Also removed were the entire agnoprotein and N-terminal parts of the VP2, while the remaining part shared all 36 single nucleotide differences present the complete BKPyV UC genome. These results were confirmed by direct nested PCR and sequencing from the UC tissue. Given the unique BKPyV UC del-NCCR and the well documented role of NCCR rearrangements for constitutive up-regulation of the BKPyV EVGR expression at the expense of LVGR expression [22], a functional analysis was performed by inserting the BKPyV UC-NCCR into the previously described bi-directional reporter construct carrying the red fluorescent protein (dsRed) and the green fluorescent protein (EGFP) as markers for EVGR and LVGR expression, respectively (Figure 2C). Following transfection of human embryonic kidney (HEK) 293 cells and single-cell analysis by flow cytometry, the UC 1A-NCCR conferred a moderate activation of EVGR expression (red) of 2 – 3 fold as compared to the archetype BKPyV ww-NCCR carrying the full-length sequence. To investigate effects of the UC-NCCR deletion on viral replication, recombinant BKPyV genomes were generated that only differed in the respective NCCR and then compared for replicative capacity following infection of hRPTEC cells (Figure 2D). The results demonstrate that the replicative capacity of the BKPyV UC-NCCR strain was significantly impaired compared to the archetype ww-NCCR. In contrast, the laboratory adapted Dunlop-NCCR or a recently identified point mutant *sp1-4*-NCCR showed a much higher replication rate (Figure 2D). Taken together, the molecular characteristics of the BKPyV-UC strain suggest that this variant significantly activates EVGR expression and contributes to LTag expression, but that this unique 17bp deletion impairs progression into the late viral life cycle and thereby offsets efficient lytic replication.

3 - Supplementary FIGURES

Figure S1 – Overview over disease progression and cancer distribution 10y ptx.

Figure S2 – Sorting and arrayCGH profiles of cancer sites 8y and 9y ptx

Figure S3 – Sorting profiles of UC sites 10y ptx

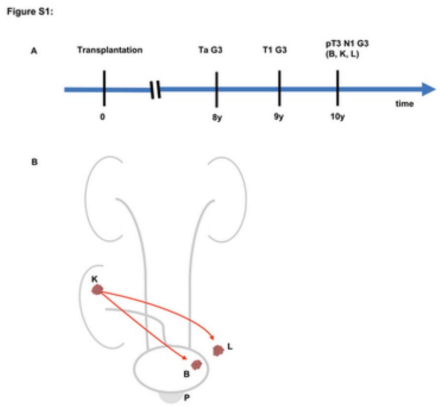


Figure S1: A. Timeline (years ptx) of clinical manifestations of UC in the kidney transplant recipient. **B.** Schematic illustration of cancer localizations and the assumed metastatic spread from the donor kidney UC. Cell-bulk indicates UC localizations. Arrows indicate assumed metastatic spread. (B: Bladder, left bladder side / apex. L: Lymphnode, left *Arteria iliaca externa.* K: Kidney, pyelon of allograft. P: Prostate, adenocarcinoma, Gleason 6 (3+3), max. diameter 7 mm).

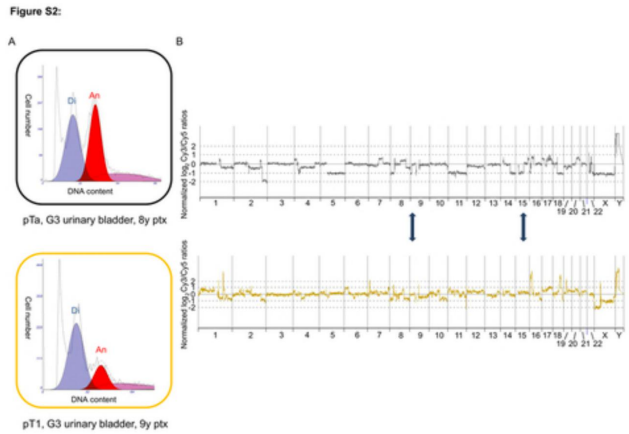


Figure S2: A. Sorting profiles. Di: Diploid population; An: Aneuploid population; X-axis: DAPI-fluorescence units x1000 – linear scale; y-axis: cell nuclei count in total numbers; blue colour: G_1/G_2-phase diploid population; red colour: G_1/G_2-phase

aneuploid population; violet colour: cells in s-phase. **B.** ArrayCGH profiles with copy number alterations of aneuploid cancer populations 8y and 9y ptx. The two samples appear unrelated; arrows indicate examples of relevant differences. X-axis: chromosome number; Y-axis: Normalized $\log_2$ Cy3/Cy5 ratios (relative count of DNA-probe signals).

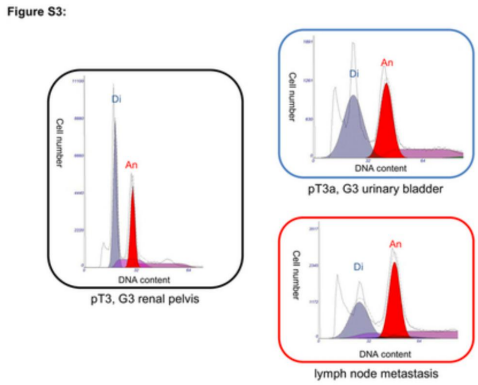
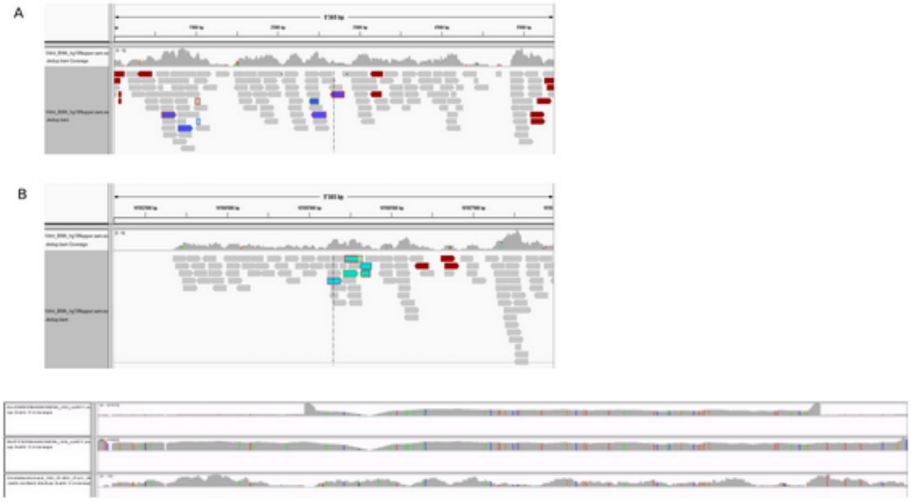


Figure S3: Di: Diploid population; An: Aneuploid population; x-axis: DAPI-fluorescence units x1000 – linear scale; y-axis: cell nuclei count in total numbers; blue colour: G_1/G_2-phase diploid population; red colour: G_1/G_2-phase aneuploid population; violet colour: cells in s-phase.



4 - Supplementary TABLES

Table S1: Summary of histopathological results.

Table S2: NGS-results of 10y ptx UC sites.

Table S3: Tabular summary of STR results.

Table S1: Summary of histopathological results.

Sample	TNM	Morphology	SV-40 Status
Bladder 8y	pTa, high-grade	non-invasive papillary UC	positive
Bladder 9y	pT1, high-grade	mainly solid UC	positive
Allograft-kidney 10y	pT3, high-grade, L1, V1, Pn0	partly necrotic, mainly micropapillary UC	positive
Bladder 10y	pT3a, pN1 (1/5), G3, L1, V0, Pn1	mainly micropapillary UC	positive
Lymphnode metastasis 10y	not available	micropapillary UC	positive

Table S2: NGS-results of 10y ptx UC sites.

		mutant allele ratio			No. of reads		
gene	mutation	kidney	bladder	lymphnode	kidney	bladder	lymphnode
ERBB4	chr2:212812097T>C	0.056	0	0.0065	339	498	153
PIK3CA	c.3100G>C	0.0895	0.0014	0	1999	2799	929
PIK3CA	chr3:178917005A>G	0.02	0.0844	0.0677	750	1007	310
PDGFRA	c.1701A>G	1	1	1	1195	1957	513
PDGFRA	c.2472C>T	0.4905	0.443	0.3766	>2000	>2000	1240
KIT	c.1621A>C	0.034	0.0745	0.146	1999	>2000	1315
KDR	InDel chr4:55962545	0.4566	0.7807	0.605	1520	>2000	524
KDR	c.1416A>T	0.5508	0.6131	0.5972	1910	1998	648
KDR	chr4:55980239C>T	1	1	0.9887	659	990	266
APC	c.4479G>A	0.0725	0.1939	0.2377	1985	1991	1296
CSF1R	chr5:149433596TG>GA	1	1	1	376	1193	625
EGFR	c.2361G>A	0.5303	0.4796	0.5311	594	1203	563
RET	c.2307G>T	0.4992	0.5298	0.4303	1238	1997	1636
RET	c.2712C>G	0.5207	0.3983	0.372	822	1996	1148
HRAS	c.81T>C	0.9325	0.8744	0.8401	696	1999	1676
FLT3	chr13:28610183A>G	1	1	0.9906	1998	1997	1909
AKT1	c.138C>A	0.4949	0.3569	0.466	681	1998	1219
TP53	c.215C>G	0.941	0.9715	0.9435	763	>2000	885
DERL3	chr22:24176287G>A	0.4773	0.6243	0.4764	484	1171	741

: Single-nucleotide polymorphism (SNP) / multi-nucleotide polymorphisms (MNP)
: mutation of unknown qualityFirst column: gene name and position of SNP/MNP/unknown mutation; second column: number shows allele ratio (mutant reads / reads analogous to reference); third column: total number of reads in according area.

Table S3: Tabular summary of STR results.

Marker	Recipient Control	Donor Control	Bladder 10y	LN 10y	Kidney 10y
D3S1358	17	**15,16,**(17)	**15**,(**16**,17)	**15**,(**16**,17)	**15**,(**16**,17)
D19S433	13/15.2	**14**,(13,15.2)	**14**,(13,15.2)	**14**	**14**,(15.2)
D2S1338	22/23	**20,24**	**20**,(22,23)	**20**	**20**,(22,**24**)
D22S1045	11	11,**15,16**	11,**15,16**	n.r.	11,**15,16**
D16S539	12	**11,13**(12)	**11**,12	**11,13**,(12)	**11,13**(12)
D18S51	13/17	**14,17,**(13)	**14,17**,(13)	**14,17**,(13)	**14,17,**(13)
D1S1656	15/16	**16,16.3**	**16.3**,(15,**16**)	**16.3**,(**16**)	**16.3**,(15)
D10S1248	13/14	**16**	**16**,(13,14)	**16**,(13)	**16**,(13,14)
D2S441	11/12	**14**	11,**14**,(12)	(12)	**14**, (11)
TH01	7/9	**6,7**	**7**,(9)	**7**	**7**,(**6**,9)
vWA	18/19	**14,16**,(18,19)	**14,16**, (18,19)	**14,16**,	**14,16**,(18,19)
D21S11	30/31.2	**28,30.2,**(30)	**30.2**,(30,31.2)	**30.2**	**30.2**,(**28**,30, 31.2)
D12S391	17/19	**16,19**,(17)	**16,19**,(17)	**16,19**,17	**16,19**,(17)
D8S1179	12/13	**10,13**,(12)	**10,13**,(12)	**10,13**,(12)	**10,13**,(12)
FGA	21/24	**22**,(21,24)	21,**22**,(24)	**22**,(21)	**22**,(21,24)
SE33	15/30.2	**18,19**,(15)	**18,19**, (15,30.2)	**18,19**	**18,19**,(15,30.2)
Amelogenin	XY	XY	XY	XY	XY

The first column shows the name of DNA-probe (marker). Numbers and letters in columns two to six show alleles (two per chromosome) where DNA-probes annealed. underlined numbers: recipient related alleles; **bold numbers**: donor related alleles; (): side profile's alleles; n.r.: no signal reported.

The molecular characterization of the BKPyV full-length genome revealed an NCCR with 17 bp deletion located in a region rarely affected in other naturally occurring BKPyV variants. Like most rearranged BKPyV NCCR variants from patients, this 17 bp deletion increases EVGR expression in expense of LVGR expression, and may contribute to the LTag expression detected in the UC. However, this deletion prevented progression of the viral replication cycle into the late phase, therefore obstructing progeny production. Integration of the viral genome into the host genome or LTag interaction with tumor suppressors can result to tumorigenesis. There was evidence of viral genome integration into the host's chromosomes.

Altogether, this study reports an oncogenic role of a rare BKPyV variant with a unique 17 bp deletion in the NCCR. This deletion activated EVGR expression in expense of LVGR expression however, a reduced viral replication probably due to integration in the cancer genome. Larger number of cases would be needed to confirm these findings.

6 DISCUSSION

We showed that the HPyV-NCCRs differ in sequences, length, number of LTag-binding sites and TFBS. These differences in the NCCRs and cellular transcription factors make-up are essential for the differences of the HPyV-NCCRs activity in the cell lines. Interestingly, some HPyVs demonstrated high EVGR or LVGR expression in cell lines derived from tissues where they were hypothesized to have tropism for, suggesting tissue tropism in the respective body compartment where those cells were derived from. LTag is vital for HPyV genome replication and LVGR expression (59, 262, 345)**.** Thus, the effect of exogenous Tags expression on HPyVs EVGR and LVGR expression was characterized. SV40 Tags were able to activate both EVGR and LVGR expression from the NCCR of all HPyVs. Additionally, MCPyV cognate Tags activated EVGR expression from the MCPyV NCCRs, expression indicating the bidirectional reporter recapitulates an essential NCCR response described for the viral life replication cycle (25, 48, 340, 346). Replication and viral production of MCPyV and BKPyV in HEK293-4T cells expressing both MCPyV Tags and SV40 Tags have been reported (347). Furthermore, BKPyV and JCPyV demonstrated productive replication in HEK293TT cells, expressing SV40 LTag (306). Taken together, these and our results support the hypothesis that cell lines expressing cognate Tags of respective HPyVs may support propagation of respective HPyVs. We also noticed that HPyVs' number of LTag binding sites correlated with EVGR expression in HEK293T cells expressing SV40 Tags, indicating a role for LTag-binding sites in HPyV EVGR and LVGR expression. The role of BKPyV LTag binding sites in EVGR and LVGR expression, which can be extrapolated to other HPyVs, was investigated for BKPyV as a model HPyV in **Section 5.3.** We showed that rr-NCCR variants of HPyV7 (HPyV7-PITT1 and -2) and HPyV9 (HPyV9-UF1) from patients displayed higher EVGR expression compared to their respective archetype, indicating that rearrangements in the NCCR of novel HPyVs may also result in gain of function for EVGR expression and probably higher viral replication rate as demonstrated for clinical variants of BKPyV and JCPyV (26, 34). This result emphasizes the fact that the NCCR is the pathogenicity determinant of HPyVs. Interestingly, the higher EVGR expression pattern of the novel HPyV (rr)-NCCRs of HPyV7 and HPyV9 differs in the cell lines, indicating that the host cell's TFs make-up

are also implicated in the NCCR-driven expression. More so, a BKPyV variant from a urothelial cancer harboring a unique 17 bp deletion in the NCCR led to a gain of function for EVGR expression (section 5.4). This BKPyV variant expresses abundant LTag in the urothelial cancer, but surprisingly showed an impaired viral replication *in vitro*. This indicates that not all NCCR variants with elevated EVGR reporter expression correspond to an increase in viral replication, especially in the context of tumorigenesis. Other components of the viral-host cells interaction such as viral-receptor entry or viral release may be involved. This impaired *in vitro* viral replication may also be due to viral integration into the host genome as reported not only in the urothelial cancer case study in **section 5.5,** but also for MCPyV integrated into the genome of Merkel cell carcinoma (348). The data are important for identifying host and viral factors vital for the replication of HPyVs in cell culture and comprehending secondary host cell tropism for HPyV beyond HPyV entry. This will go a long way in understanding the basic biology of these important human viruses. However, further studies are needed to determine the actual replication of some of the HPyV-NCCRs with high activity in cell lines.

In the following part of the thesis we focused on the regulation of HPyVs using BKPyV as a model HPyV. A role of LTag binding motifs in EVGR and LVGR fold expression in cell lines constitutively expressing LTag was suspected, although a role of each particular LTag binding site is unknown. LTag-binding motif of SV40 have been greatly studied (59)**.** The key feature of LTag in the NCCR of SV40 is its ability to bind the motifs GRGGC located at the ori, to regulate viral transcription and promote viral genome replication (59, 262, 345). All HPyVs including SV40 share a common genome architecture and all contain LTag binding sites in their NCCRs. However, there are some important features in which HPyVs differ. For instance, their secondary host cell tropism (NCCR activity in host cells), NCCR length, viral control elements for example, ori, EVGR and LVGR promoters, and location and number of LTag-binding sites. All HPyVs encode at least one LTag-binding motif located mostly at the ori. Since the roles of the SV40 LTag motifs are greatly known and the motifs between SV40, BKPyV and JCPyV at the ori are the same, we hypothesized that the role of the LTag binding motif in BKPyV is similar to those of SV40, which can be extrapolated to other HPyVs. The EVGR and LVGR expression of BKPyV ww-NCCR reporter with those of BKPyV ww-NCCR defective in one or

more LTag binding motifs was compared in HEK293 cells expressing no, small or large amounts of SV40 LTag. Eight different combinations of permutations were made in the BKPyV ww-NCCR LTag-binding motif (termed A, B, C, D, AC, ABC, ACD and ABCD). Our results revealed that EVGR expression of the ww-NCCR, increases in HEK239T cells and then decreases in HEK293TT cells. Mutating the LTag binding *site A* or combination of LTag binding *site A* with other LTag binding sites interestingly, abrogated this EVGR expression decrease in HEK293TT cells. This indicates that LTag may interact with LTag binding *site A* to exhibit this decrease**.** Mutation of other LTag binding sites together with LTag-binding *site A* did not affect this EVGR decrease in HEK293TT cells. With respect to the LVGR expression, all LTag-binding site mutants showed a proportional increase in LVGR expression in the presence of increasing amounts of LTag expression. Nonetheless, this proportionate LVGR expression is abrogated in mutants containing both *A* and *D* LTag-binding sites mutation. Intriguing, mutants containing either *A* or *D* LTag binding sites still exhibited the proportional LVGR expression. This result suggests that LTag-binding *sites A* and *D* may be involved in the LVGR-expression increase in the present of abundant LTag expression (HEK293TT cells) or the increase could be due to plasmid copy number effects **(Fig. 4C).** However, our preliminary reporter plasmid quantification by qPCR did not exactly correlate with the reporter constructs expression, thus indicating the expressing differences may not be due to plasmid copy number effects. The results obtained from investigating the LTag-binding sites of BKPyV can be extrapolated to other HPyV, such as JCPyV with similar LTag-binding sites organization. In this study, we analyzed the effect of SV40 LTag on the BKPyV ww-NCCR mutants defective of LTag-binding motif knowing that one cannot exclude that BKPyV cognate LTag might give different results. Nonetheless, the SV40 and BKPyV LTag have great similarity (47) and the fact that BKPyV can replicate in HEK293TT cells constitutively expressing SV40 LTag argued against this possibility (306)**.** EVGR and LVGR expression were analyzed using the bi-directional reporter vector omitting viral replication. Thus, further experiments are required to quantify the viral replication potential of the various NCCR mutants defective of LTag-binding motif by cloning them in a BKPyV-dunlop backbone and perform viral replication studies in cell culture as previously described (34)**.**

In conclusion, this thesis gave novel insights into the host and viral factors that may be vital for identifying suitable cell culture systems for HPyVs replication and prediction of HPyVs secondary host cell tropism. This will greatly enhance our understanding of the basic biology of HPyVs and development models for anti-viral testing. Finally, the effect of LTag-binding sites mutageneses suggests LTag-binding sites A and D may play a crucial role in the regulation of BKPyV`s EVGR and LVGR expression.

7 SUPPLEMENTARY MATERIAL AND METHODS

Dual reporter design and Primers

We designed a new-smaller bi-directional reporter pRG13 **(Fig. 1A)** based on the phRG reporter described in Gosert et al **(Fig. 13),** (25). The reporter lacks the SV40 early promoter, beta globulin polyadenylation signal and Hygromycin gene sequences found in the phRG **(Fig. 13**). Furthermore, a derivative reporter vector (pRG13D12) which lacks only 12-nucleotide sequence of BKPyV ww agno gene (nucleotides number 4-15) compared to pRG13 was designed. The pRG13 and pRG13D12 reporters were designed as follows: a pBI-CMV2 plasmid **(Fig. 14)** containing a bidirectional CMV promoter, a green fluorescent protein reporter gene from the *Aequorea coerulescens* jellyfish (AcGFP1), Ampicillin resistance gene, SV40 poly-adenylation signals and Colicin E1 origin of replication was ordered (clontech, Germany). The plasmid was digested with NotI and ApaI restriction enzymes (New Englands Biolabs, England). The digestion product was visualized on a 1% agarose gel (Thermo Fischer Scientific, Switzerland). A 2385 bps fragment comprising of the ampicillin resistance gene, SV40 poly-adenylation signals and Colicin E1 was purified from the gel with the nucleospin extract II columns (Macherey-Nagel, Switzerland). Next, from the phGR reporter vector (25) a 1803 bps fragment comprising of the enhanced green fluorescence protein (EGFP), BKPyV Dunlop NCCR containing the 12-nucleotide sequence and the *Discosoma striata* red fuorescent protein 2 (DsRed2) was PCR amplified with primers **(Table 3).** The PCR product was digested with ApaI and NotI and then purified with the nucleospin extract II columns. DNA concentrations were estimated using a NanoDrop 1000 spectrophotometer (Thermo Fischer Scientific, Switzerland). The 1803 bps fragment was ligated using T4 DNA ligase (New England Biolabs, England) to the ApaI/NotI digested pBI-CMV2 vector at insert/vector ratio of 3:1. Chemo-competent *Escherichia coli* DH5α cells (New England Biolabs, England) were transformed with the ligated product. Seven colonies were amplified in LB (Luria-Bertani) liquid medium (Thermo Fischer Scientific, USA) and plasmids were purified with the QIAprep Spin Miniprep Kit (Qiagen, Switzerland). Digestion of positive clones with ApaI and NotI (New

England Biolabs, England) was used to identify positive clones. For visualization, the digestion products were stained with ethidium bromide after gel electrophoresis on a 1% agarose gel. Clones with the correct insert gave band sizes of 2385 and 1803 bps. Using different forward and reverse primers (**Table 4**), the entire plasmid of one positive clone was sequenced using the 3130 Genetic Analyzer Sanger sequencer (Applied Biosystems, Switzerland). The newly designed reporter vector was called pRG13**.** The BKPyV Dunlop NCCR in the new designed reporter contained the 12-nucleotides sequence. All HPyVs NCCRs cloned in the pRG13 lack the 12-nucleotides sequence, hence the name pRG13D12. HPyV NCCRs lacking the 12-nucleotide sequences were synthesized (Eurogentec S.A, Belgium). The NCCRs were excised from the pUC57 vector using the restriction enzymes BssHII and MluI (New England Biolabs, England). Excised DNA products were visualized on 1% agarose gel and NCCR DNAs were purified with the nucleospin extract II columns. Next, the pRG13 reporter vector containing Dunlop NCCR cloned via BssHII and MluI restriction sites was digested with BssHII and MluI, and visualized on 1% agarose gel. The plasmid backbone fragment (3782 bps) was purified with the nucleospin extract II column. The concentrations of the purified DNA fragments (NCCR and plasmid backbone) were estimated using a NanoDrop 1000 spectrophotometer. The NCCR DNA fragments were ligated using T4 DNA ligase to the BssHII/MluI digested pRG13 vector at an insert/vector ratio of 3:1. Chemo-competent *Escherichia coli* DH5α cells were then transformed with the ligated product. For each of the NCCR cloned, 10 colonies were PCR-screened with HPyV NCCR-specific reverse primers and a DsRed2 forward primer (Ds_Red-F1) **(Table 5)**. Colonies carrying the right inserts gave an approximate size of 700 bps. Using the 3130 Genetic Analyzer Sanger sequencer, two colonies with the right inserts were sequenced using HPyV NCCR-specific primers **(Table 5)**. Colonies with the correct NCCR sequences was maximally amplified (maxi prep) in 400mL LB liquid medium. The plasmids were purified with the pure yield plasmid maxiprep System (Promega, Switzerland). The maxiprep plasmids were then re-sequenced.

Table 3: Primers for the PCR amplification of the 1803 bps fragment using phRG as template.

Primer names	Primer sequences
Not1_DsRed_F	CTGAGCGGCCGCCTACAGGAACAGGTGGTGG
APa1_EGFP_R	CTGAGGGCCCTTACTTGTACAGCTCGTCCATGC

Table 4: Primers for sequencing the entire pRG13 reporter vector

Primer names	Primer sequences
BKV_at_nccr_F	AAGCCTCCACACCCTTACTAC
BKV_at_nccr_R	ACTCCTTTGGCCAGTTTCCAC
Ds_Red_F_1	CATGGTCTTCTTCTGCATCAC
Ds_Red_R_1	TCCAAGGTGTACGTGAAGCAC
DsRed_F	CTATGGAAAAACGCCAGCAACG
DsRed_R	CACAAGGCCCTGAAGCTGAAG G
EGFP_F	CAGCTCGCCGACCACTACCAG
EGFP_R	GAATAAGGGCGACACGGAAATG
GFP_R1	GTGCTCAGGTAGTGGTTGTC
GFP_F1	CTGAAGTTCATCTGCACCAC
puc19_F_1	AGTGCCACCTGACGTCTAAG
puc19_R_1	CTTACAGACAAGCTGTGACCGTC
puc19_F_2	AGCTCACGCTGTAGGTATCTC
puc19_R_2	TACCTCGCTCTGCTAATCCTG
puc19_F_3	GTCTGACGCTCAGTGGAACG
puc19_R_3	GTCAGGCAACTATGGATGAAC
puc19_R_4	AGCTAGAGTAAGTAGTTCGCCAG
puc19_F_4	CAACGATCGGAGGACCGAAG
puc19_R_5	TGGTGAGTACTCAACCAAGTC
puc19_F_5	GTATCCGCTCATGAGACAATAAC

Table 5: HPyV NCCR-specific primers for sequencing HPyV NCCR reporter constructs

Primer names	Primer sequences
BKV_at_nccr_F	AAGCCTCCACACCCTTACTAC
BKV_at_nccr_R	ACTCCTTTGGCCAGTTTCCAC
JCV_at_nccr_F	TACTTCTGAGTAAGCTTGGAG
JCV_at_nccr_R	TATAACTGCCAGTGGCATGC
KIPyV_nccr_F	GCACAGTGATTGACAGTTGTG
KIPyV_nccr_R	CAGTCTTACCTGTGTCAAGGACAG
WUPyV_nccr_F	GCTAAGCATGATTGACAGTGTG
WUPyV_nccr_R	CCAAGGACGTCTCTGTTAAGC
MCPyV_nccr_F	GAGCTACCTCACTAAGGAGTG
MCPyV_nccr_R	TCTCCAGGAAATGAGTCAATGC
HPyV-6_nccr_F	CTGTGCAGGCTTACGCAACTG
HPyV-6_nccr_R	GTGGTCAACCTCAGTTATGACTG
HPyV7_nccr_F	AGGTGGAGGAGGAAATGTCTG
HPyV7_nccr_R	CTAAGTCCTTGACACAGCAC
TSPyV_nccr_F	GAGGAAGTGCCAAGCCTCAAG
TSPyV_nccr_R	CTTCCTGTGTTTGTGGCAGTATG
HPyV9_nccr_F	CAACAAGAGAAGGAGGCAAG
HPyV9_nccr_R	GAAACCGCAATGCACTTGTTTG
HPyV10_nccr_F	AGACAGAGGCTTAGGCCTC
HPyV10_nccr_R	CTTACTGTACTTTGGAGTATTTCC
STLPyV_nccr_F	CTGTCTCTAAAGAGGCTGAGG
STLPyV_nccr_R	ACATAATGCTGTTGTTGGCTCG
HPyV12_nccr_F	CAGTACAGTGTACGTGCCTGG
HPyV12_nccr_R	CTTTAGTAGATCTTCTAGAAGGTTCAGG
NJPyV-2013_nccr_F	CTGCCTAGAAGACCTTGCAGC
NJPyV-2013_nccr_R	CAGTAACTTGCTTTCTCTTGTTTGG

Table 6: Proposed genera of HPyVs including LIPyV yet to be confirmed as a HPyV, adapted from (2).

Human polyomaviruses	**GenBank accession number**
Genus Alphapolyomavirus	
MCPyV	HM011556
TSPyV	GU989205
HPyV9	HQ696595
HPyV12	JX308829
NJPyV	KF954417
LIPyV	KY404016
Genus Betapolyomavirus	
BKPyV	V01108
JCPyV	AB038249
KIPyV	EF444549
WUPyV	EF444549
Genus Deltapolyomavirus	
HPyV6	HM011563
HPyV7	HM011567
HPyV10	JX262162
STLPyV	JX463183

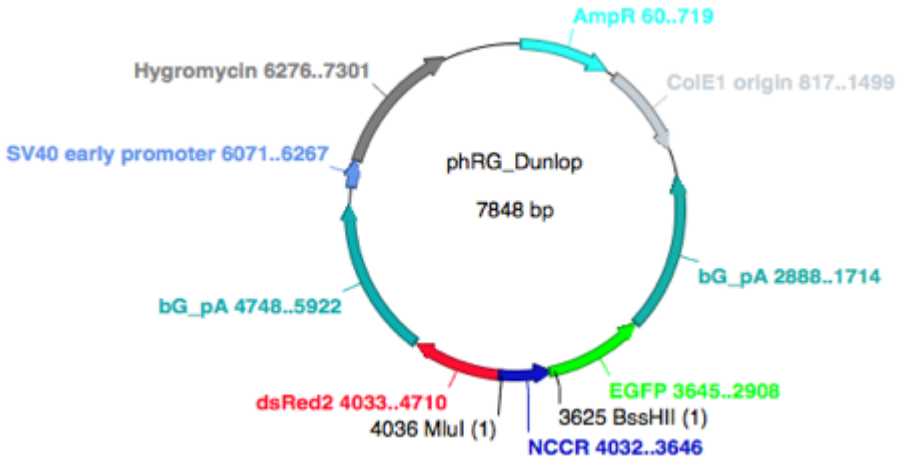


Fig. 13: Vector map of phRG. The bi-directional reporter vector phRG recapitulating the HPyV genome contains the NCCR cloned via the MluI/BssHII sites in the forward orientation (as in the viral genome), dsRed2 as a marker of EVGR expression and EGFP as marker of LVGR expression. It also contains Beta globulin poly-adenylation signal for stabilizing the dsRed2 and EGFP messages, SV40 early promoter for driving

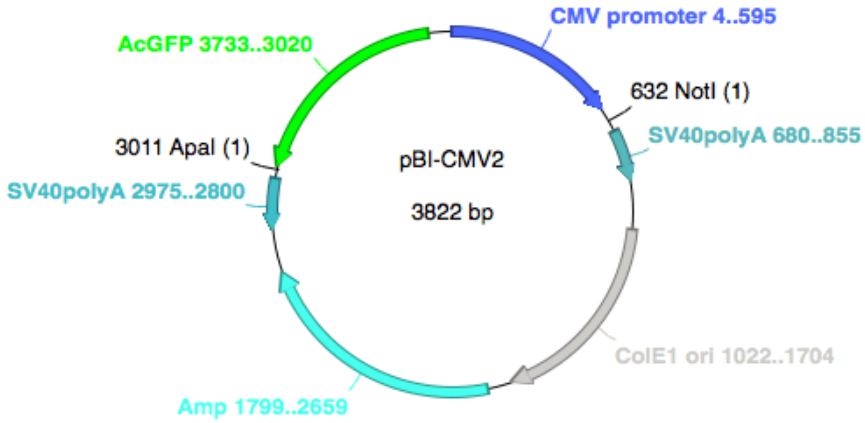


Fig. 14: **Vector map of pBI-CMV2.** The pBI-CMV2 vector contains a bi-directional CMV promoter, SV40 poly-adenylation signals for stabilizing messages, AcGFP reporter gene, Colicin E1 origin for bacteria replication and Ampicillin resistance gene for selection in bacteria.

Hygromycin gene, Hygromycin resistance gene for selection in eukaryotic cells, the ampicillin resistance gene for selection in bacteria and Colicin E1 origin for bacteria replication.

Plasmid DNA Sequencing

Dye-terminator sequencing method was used to sequence plasmid DNA. Plasmid DNA was first PCR amplified to obtain sufficient amounts of template DNA concentration. Forward and reverse amplification was performed separately. For the sequencing PCR, a master mix of 4 μL Big Dye vs. 3.1 (Applied Biosystem, Switzerland), 2 μL 5x buffer (applied Biosystem, Switzerland), 3.2 μL primer for each forward and reverse reaction (final concentration of 160 nM) was used. 9.2 μL of the master mix was pipetted into each PCR tube (Applied Biosystem, Switzerland). 400ng template DNA was added to the mix. The final volume was adjusted to 20μL with H_2O. Tubes were tightly closed with caps and transferred to the Veritti thermal cycler **(**Applied Biosystem, Switzerland**)**. The PCR reaction was run with the program SEQAL:

94° C 5 min

96° C 30 sec
55° C 15 sec } 35 cycles
60° C 4 min

cooling at 10°C

After the PCR reaction, the PCR product was purified with the NucleoSEQ columns (Macherey-Nagel, Switzerland) according to the manufacturer's protocol. In brief, the PCR sample was applied carefully in the center of the column and spin at 750x g in an Eppendorf tube to recover the purified samples. The purified sample was put in a 96-well plate (applied Biosystems, Switzerland), closed tightly with septum and put into the sequencing machine. The sequencing results were analyzed with SeqMan Pro software, version 14.0.0.

8 REFERENCES

1. Imperiale MJ, Major, E.O. Polyomaviruses. In: Knipe DM, Howley, P.M., editor. Fields Virology. 1. 5th ed. Philadelphia, PA, USA: Lippincott Williams & Wilkins, Wolters Kluwer; 2007. p. 2263-98.
2. Polyomaviridae Study Group of the International Committee on Taxonomy of V, Calvignac-Spencer S, Feltkamp MC, Daugherty MD, Moens U, Ramqvist T, et al. A taxonomy update for the family Polyomaviridae. Arch Virol. 2016;161(6):1739-50.
3. Gardner SD, Field AM, Coleman DV, Hulme B. New human papovavirus (B.K.) isolated from urine after renal transplantation. Lancet. 1971;1(7712):1253-57.
4. Padgett BL, Walker DL, ZuRhein GM, Eckroade RJ, Dessel BH. Cultivation of papova-like virus from human brain with progressive multifocal leucoencephalopathy. Lancet. 1971;1(7712):1257-60.
5. Bialasiewicz S, Whiley DM, Lambert SB, Wang D, Nissen MD, Sloots TP. A newly reported human polyomavirus, KI virus, is present in the respiratory tract of Australian children. J Clin Virol. 2007;40(1):15-8.
6. Gaynor AM, Nissen MD, Whiley DM, Mackay IM, Lambert SB, Wu G, et al. Identification of a novel polyomavirus from patients with acute respiratory tract infections. PLoS Pathog. 2007;3(5):e64.
7. Feng H, Shuda M, Chang Y, Moore PS. Clonal integration of a polyomavirus in human Merkel cell carcinoma. Science. 2008;319(5866):1096-100.
8. Schowalter RM, Pastrana DV, Pumphrey KA, Moyer AL, Buck CB. Merkel cell polyomavirus and two previously unknown polyomaviruses are chronically shed from human skin. Cell Host Microbe. 2010;7(6):509-15.
9. van der Meijden E, Janssens RW, Lauber C, Bouwes Bavinck JN, Gorbalenya AE, Feltkamp MC. Discovery of a new human polyomavirus associated with trichodysplasia spinulosa in an immunocompromized patient. PLoS Pathog. 2010;6(7):e1001024.
10. Scuda N, Hofmann J, Calvignac-Spencer S, Ruprecht K, Liman P, Kuhn J, et al. A novel human polyomavirus closely related to the african green monkey-derived lymphotropic polyomavirus. J Virol. 2011;85(9):4586-90.
11. Siebrasse EA, Reyes A, Lim ES, Zhao G, Mkakosya RS, Manary MJ, et al. Identification of MW polyomavirus, a novel polyomavirus in human stool. J Virol. 2012;86(19):10321-6.

12. Lim ES, Reyes A, Antonio M, Saha D, Ikumapayi UN, Adeyemi M, 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):295-303.
13. Korup S, Rietscher J, Calvignac-Spencer S, Trusch F, Hofmann J, Moens U, et al. Identification of a novel human polyomavirus in organs of the gastrointestinal tract. PLoS One. 2013;8(3):e58021.
14. Mishra N, Pereira M, Rhodes RH, An P, Pipas JM, Jain K, et al. Identification of a novel polyomavirus in a pancreatic transplant recipient with retinal blindness and vasculitic myopathy. J Infect Dis. 2014;210(10):1595-9.
15. Gheit T, Dutta S, Oliver J, Robitaille A, Hampras S, Combes JD, et al. Isolation and characterization of a novel putative human polyomavirus. Virology. 2017;506:45-54.
16. Dalianis T, Hirsch HH. Human polyomaviruses in disease and cancer. Virology. 2013;437(2):63-72.
17. White MK, Gordon J, Khalili K. The rapidly expanding family of human polyomaviruses: recent developments in understanding their life cycle and role in human pathology. PLoS Pathog. 2013;9(3):e1003206.
18. Moens U, Krumbholz A, Ehlers B, Zell R, Johne R, Calvignac-Spencer S, et al. Biology, evolution, and medical importance of polyomaviruses: An update. Infect Genet Evol. 2017;54:18-38.
19. Chang D, Fung CY, Ou WC, Chao PC, Li SY, Wang M, et al. Self-assembly of the JC virus major capsid protein, VP1, expressed in insect cells. J Gen Virol. 1997;78:1435-9.
20. Liddington RC, Yan Y, Moulai J, Sahli R, Benjamin TL, Harrison SC. Structure of simian virus 40 at 3.8-A resolution. Nature. 1991;354(6351):278-84.
21. Montross L, Watkins S, Moreland RB, Mamon H, Caspar DL, Garcea RL. Nuclear assembly of polyomavirus capsids in insect cells expressing the major capsid protein VP1. J Virol. 1991;65(9):4991-8.
22. Imperiale MJ, Jiang M. What DNA Viral Genomic Rearrangements Tell Us about Persistence. J Virol. 2015;89(4):1948-50.
23. White MK, Safak M, Khalili K. Regulation of gene expression in primate polyomaviruses. J Virol. 2009;83(21):10846-56.
24. Vaz B, Cinque P, Pickhardt M, Weber T. Analysis of the transcriptional control region in progressive multifocal leukoencephalopathy. J Neurovirol. 2000;6(5):398-409.

25. Gosert R, Rinaldo CH, Funk GA, Egli A, Ramos E, Drachenberg CB, et al. Polyomavirus BK with rearranged noncoding control region emerge in vivo in renal transplant patients and increase viral replication and cytopathology. J Exp Med. 2008;205(4):841-52.
26. Gosert R, Kardas P, Major EO, Hirsch HH. Rearranged JC virus noncoding control regions found in progressive multifocal leukoencephalopathy patient samples increase virus early gene expression and replication rate. J Virol. 2010;84(20):10448-56.
27. Sharma PM, Gupta G, Vats A, Shapiro R, Randhawa PS. Polyomavirus BK non-coding control region rearrangements in health and disease. J Med Virol. 2007;79(8):1199-207.
28. Moens U, Rasheed K, Abdulsalam I, Sveinbjornsson B. The role of Merkel cell polyomavirus and other human polyomaviruses in emerging hallmarks of cancer. Viruses. 2015;7(4):1871-901.
29. Stettner MR, Nance JA, Wright CA, Kinoshita Y, Kim WK, Morgello S, et al. SMAD proteins of oligodendroglial cells regulate transcription of JC virus early and late genes coordinately with the Tat protein of human immunodeficiency virus type 1. J Gen Virol. 2009;90(Pt 8):2005-14.
30. Romagnoli L, Sariyer IK, Tung J, Feliciano M, Sawaya BE, Del Valle L, et al. Early growth response-1 protein is induced by JC virus infection and binds and regulates the JC virus promoter. Virology. 2008;375(2):331-41.
31. Ranganathan PN, Khalili K. The transcriptional enhancer element, kappa B, regulates promoter activity of the human neurotropic virus, JCV, in cells derived from the CNS. Nucleic Acids Res. 1993;21(8):1959-64.
32. Wegner M, Drolet DW, Rosenfeld MG. Regulation of JC virus by the POU-domain transcription factor Tst-1: implications for progressive multifocal leukoencephalopathy. Proc Natl Acad Sci U S A. 1993;90(10):4743-7.
33. Renner K, Leger H, Wegner M. The POU domain protein Tst-1 and papovaviral large tumor antigen function synergistically to stimulate glia-specific gene expression of JC virus. Proc Natl Acad Sci U S A. 1994;91(14):6433-7.
34. Bethge T, Ajuh E, Hirsch HH. Imperfect Symmetry of Sp1 and Core Promoter Sequences Regulates Early and Late Virus Gene Expression of the Bidirectional BK Polyomavirus Noncoding Control Region. J Virol. 2016;90(22):10083-101.
35. Ambalathingal GR, Francis RS, Smyth MJ, Smith C, Khanna R. BK Polyomavirus: Clinical Aspects, Immune Regulation, and Emerging Therapies. Clin Microbiol Rev. 2017;30(2):503-28.

36. Li PP, Naknanishi A, Tran MA, Ishizu K, Kawano M, Phillips M, et al. Importance of Vp1 calcium-binding residues in assembly, cell entry, and nuclear entry of simian virus 40. J Virol. 2003;77(13):7527-38.
37. Tsai B, Gilbert JM, Stehle T, Lencer W, Benjamin TL, Rapoport TA. Gangliosides are receptors for murine polyoma virus and SV40. Embo J. 2003;22(17):4346-55.
38. Low JA, Magnuson B, Tsai B, Imperiale MJ. Identification of gangliosides GD1b and GT1b as receptors for BK virus. J Virol. 2006;80(3):1361-6.
39. Assetta B, Maginnis MS, Gracia Ahufinger I, Haley SA, Gee GV, Nelson CD, et al. 5-HT2 receptors facilitate JC polyomavirus entry. J Virol. 2013;87(24):13490-8.
40. Schelhaas M, Malmstrom J, Pelkmans L, Haugstetter J, Ellgaard L, Grunewald K, et al. Simian Virus 40 depends on ER protein folding and quality control factors for entry into host cells. Cell. 2007;131(3):516-29.
41. Eash S, Atwood WJ. Involvement of cytoskeletal components in BK virus infectious entry. J Virol. 2005;79(18):11734-41.
42. Jiang M, Abend JR, Tsai B, Imperiale MJ. Early events during BK virus entry and disassembly. J Virol. 2009;83(3):1350-8.
43. Dupzyk A, Tsai B. How Polyomaviruses Exploit the ERAD Machinery to Cause Infection. Viruses. 2016;8(9).
44. Walczak CP, Tsai B. A PDI family network acts distinctly and coordinately with ERp29 to facilitate polyomavirus infection. J Virol. 2011;85(5):2386-96.
45. Nakanishi A, Itoh N, Li PP, Handa H, Liddington RC, Kasamatsu H. Minor capsid proteins of simian virus 40 are dispensable for nucleocapsid assembly and cell entry but are required for nuclear entry of the viral genome. J Virol. 2007;81(8):3778-85.
46. Imperiale MJ. The human polyomaviruses: An overview. In: Khalili K, Stoner, G. L., editor. Human Polyomaviruses: Molecular and Clinical Perspectives: Wiley-Liss; 2001. p. 53-71.
47. Barbanti-Brodano G, Sabbioni S, Martini F, Negrini M, Corallini A, Tognon M. BK virus, JC virus and Simian Virus 40 infection in humans, and association with human tumors. Adv Exp Med Biol. 2006;577:319-41.
48. DeCaprio JA, Garcea RL. A cornucopia of human polyomaviruses. Nat Rev Microbiol. 2013;11(4):264-76.

49. Harrison C, Jiang T, Banerjee P, Meinke G, D'Abramo CM, Schaffhausen B, et al. Polyomavirus large T antigen binds symmetrical repeats at the viral origin in an asymmetrical manner. J Virol. 2013;87(24):13751-9.
50. Bargonetti J, Reynisdottir I, Friedman PN, Prives C. Site-specific binding of wild-type p53 to cellular DNA is inhibited by SV40 T antigen and mutant p53. Genes Dev. 1992;6(10):1886-98.
51. Pipas JM. SV40: Cell transformation and tumorigenesis. Virology. 2009;384(2):294-303.
52. Khalili K, Sariyer IK, Safak M. Small tumor antigen of polyomaviruses: role in viral life cycle and cell transformation. J Cell Physiol. 2008;215(2):309-19.
53. Frost JA, Alberts AS, Sontag E, Guan K, Mumby MC, Feramisco JR. Simian virus 40 small t antigen cooperates with mitogen-activated kinases to stimulate AP-1 activity. Mol Cell Biol. 1994;14(9):6244-52.
54. Seo GJ, Fink LH, O'Hara B, Atwood WJ, Sullivan CS. Evolutionarily conserved function of a viral microRNA. J Virol. 2008;82(20):9823-8.
55. Sullivan CS, Grundhoff AT, Tevethia S, Pipas JM, Ganem D. SV40-encoded microRNAs regulate viral gene expression and reduce susceptibility to cytotoxic T cells. Nature. 2005;435(7042):682-6.
56. Seo GJ, Chen CJ, Sullivan CS. Merkel cell polyomavirus encodes a microRNA with the ability to autoregulate viral gene expression. Virology. 2009;383(2):183-7.
57. Bauman Y, Mandelboim O. MicroRNA based immunoevasion mechanism of human polyomaviruses. RNA Biol. 2011;8(4):591-4.
58. Broekema NM, Imperiale MJ. miRNA regulation of BK polyomavirus replication during early infection. Proc Natl Acad Sci U S A. 2013;110(20):8200-5.
59. Fanning E, Knippers R. Structure and function of simian virus 40 large tumor antigen. Annu Rev Biochem. 1992;61:55-85.
60. Raghava S, Giorda KM, Romano FB, Heuck AP, Hebert DN. The SV40 late protein VP4 is a viroporin that forms pores to disrupt membranes for viral release. PLoS Pathog. 2011;7(6):e1002116.
61. Khalili K, White MK, Sawa H, Nagashima K, Safak M. The agnoprotein of polyomaviruses: a multifunctional auxiliary protein. J Cell Physiol. 2005;204(1):1-7.
62. Allander T, Andreasson K, Gupta S, Bjerkner A, Bogdanovic G, Persson MA, et al. Identification of a third human polyomavirus. J Virol. 2007;81(8):4130-6.

63. Shishido-Hara Y, Nagashima K. Synthesis and assembly of polyomavirus virions. In: Kamel Khalili GLS, editor. Human Polyomaviruses: Molecular and Clinical Perspectives. Wiley-Liss.,Inc.: Wiley-Liss.,Inc.; 2001. p. 149-77.
64. Daniels R, Rusan NM, Wilbuer AK, Norkin LC, Wadsworth P, Hebert DN. Simian virus 40 late proteins possess lytic properties that render them capable of permeabilizing cellular membranes. J Virol. 2006;80(13):6575-87.
65. Suzuki T, Orba Y, Okada Y, Sunden Y, Kimura T, Tanaka S, et al. The human polyoma JC virus agnoprotein acts as a viroporin. PLoS Pathog. 2010;6(3):e1000801.
66. Henriksen S, Hansen T, Bruun JA, Rinaldo CH. The Presumed Polyomavirus Viroporin VP4 of Simian Virus 40 or Human BK Polyomavirus Is Not Required for Viral Progeny Release. J Virol. 2016;90(22):10398-413.
67. 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):414-21.
68. Sweet BH, Hilleman MR. The vacuolating virus, S.V. 40. Proc Soc Exp Biol Med. 1960;105:420-7.
69. Shein HM, Enders JF. Transformation induced by simian virus 40 in human renal cell cultures. I. Morphology and growth characteristics. Proc Natl Acad Sci U S A. 1962;48:1164-72.
70. Shah KV. SV40 and human cancer: a review of recent data. Int J Cancer. 2007;120(2):215-23.
71. Rivera Z, Strianese O, Bertino P, Yang H, Pass H, Carbone M. The relationship between simian virus 40 and mesothelioma. Curr Opin Pulm Med. 2008;14(4):316-21.
72. Strickler HD, Goedert JJ, Devesa SS, Lahey J, Fraumeni JF, Jr., Rosenberg PS. Trends in U.S. pleural mesothelioma incidence rates following simian virus 40 contamination of early poliovirus vaccines. J Natl Cancer Inst. 2003;95(1):38-45.
73. Sabatier J, Uro-Coste E, Benouaich A, Boetto S, Gigaud M, Tremoulet M, et al. Immunodetection of SV40 large T antigen in human central nervous system tumours. J Clin Pathol. 2005;58(4):429-31.
74. Felsani A, Mileo AM, Paggi MG. Retinoblastoma family proteins as key targets of the small DNA virus oncoproteins. Oncogene. 2006;25(38):5277-85.
75. Martini F, Corallini A, Balatti V, Sabbioni S, Pancaldi C, Tognon M. Simian virus 40 in humans. Infect Agent Cancer. 2007;2:13.

76. Rinaldo CH, Tylden GD, Sharma BN. The human polyomavirus BK (BKPyV): virological background and clinical implications. APMIS. 2013;121:728–45.
77. Kean JM, Rao S, Wang M, Garcea RL. Seroepidemiology of human polyomaviruses. PLoS Pathog. 2009;5(3):e1000363.
78. Knowles WA. Discovery and epidemiology of the human polyomaviruses BK virus (BKV) and JC virus (JCV). Adv Exp Med Biol. 2006;577:19-45.
79. Goudsmit J, Wertheim-van Dillen P, van Strien A, van der Noordaa J. The role of BK virus in acute respiratory tract disease and the presence of BKV DNA in tonsils. J Med Virol. 1982;10(2):91-9.
80. Bofill-Mas S, Formiga-Cruz M, Clemente-Casares P, Calafell F, Girones R. Potential transmission of human polyomaviruses through the gastrointestinal tract after exposure to virions or viral DNA. J Virol. 2001;75(21):10290-9.
81. Doerries K. Human polyomavirus JC and BK persistent infection. Adv Exp Med Biol. 2006;577:102-16.
82. Heritage J, Chesters PM, McCance DJ. The persistence of papovavirus BK DNA sequences in normal human renal tissue. J Med Virol. 1981;8(2):143-50.
83. Chesters PM, Heritage J, McCance DJ. Persistence of DNA sequences of BK virus and JC virus in normal human tissues and in diseased tissues. J Infect Dis. 1983;147(4):676-84.
84. Rekvig OP, Moens U. Polyomavirus BK and Autoimmunity to Nucleosomes. graft [Internet]. 2002; 5(36):[36-45 pp.].
85. Burger-Calderon R, Madden V, Hallett RA, Gingerich AD, Nickeleit V, Webster-Cyriaque J. Replication of oral BK virus in human salivary gland cells. J Virol. 2014;88(1):559-73.
86. Jeffers LK, Madden V, Webster-Cyriaque J. BK virus has tropism for human salivary gland cells in vitro: implications for transmission. Virology. 2009;394(2):183-93.
87. Abend JR, Jiang M, Imperiale MJ. BK virus and human cancer: innocent until proven guilty. Semin Cancer Biol. 2009;19(4):252-60.
88. Dorries K, Sbiera S, Drews K, Arendt G, Eggers C, Dorries R. Association of human polyomavirus JC with peripheral blood of immunoimpaired and healthy individuals. J Neurovirol. 2003;9 Suppl 1:81-7.
89. Ravichandran V, Major EO. Viral proteomics: a promising approach for understanding JC virus tropism. Proteomics. 2006;6(20):5628-36.

90. Boothpur R, Brennan DC. Human polyoma viruses and disease with emphasis on clinical BK and JC. J Clin Virol. 2010;47(4):306-12.
91. Maginnis MS, Atwood WJ. JC virus: an oncogenic virus in animals and humans? Semin Cancer Biol. 2009;19(4):261-9.
92. Abend JR, Low JA, Imperiale MJ. Global effects of BKV infection on gene expression in human primary kidney epithelial cells. Virology. 2010;397(1):73-9.
93. Jiang M, Abend JR, Johnson SF, Imperiale MJ. The role of polyomaviruses in human disease. Virology. 2009;384(2):266-73.
94. Berger JR, Concha M. Progressive multifocal leukoencephalopathy: the evolution of a disease once considered rare. J Neurovirol. 1995;1(1):5-18.
95. Bogdanovic G, Priftakis P, Giraud G, Kuzniar M, Ferraldeschi R, Kokhaei P, et al. Association between a high BK virus load in urine samples of patients with graft-versus-host disease and development of hemorrhagic cystitis after hematopoietic stem cell transplantation. J Clin Microbiol. 2004;42(11):5394-6.
96. Erard V, Storer B, Corey L, Nollkamper J, Huang ML, Limaye A, et al. BK virus infection in hematopoietic stem cell transplant recipients: frequency, risk factors, and association with postengraftment hemorrhagic cystitis. Clin Infect Dis. 2004;39(12):1861-5.
97. Baksh FK, Finkelstein SD, Swalsky PA, Stoner GL, Ryschkewitsch CF, Randhawa P. Molecular genotyping of BK and JC viruses in human polyomavirus-associated interstitial nephritis after renal transplantation. Am J Kidney Dis. 2001;38(2):354-65.
98. Venter M, Visser A, Lassauniere R. Human polyomaviruses, WU and KI in HIV exposed children with acute lower respiratory tract infections in hospitals in South Africa. J Clin Virol. 2009;44(3):230-4.
99. Moens U, Johannessen M, Barcena-Panero A, Gerits N, Van Ghelue M. Emerging polyomavirus in the human population. RIF 1(2):59-93 (2010): University of Tromsö, Faculty of Health Sciences, Institute of Medical Biology, N-9037 Tromsö, Norway
Carlos III Health Institute, National Centre of Microbiology, Service of Diagnostics in Microbiology, Laboratory of Isolation and Viral Detection, Majadahonda, Madrid, Spain
University Hospital Northern-Norway, Department of Medical Genetics, N-9038 Tromsö, Norway; 2010 2010.

100. Csoma E, Meszaros B, Asztalos L, Gergely L. WU and KI polyomaviruses in respiratory, blood and urine samples from renal transplant patients. J Clin Virol. 2015;64:28-33.
101. Barzon L, Squarzon L, Militello V, Trevisan M, Porzionato A, Macchi V, et al. WU and KI polyomaviruses in the brains of HIV-positive patients with and without progressive multifocal leukoencephalopathy. J Infect Dis. 2009;200(11):1755-8.
102. Siebrasse EA, Nguyen NL, Smith C, Simmonds P, Wang D. Immunohistochemical detection of KI polyomavirus in lung and spleen. Virology. 2014;468-470C:178-84.
103. Abedi Kiasari B, Vallely PJ, Corless CE, Al-Hammadi M, Klapper PE. Age-related pattern of KI and WU polyomavirus infection. J Clin Virol. 2008;43(1):123-5.
104. Gunel C, Kirdar S, Omurlu IK, Agdas F. Detection of the Epstein-Barr virus, Human Bocavirus and novel KI and KU polyomaviruses in adenotonsillar tissues. Int J Pediatr Otorhinolaryngol. 2015;79(3):423-7.
105. Babakir-Mina M, Ciccozzi M, Dimonte S, Farchi F, Valdarchi C, Rezza G, et al. Identification of the novel KI polyomavirus in the respiratory tract of an Italian patient. J Med Virol. 2008;80(11):2012-4.
106. Kleines M, Scheithauer S, Hengst M, Honnef D, Ritter K, Muhler E, et al. Low to medium WU-virus titers in young children with lower respiratory tract infections. Intervirology. 2008;51(6):444-6.
107. van der Zalm MM, Rossen JW, van Ewijk BE, Wilbrink B, van Esch PC, Wolfs TF, et al. Prevalence and pathogenicity of WU and KI polyomaviruses in children, the Netherlands. Emerg Infect Dis. 2008;14(11):1787-9.
108. Foulongne V, Brieu N, Jeziorski E, Chatain A, Rodiere M, Segondy M. KI and WU polyomaviruses in children, France. Emerg Infect Dis. 2008;14(3):523-5.
109. Yuan XH, Jin Y, Xie ZP, Gao HC, Xu ZQ, Zheng LS, et al. Prevalence of human KI and WU polyomaviruses in children with acute respiratory tract infection in China. J Clin Microbiol. 2008;46(10):3522-5.
110. Ren L, Gonzalez R, Xie Z, Zhang J, Liu C, Li J, et al. WU and KI polyomavirus present in the respiratory tract of children, but not in immunocompetent adults. J Clin Virol. 2008;43(3):330-3.
111. Neske F, Blessing K, Ullrich F, Prottel A, Wolfgang Kreth H, Weissbrich B. WU polyomavirus infection in children, Germany. Emerg Infect Dis. 2008;14(4):680-1.

112. Neske F, Blessing K, Prottel A, Ullrich F, Kreth HW, Weissbrich B. Detection of WU polyomavirus DNA by real-time PCR in nasopharyngeal aspirates, serum, and stool samples. J Clin Virol. 2009;44(2):115-8.
113. Lin F, Zheng M, Li H, Zheng C, Li X, Rao G, et al. WU polyomavirus in children with acute lower respiratory tract infections, China. J Clin Virol. 2008;42(1):94-102.
114. Abed Y, Wang D, Boivin G. WU polyomavirus in children, Canada. Emerg Infect Dis. 2007;13(12):1939-41.
115. Wattier RL, Vazquez M, Weibel C, Shapiro ED, Ferguson D, Landry ML, et al. Role of human polyomaviruses in respiratory tract disease in young children. Emerg Infect Dis. 2008;14(11):1766-8.
116. Bialasiewicz S, Whiley DM, Lambert SB, Jacob K, Bletchly C, Wang D, et al. Presence of the newly discovered human polyomaviruses KI and WU in Australian patients with acute respiratory tract infection. J Clin Virol. 2008;41(2):63-8.
117. Bialasiewicz S, Whiley DM, Lambert SB, Nissen MD, Sloots TP. Detection of BK, JC, WU, or KI polyomaviruses in faecal, urine, blood, cerebrospinal fluid and respiratory samples. J Clin Virol. 2009;45(3):249-54.
118. Nguyen NL, Le BM, Wang D. Serologic evidence of frequent human infection with WU and KI polyomaviruses. Emerg Infect Dis. 2009;15(8):1199-205.
119. Neske F, Prifert C, Scheiner B, Ewald M, Schubert J, Opitz A, et al. High prevalence of antibodies against polyomavirus WU, polyomavirus KI, and human bocavirus in German blood donors. BMC Infect Dis.10:215.
120. Rao S, Garcea RL, Robinson CC, Simoes EA. WU and KI polyomavirus infections in pediatric hematology/oncology patients with acute respiratory tract illness. J Clin Virol. 2011;52(1):28-32.
121. Babakir-Mina M, Ciccozzi M, Alteri C, Polchi P, Picardi A, Greco F, et al. Excretion of the novel polyomaviruses KI and WU in the stool of patients with hematological disorders. J Med Virol. 2009;81(9):1668-73.
122. Barzon L, Squarzon L, Pacenti M, Scotton PG, Palu G. Detection of WU polyomavirus in cerebrospinal fluid specimen from a patient with AIDS and suspected progressive multifocal leukoencephalopathy. J Infect Dis. 2009;200(2):314-5.
123. Debiaggi M, Canducci F, Brerra R, Sampaolo M, Marinozzi MC, Parea M, et al. Molecular epidemiology of KI and WU polyomaviruses in infants with acute

respiratory disease and in adult hematopoietic stem cell transplant recipients. J Med Virol. 2010;82(1):153-6.

124. Mourez T, Bergeron A, Ribaud P, Scieux C, de Latour RP, Tazi A, et al. Polyomaviruses KI and WU in immunocompromised patients with respiratory disease. Emerg Infect Dis. 2009;15(1):107-9.

125. Sharp CP, Norja P, Anthony I, Bell JE, Simmonds P. Reactivation and Mutation of Newly Discovered WU, KI, and Merkel Cell Carcinoma Polyomaviruses in Immunosuppressed Individuals. J Infect Dis. 2009;199(3):398-404.

126. Babakir-Mina M, Ciccozzi M, Perno CF, Ciotti M. The novel KI, WU, MC polyomaviruses: possible human pathogens? New Microbiol. 2011;34(1):1-8.

127. Bergallo M, Costa C, Terlizzi ME, Astegiano S, Curtoni A, Solidoro P, et al. Quantitative detection of the new polyomaviruses KI, WU and Merkel cell virus in transbronchial biopsies from lung transplant recipients. J Clin Pathol. 2010;63(8):722-5.

128. Babakir-Mina M, Ciccozzi M, Campitelli L, Aquaro S, Lo Coco A, Perno CF, et al. Identification of the novel KI Polyomavirus in paranasal and lung tissues. J Med Virol. 2009;81(3):558-61.

129. Giraud G, Ramqvist T, Ragnarsson-Olding B, Dalianis T. DNA from BK virus and JC virus and from KI, WU, and MC polyomaviruses as well as from simian virus 40 is not detected in non-UV-light-associated primary malignant melanomas of mucous membranes. J Clin Microbiol. 2008;46(11):3595-8.

130. Moens U, Johannessen M. Human polyomaviruses and cancer: expanding repertoire. J Dtsch Dermatol Ges. 2008;6(9):704-8.

131. Pulitzer MP, Amin BD, Busam KJ. Merkel cell carcinoma: review. Adv Anat Pathol. 2009;16(3):135-44.

132. Delbue S, Tremolada S, Elia F, Carloni C, Amico S, Tavazzi E, et al. Lymphotropic polyomavirus is detected in peripheral blood from immunocompromised and healthy subjects. J Clin Virol. 2009;47(2):156-60.

133. Shuda M, Arora R, Kwun HJ, Feng H, Sarid R, Fernandez-Figueras MT, 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):1243-9.

134. Chang Y, Moore PS. Merkel cell carcinoma: a virus-induced human cancer. Annu Rev Pathol. 2012;7:123-44.

135. Mertz KD, Paasinen A, Arnold A, Baumann M, Offner F, Willi N, et al. Merkel cell polyomavirus large T antigen is detected in rare cases of nonmelanoma skin cancer. J Cutan Pathol. 2013;40(6):543-9.
136. Baez CF, Guimaraes MA, Martins RA, Zalona AC, Cossatis JJ, Zalis MG, et al. Detection of Merkel cell polyomavirus in oral samples of renal transplant recipients without Merkel cell carcinoma. J Med Virol. 2013;85(11):2016-9.
137. Signorini L, Belingheri M, Ambrogi F, Pagani E, Binda S, Ticozzi R, et al. High frequency of Merkel cell polyomavirus DNA in the urine of kidney transplant recipients and healthy controls. J Clin Virol. 2014;61(4):565-70.
138. Rockett RJ, Sloots TP, Bowes S, O'Neill N, Ye S, Robson J, et al. Detection of novel polyomaviruses, TSPyV, HPyV6, HPyV7, HPyV9 and MWPyV in feces, urine, blood, respiratory swabs and cerebrospinal fluid. PLoS One. 2013;8(5):e62764.
139. Rennspiess D, Pujari S, Keijzers M, Abdul-Hamid MA, Hochstenbag M, Dingemans AM, et al. Detection of human polyomavirus 7 in human thymic epithelial tumors. J Thorac Oncol. 2015;10(2):360-6.
140. Ho J, Jedrych JJ, Feng H, Natalie AA, Grandinetti L, Mirvish E, et al. Human Polyomavirus 7-Associated Pruritic Rash and Viremia in Transplant Recipients. J Infect Dis. 2014:1-6.
141. Nguyen KD, Lee EE, Yue Y, Stork J, Pock L, North JP, et al. Human polyomavirus 6 and 7 are associated with pruritic and dyskeratotic dermatoses. J Am Acad Dermatol. 2017;76(5):932-40 e3.
142. Duncavage EJ, Pfeifer JD. Human polyomaviruses 6 and 7 are not detectable in Merkel cell polyomavirus-negative Merkel cell carcinoma. J Cutan Pathol. 2011;38(10):790-6.
143. Haycox CL, Kim S, Fleckman P, Smith LT, Piepkorn M, Sundberg JP, et al. Trichodysplasia spinulosa--a newly described folliculocentric viral infection in an immunocompromised host. J Investig Dermatol Symp Proc. 1999;4(3):268-71.
144. Kazem S, van der Meijden E, Feltkamp MC. The trichodysplasia spinulosa-associated polyomavirus; virological background and clinical implications. APMIS. 2013;121:770–82.
145. Izakovic J, Buchner SA, Duggelin M, Guggenheim R, Itin PH. [Hair-like hyperkeratoses in patients with kidney transplants. A new cyclosporin side-effect]. Hautarzt. 1995;46(12):841-6.

146. Kazem S, van der Meijden E, Wang RC, Rosenberg AS, Pope E, Benoit T, et al. Polyomavirus-associated Trichodysplasia spinulosa involves hyperproliferation, pRB phosphorylation and upregulation of p16 and p21. PLoS One. 2014;9(10):e108947.
147. Sadeghi M, Aaltonen LM, Hedman L, Chen T, Soderlund-Venermo M, Hedman K. Detection of TS polyomavirus DNA in tonsillar tissues of children and adults: evidence for site of viral latency. J Clin Virol. 2014;59(1):55-8.
148. Fischer MK, Kao GF, Nguyen HP, Drachenberg CB, Rady PL, Tyring SK, et al. Specific detection of trichodysplasia spinulosa-associated polyomavirus DNA in skin and renal allograft tissues in a patient with trichodysplasia spinulosa. Arch Dermatol. 2012;148(6):726-33.
149. Tsuzuki S, Fukumoto H, Mine S, Sato N, Mochizuki M, Hasegawa H, et al. Detection of trichodysplasia spinulosa-associated polyomavirus in a fatal case of myocarditis in a seven-month-old girl. Int J Clin Exp Pathol. 2014;7(8):5308-12.
150. van der Meijden E, Kazem S, Burgers MM, Janssens R, Bouwes Bavinck JN, de Melker H, et al. Seroprevalence of trichodysplasia spinulosa-associated polyomavirus. Emerg Infect Dis. 2011;17(8):1355-63.
151. Delbue S, Tremolada S, Branchetti E, Elia F, Gualco E, Marchioni E, et al. First identification and molecular characterization of lymphotropic polyomavirus in peripheral blood from patients with leukoencephalopathies. J Clin Microbiol. 2008;46(7):2461-2.
152. Brade L, Muller-Lantzsch N, zur Hausen H. B-lymphotropic papovavirus and possibility of infections in humans. J Med Virol. 1981;6(4):301-8.
153. Trusch F, Klein M, Finsterbusch T, Kuhn J, Hofmann J, Ehlers B. Seroprevalence of human polyomavirus 9 and cross-reactivity to African green monkey-derived lymphotropic polyomavirus. J Gen Virol. 2012;93(Pt 4):698-705.
154. Sauvage V, Foulongne V, Cheval J, Ar Gouilh M, Pariente K, Dereure O, et al. Human polyomavirus related to African green monkey lymphotropic polyomavirus. Emerg Infect Dis. 2011;17(8):1364-70.
155. Lednicky JA, Butel JS, Luetke MC, Loeb JC. Complete genomic sequence of a new Human polyomavirus 9 strain with an altered noncoding control region. Virus Genes. 2014;49(3):490-2.
156. Buck CB, Phan GQ, Raiji MT, Murphy PM, McDermott DH, McBride AA. Complete genome sequence of a tenth human polyomavirus. J Virol. 2012;86(19):10887.

157. Yu G, Greninger AL, Isa P, Phan TG, Martinez MA, de la Luz Sanchez M, et al. Discovery of a novel polyomavirus in acute diarrheal samples from children. PLoS One. 2012;7(11):e49449.
158. Wieland U, Silling S, Hellmich M, Potthoff A, Pfister H, Kreuter A. Human polyomaviruses 6, 7, 9, 10 and Trichodysplasia spinulosa-associated polyomavirus in HIV-infected men. J Gen Virol. 2014;95(Pt 4):928-32.
159. Pastrana DV, Fitzgerald PC, Phan GQ, Raiji MT, Murphy PM, McDermott DH, et al. A divergent variant of the eleventh human polyomavirus species, saint louis polyomavirus. Genome Announc. 2013;1(5):e00812-13.
160. Lim ES, Meinerz NM, Primi B, Wang D, Garcea RL. Common exposure to STL polyomavirus during childhood. Emerg Infect Dis. 2014;20(9):1559-61.
161. Hirsch HH. BK virus: opportunity makes a pathogen. Clin Infect Dis. 2005;41(3):354-60.
162. Reploeg MD, Storch GA, Clifford DB. Bk virus: a clinical review. Clin Infect Dis. 2001;33(2):191-202.
163. Akazawa Y, Terada Y, Yamane T, Tanaka S, Aimoto M, Koh H, et al. Fatal BK virus pneumonia following stem cell transplantation. Transpl Infect Dis. 2012;14(6):E142-6.
164. Chittick P, Williamson JC, Ohl CA. BK virus encephalitis: case report, review of the literature, and description of a novel treatment modality. Ann Pharmacother. 2013;47(9):1229-33.
165. Coleman DV, Mackenzie EF, Gardner SD, Poulding JM, Amer B, Russell WJ. Human polyomavirus (BK) infection and ureteric stenosis in renal allograft recipients. J Clin Pathol. 1978;31(4):338-47.
166. Delbue S, Ferrante P, Provenzano M. Polyomavirus BK and prostate cancer: an unworthy scientific effort? Oncoscience. 2014;1(4):296-303.
167. Lundstig A, Dillner J. Serological diagnosis of human polyomavirus infection. Adv Exp Med Biol. 2006;577:96-101.
168. Niv Y, Goel A, Boland CR. JC virus and colorectal cancer: a possible trigger in the chromosomal instability pathways. Curr Opin Gastroenterol. 2005;21(1):85-9.
169. Brew BJ, Davies NW, Cinque P, Clifford DB, Nath A. Progressive multifocal leukoencephalopathy and other forms of JC virus disease. Nat Rev Neurol. 2010;6(12):667-79.

170. Gossai A, Waterboer T, Nelson HH, Michel A, Willhauck-Fleckenstein M, Farzan SF, et al. Seroepidemiology of Human Polyomaviruses in a US Population. Am J Epidemiol. 2016;183(1):61-9.
171. Toptan T, Yousem SA, Ho J, Matsushima Y, Stabile LP, Fernandez-Figueras MT, et al. Survey for human polyomaviruses in cancer. JCI Insight. 2016;1(2):1-14.
172. Nicol JT, Robinot R, Carpentier A, Carandina G, Mazzoni E, Tognon M, et al. Age-specific seroprevalences of merkel cell polyomavirus, human polyomaviruses 6, 7, and 9, and trichodysplasia spinulosa-associated polyomavirus. Clin Vaccine Immunol. 2013;20(3):363-8.
173. Pastrana DV, Tolstov YL, Becker JC, Moore PS, Chang Y, Buck CB. Quantitation of human seroresponsiveness to Merkel cell polyomavirus. PLoS Pathog. 2009;5(9):e1000578.
174. van der Meijden E, Bialasiewicz S, Rockett RJ, Tozer SJ, Sloots TP, Feltkamp MC. Different serologic behavior of MCPyV, TSPyV, HPyV6, HPyV7 and HPyV9 polyomaviruses found on the skin. PLoS One. 2013;8(11):e81078.
175. Ehlers B, Wieland U. The novel human polyomaviruses HPyV6, 7, 9 and beyond. APMIS. 2013;121:783–95.
176. Nicol JT, Touze A, Robinot R, Arnold F, Mazzoni E, Tognon M, et al. Seroprevalence and cross-reactivity of human polyomavirus 9. Emerg Infect Dis. 2012;18(8):1329-32.
177. Berrios C, Jung J, Primi B, Wang M, Pedamallu C, Duke F, et al. Malawi polyomavirus is a prevalent human virus that interacts with known tumor suppressors. J Virol. 2015;89(1):857-62.
178. Nicol JT, Leblond V, Arnold F, Guerra G, Mazzoni E, Tognon M, et al. Seroprevalence of human Malawi polyomavirus. J Clin Microbiol. 2014;52(1):321-3.
179. Hirsch HH. Polyoma and Papilloma Virus Infections after Hematopoietic Stem Cell or Solid Organ Transplantation. In: Bowden P, Ljungman, P., Snydman, D.R., editor. Transplant Infections. Third ed. Philadelphia, Baltimore, New York, London, Buenos Aires, Hong Kong, Sydney, Tokyo.: Lippincott Williams & Wilkins; 2010. p. 465-82.
180. Geetha D, Tong BC, Racusen L, Markowitz JS, Westra WH. Bladder carcinoma in a transplant recipient: evidence to implicate the BK human polyomavirus as a causal transforming agent. Transplantation. 2002;73(12):1933-6.

181. Nickeleit V, Klimkait T, Binet IF, Dalquen P, Del Zenero V, Thiel G, et al. Testing for polyomavirus type BK DNA in plasma to identify renal-allograft recipients with viral nephropathy. N Engl J Med. 2000;342(18):1309-15.
182. Safdar A, Rubocki RJ, Horvath JA, Narayan KK, Waldron RL. Fatal immune restoration disease in human immunodeficiency virus type 1-infected patients with progressive multifocal leukoencephalopathy: impact of antiretroviral therapy-associated immune reconstitution. Clin Infect Dis. 2002;35(10):1250-7.
183. Kappos L, Bates D, Hartung HP, Havrdova E, Miller D, Polman CH, et al. Natalizumab treatment for multiple sclerosis: recommendations for patient selection and monitoring. Lancet Neurol. 2007;6(5):431-41.
184. Bredholt G, Olaussen E, Moens U, Rekvig OP. Linked production of antibodies to mammalian DNA and to human polyomavirus large T antigen: footprints of a common molecular and cellular process? Arthritis Rheum. 1999;42(12):2583-92.
185. Fredriksen K, Osei A, Sundsfjord A, Traavik T, Rekvig OP. On the biological origin of anti-double-stranded (ds) DNA antibodies: systemic lupus erythematosus-related anti-dsDNA antibodies are induced by polyomavirus BK in lupus-prone (NZBxNZW) F1 hybrids, but not in normal mice. Eur J Immunol. 1994;24(1):66-70.
186. Mortensen ES, Fenton KA, Rekvig OP. Lupus nephritis: the central role of nucleosomes revealed. Am J Pathol. 2008;172(2):275-83.
187. Hirsch HH, Knowles W, Dickenmann M, Passweg J, Klimkait T, Mihatsch MJ, et al. Prospective study of polyomavirus type BK replication and nephropathy in renal-transplant recipients. N Engl J Med. 2002;347(7):488-96.
188. Bressollette-Bodin C, Coste-Burel M, Hourmant M, Sebille V, Andre-Garnier E, Imbert-Marcille BM. A prospective longitudinal study of BK virus infection in 104 renal transplant recipients. Am J Transplant. 2005;5(8):1926-33.
189. Binet I, Nickeleit V, Hirsch HH, Prince O, Dalquen P, Gudat F, et al. Polyomavirus disease under new immunosuppressive drugs: a cause of renal graft dysfunction and graft loss. Transplantation. 1999;67(6):918-22.
190. Nickeleit V, Mihatsch MJ. Polyomavirus nephropathy in native kidneys and renal allografts: an update on an escalating threat. Transpl Int. 2006;19(12):960-73.
191. Nickeleit V, Singh HK, Mihatsch MJ. Polyomavirus nephropathy: morphology, pathophysiology, and clinical management. Curr Opin Nephrol Hypertens. 2003;12(6):599-605.

192. Nickeleit V, Hirsch HH, Zeiler M, Gudat F, Prince O, Thiel G, et al. BK-virus nephropathy in renal transplants-tubular necrosis, MHC-class II expression and rejection in a puzzling game. Nephrol Dial Transplant. 2000;15(3):324-32.
193. Mengel M, Marwedel M, Radermacher J, Eden G, Schwarz A, Haller H, et al. Incidence of polyomavirus-nephropathy in renal allografts: influence of modern immunosuppressive drugs. Nephrol Dial Transplant. 2003;18(6):1190-6.
194. Drachenberg CB, Hirsch HH, Ramos E, Papadimitriou JC. Polyomavirus disease in renal transplantation: review of pathological findings and diagnostic methods. Hum Pathol. 2005;36(12):1245-55.
195. Vats A, Randhawa PS, Shapiro R. Diagnosis and treatment of BK virus-associated transplant nephropathy. Adv Exp Med Biol. 2006;577:213-27.
196. Drachenberg CB, Papadimitriou JC, Wali R, Nogueira J, Mendley S, Hirsch HH, et al. Improved outcome of polyoma virus allograft nephropathy with early biopsy. Transplant Proc. 2004;36(3):758-9.
197. Bonvoisin C, Weekers L, Xhignesse P, Grosch S, Milicevic M, Krzesinski JM. Polyomavirus in renal transplantation: a hot problem. Transplantation. 2008;85(7 Suppl):S42-8.
198. Pavlakis M, Haririan A, Klassen DK. BK virus infection after non-renal transplantation. Adv Exp Med Biol. 2006;577:185-9.
199. Bohl DL, Storch GA, Ryschkewitsch C, Gaudreault-Keener M, Schnitzler MA, Major EO, et al. Donor origin of BK virus in renal transplantation and role of HLA C7 in susceptibility to sustained BK viremia. Am J Transplant. 2005;5(9):2213-21.
200. Ramos E, Drachenberg CB, Portocarrero M, Wali R, Klassen DK, Fink JC, et al. BK virus nephropathy diagnosis and treatment: experience at the University of Maryland Renal Transplant Program. Clin Transpl. 2002:143-53.
201. Ramos E, Drachenberg CB, Papadimitriou JC, Hamze O, Fink JC, Klassen DK, et al. Clinical course of polyoma virus nephropathy in 67 renal transplant patients. J Am Soc Nephrol. 2002;13(8):2145-51.
202. Bernhoff E, Gutteberg TJ, Sandvik K, Hirsch HH, Rinaldo CH. Cidofovir Inhibits Polyomavirus BK Replication in Human Renal Tubular Cells Downstream of Viral Early Gene Expression. Am J Transplant. 2008;8:1413-22.

203. Blanckaert K, De Vriese AS. Current recommendations for diagnosis and management of polyoma BK virus nephropathy in renal transplant recipients. Nephrol Dial Transplant. 2006;21(12):3364-7.
204. Bjorang O, Tveitan H, Midtvedt K, Broch LU, Scott H, Andresen PA. Treatment of polyomavirus infection with cidofovir in a renal-transplant recipient. Nephrol Dial Transplant. 2002;17(11):2023-5.
205. Kadambi PV, Josephson MA, Williams J, Corey L, Jerome KR, Meehan SM, et al. Treatment of refractory BK virus-associated nephropathy with cidofovir. Am J Transplant. 2003;3(2):186-91.
206. Kuypers DR, Vandooren AK, Lerut E, Evenepoel P, Claes K, Snoeck R, et al. Adjuvant low-dose cidofovir therapy for BK polyomavirus interstitial nephritis in renal transplant recipients. Am J Transplant. 2005;5(8):1997-2004.
207. Faguer S, Hirsch HH, Kamar N, Guilbeau-Frugier C, Ribes D, Guitard J, et al. Leflunomide treatment for polyomavirus BK-associated nephropathy after kidney transplantation. Transpl Int. 2007;20(11):962-9.
208. Josephson MA, Gillen D, Javaid B, Kadambi P, Meehan S, Foster P, et al. Treatment of renal allograft polyoma BK virus infection with leflunomide. Transplantation. 2006;81(5):704-10.
209. Williams JW, Javaid B, Kadambi PV, Gillen D, Harland R, Thistlewaite JR, et al. Leflunomide for polyomavirus type BK nephropathy. N Engl J Med. 2005;352(11):1157-8.
210. Rinaldo CH, Hirsch HH. Antivirals for the treatment of polyomavirus BK replication. Expert Rev Anti Infect Ther. 2007;5(1):105-15.
211. Sessa A, Esposito A, Giliberti A, Bergallo M, Costa C, Rossano R, et al. BKV reactivation in renal transplant recipients: diagnostic and therapeutic strategy--case reports. Transplant Proc. 2008;40(6):2055-8.
212. Sener A, House AA, Jevnikar AM, Boudville N, McAlister VC, Muirhead N, et al. Intravenous immunoglobulin as a treatment for BK virus associated nephropathy: one-year follow-up of renal allograft recipients. Transplantation. 2006;81(1):117-20.
213. Nickeleit V, Hirsch HH, Binet IF, Gudat F, Prince O, Dalquen P, et al. Polyomavirus infection of renal allograft recipients: from latent infection to manifest disease. J Am Soc Nephrol. 1999;10(5):1080-9.

214. Hingorani S. Renal Complications of Hematopoietic-Cell Transplantation. N Engl J Med. 2016;374(23):2256-67.
215. Cesaro S, Dalianis T, Rinaldo CH, Koskenvuo M, Einsele H, Hirsch HH. ECIL 6 - Guidelines for the Prevention, Diagnosis, and Treatment of BK Polyomavirus Disease in Stem Cell Transplant Patients 2016 [
216. Apperley JF, Rice SJ, Bishop JA, Chia YC, Krausz T, Gardner SD, et al. Late-onset hemorrhagic cystitis associated with urinary excretion of polyomaviruses after bone marrow transplantation. Transplantation. 1987;43(1):108-12.
217. Azzi A, Cesaro S, Laszlo D, Zakrzewska K, Ciappi S, De Santis R, et al. Human polyomavirus BK (BKV) load and haemorrhagic cystitis in bone marrow transplantation patients. J Clin Virol. 1999;14(2):79-86.
218. Bedi A, Miller CB, Hanson JL, Goodman S, Ambinder RF, Charache P, et al. Association of BK virus with failure of prophylaxis against hemorrhagic cystitis following bone marrow transplantation. J Clin Oncol. 1995;13(5):1103-9.
219. Egli A, Binggeli S, Bodaghi S, Dumoulin A, Funk GA, Khanna N, et al. Cytomegalovirus and polyomavirus BK posttransplant. Nephrol Dial Transplant. 2007;22 Suppl 8:viii72-viii82.
220. Dropulic LK, Jones RJ. Polyomavirus BK infection in blood and marrow transplant recipients. Bone Marrow Transplant. 2008;41(1):11-8.
221. Maertens J, Cesaro S, Maschmeyer G, Einsele H, Donnelly JP, Alanio A, et al. ECIL guidelines for preventing Pneumocystis jirovecii pneumonia in patients with haematological malignancies and stem cell transplant recipients. J Antimicrob Chemother. 2016;71(9):2397-404.
222. Held TK, Biel SS, Nitsche A, Kurth A, Chen S, Gelderblom HR, et al. Treatment of BK virus-associated hemorrhagic cystitis and simultaneous CMV reactivation with cidofovir. Bone Marrow Transplant. 2000;26(3):347-50.
223. Leung AY, Chan MT, Yuen KY, Cheng VC, Chan KH, Wong CL, et al. Ciprofloxacin decreased polyoma BK virus load in patients who underwent allogeneic hematopoietic stem cell transplantation. Clin Infect Dis. 2005;40(4):528-37.
224. Savona MR, Newton D, Frame D, Levine JE, Mineishi S, Kaul DR. Low-dose cidofovir treatment of BK virus-associated hemorrhagic cystitis in recipients of hematopoietic stem cell transplant. Bone Marrow Transplant. 2007;39(12):783-7.

225. Hirsch HH, Yakhontova K, Lu M, Manzetti J. BK Polyomavirus Replication in Renal Tubular Epithelial Cells Is Inhibited by Sirolimus, but Activated by Tacrolimus Through a Pathway Involving FKBP-12. Am J Transplant. 2016;16(3):821-32.
226. Hirsch HH, Kardas P, Kranz D, Leboeuf C. The human JC polyomavirus (JCPyV): virological background and clinical implications. APMIS. 2013;121:685–727.
227. Kuhle J, Gosert R, Buhler R, Derfuss T, Sutter R, Yaldizli O, et al. Management and outcome of CSF-JC virus PCR-negative PML in a natalizumab-treated patient with MS. Neurology. 2011;77(23):2010-6.
228. Kappos L, Bates D, Edan G, Eraksoy M, Garcia-Merino A, Grigoriadis N, et al. Natalizumab treatment for multiple sclerosis: updated recommendations for patient selection and monitoring. Lancet Neurol. 2011;10(8):745-58.
229. Viscidi RP, Khanna N, Tan CS, Li X, Jacobson L, Clifford DB, et al. JC virus antibody and viremia as predictors of progressive multifocal leukoencephalopathy in human immunodeficiency virus-1-infected individuals. Clin Infect Dis. 2011;53(7):711-5.
230. Berger JR, Chauhan A, Galey D, Nath A. Epidemiological evidence and molecular basis of interactions between HIV and JC virus. J Neurovirol. 2001;7(4):329-38.
231. Major EO, Amemiya K, Tornatore CS, Houff SA, Berger JR. Pathogenesis and molecular biology of progressive multifocal leukoencephalopathy, the JC virus-induced demyelinating disease of the human brain. Clin Microbiol Rev. 1992;5(1):49-73.
232. Khanna N, Wolbers M, Mueller NJ, Garzoni C, Du Pasquier RA, Fux CA, et al. JC virus-specific immune responses in human immunodeficiency virus type 1 patients with progressive multifocal leukoencephalopathy. J Virol. 2009;83(9):4404-11.
233. Brooks BR, Walker DL. Progressive multifocal leukoencephalopathy. Neurol Clin. 1984;2(2):299-313.
234. Berger JR, Levy RM, Flomenhoft D, Dobbs M. Predictive factors for prolonged survival in acquired immunodeficiency syndrome-associated progressive multifocal leukoencephalopathy. Ann Neurol. 1998;44(3):341-9.
235. Khanna N, Elzi, L., Mueller, N.J., Garzoni, C., Cavassini, M., Fux, C.A., Vernazza, P., Bernasconi, E., Battegay, M., Hirsch, H.H., for the Swiss HIV Cohort Study. Incidence and Outcome of Progressive Multifocal Leukoencephalopathy in 20 years of the Swiss HIV Cohort Study. Clinical Infectious Diseases. 2009;48:1459-66.

236. Berger JR, Major EO. Progressive multifocal leukoencephalopathy. Semin Neurol. 1999;19(2):193-200.
237. Whiteman ML, Post MJ, Berger JR, Tate LG, Bell MD, Limonte LP. Progressive multifocal leukoencephalopathy in 47 HIV-seropositive patients: neuroimaging with clinical and pathologic correlation. Radiology. 1993;187(1):233-40.
238. Houff SA, Major EO, Katz DA, Kufta CV, Sever JL, Pittaluga S, et al. Involvement of JC virus-infected mononuclear cells from the bone marrow and spleen in the pathogenesis of progressive multifocal leukoencephalopathy. N Engl J Med. 1988;318(5):301-5.
239. von Einsiedel RW, Samorei IW, Pawlita M, Zwissler B, Deubel M, Vinters HV. New JC virus infection patterns by in situ polymerase chain reaction in brains of acquired immunodeficiency syndrome patients with progressive multifocal leukoencephalopathy. J Neurovirol. 2004;10(1):1-11.
240. Yiannoutsos CT, Major EO, Curfman B, Jensen PN, Gravell M, Hou J, et al. Relation of JC virus DNA in the cerebrospinal fluid to survival in acquired immunodeficiency syndrome patients with biopsy-proven progressive multifocal leukoencephalopathy. Ann Neurol. 1999;45(6):816-21.
241. Koralnik IJ, Boden D, Mai VX, Lord CI, Letvin NL. JC virus DNA load in patients with and without progressive multifocal leukoencephalopathy. Neurology. 1999;52(2):253-60.
242. White FA, 3rd, Ishaq M, Stoner GL, Frisque RJ. JC virus DNA is present in many human brain samples from patients without progressive multifocal leukoencephalopathy. J Virol. 1992;66(10):5726-34.
243. Berger JR. Progressive multifocal leukoencephalopathy in acquired immunodeficiency syndrome: explaining the high incidence and disproportionate frequency of the illness relative to other immunosuppressive conditions. J Neurovirol. 2003;9 Suppl 1:38-41.
244. Khalili K, Gordon J, White MK. The polyomavirus, JCV and its involvement in human disease. Adv Exp Med Biol. 2006;577:274-87.
245. Chowdhury M, Taylor JP, Chang CF, Rappaport J, Khalili K. Evidence that a sequence similar to TAR is important for induction of the JC virus late promoter by human immunodeficiency virus type 1 Tat. J Virol. 1992;66(12):7355-61.

246. Alstadhaug K, Croughs T, Henriksen S, Leboeuf C, Sereti I, Hirsch H, et al. Treatment of progressive multifocal leukoencephalopathy with Interleukin 7. JAMA Neurology; 2014 June 30, 2014.
247. Lautenschlager I, Jahnukainen T, Kardas P, Lohi J, Auvinen E, Mannonen L, et al. A Case of Primary JC Polyomavirus Infection-Associated Nephropathy. Am J Transplant. 2014;14(12):2887-92.
248. Drachenberg CB, Hirsch HH, Papadimitriou JC, Gosert R, Wali RK, Munivenkatappa R, et al. Polyomavirus BK versus JC replication and nephropathy in renal transplant recipients: a prospective evaluation. Transplantation. 2007;84(3):323-30.
249. Boldorini R, Omodeo-Zorini E, Nebuloni M, Benigni E, Vago L, Ferri A, et al. Lytic JC virus infection in the kidneys of AIDS subjects. Mod Pathol. 2003;16(1):35-42.
250. Lopez V, Gutierrez C, Sola E, Garcia I, Burgos D, Cabello M, et al. Does JC polyomavirus cause nephropathy in renal transplant patients? Transplant Proc. 2010;42(8):2889-91.
251. Razonable RR, Brown RA, Humar A, Covington E, Alecock E, Paya CV. A longitudinal molecular surveillance study of human polyomavirus viremia in heart, kidney, liver, and pancreas transplant patients. J Infect Dis. 2005;192(8):1349-54.
252. Wu JH, Nguyen HP, Rady PL, Tyring SK. Molecular insight into the viral biology and clinical features of trichodysplasia spinulosa. Br J Dermatol. 2016;174(3):490-8.
253. Osswald SS, Kulick KB, Tomaszewski MM, Sperling LC. Viral-associated trichodysplasia in a patient with lymphoma: a case report and review. J Cutan Pathol. 2007;34(9):721-5.
254. Babakir-Mina M, Ciccozzi M, Perno CF, Ciotti M. The human polyomaviruses KI and WU: Virological background and clinical implications. APMIS. 2013;121:746–54.
255. Ho J, Jedrych JJ, Feng H, Natalie AA, Grandinetti L, Mirvish E, et al. Human polyomavirus 7-associated pruritic rash and viremia in transplant recipients. J Infect Dis. 2015;211(10):1560-5.
256. Schrama D, Groesser L, Ugurel S, Hafner C, Pastrana DV, Buck CB, et al. Presence of human polyomavirus 6 in mutation-specific BRAF inhibitor-induced epithelial proliferations. JAMA Dermatol. 2014;150(11):1180-6.
257. Delbue S, Elia F, Signorini L, Bella R, Villani S, Marchioni E, et al. Human polyomavirus 6 DNA in the cerebrospinal fluid of an HIV-positive patient with leukoencephalopathy. J Clin Virol. 2015;68:24-7.

258. Carter JJ, Madeleine MM, Wipf GC, Garcea RL, Pipkin PA, Minor PD, et al. Lack of serologic evidence for prevalent simian virus 40 infection in humans. J Natl Cancer Inst. 2003;95(20):1522-30.
259. Knowles WA, Pipkin P, Andrews N, Vyse A, Minor P, Brown DW, et al. Population-based study of antibody to the human polyomaviruses BKV and JCV and the simian polyomavirus SV40. J Med Virol. 2003;71(1):115-23.
260. Cheng J, DeCaprio JA, Fluck MM, Schaffhausen BS. Cellular transformation by Simian Virus 40 and Murine Polyoma Virus T antigens. Semin Cancer Biol. 2009;19(4):218-28.
261. Moens U, Van Ghelue M, Johannessen M. Oncogenic potentials of the human polyomavirus regulatory proteins. Cell Mol Life Sci. 2007;64(13):1656-78.
262. Topalis D, Andrei G, Snoeck R. The large tumor antigen: a "Swiss Army knife" protein possessing the functions required for the polyomavirus life cycle. Antiviral Res. 2013;97(2):122-36.
263. Borchert S, Czech-Sioli M, Neumann F, Schmidt C, Wimmer P, Dobner T, 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):3144-60.
264. Sontag E, Sontag JM, Garcia A. Protein phosphatase 2A is a critical regulator of protein kinase C zeta signaling targeted by SV40 small t to promote cell growth and NF-kappaB activation. Embo J. 1997;16(18):5662-71.
265. Ehlers B, Moens U. Genome analysis of non-human primate polyomaviruses. Infect Genet Evol. 2014;26:283-94.
266. Kwun HJ, Shuda M, Camacho CJ, Gamper AM, Thant M, Chang Y, et al. Restricted Protein Phosphatase 2A Targeting by Merkel Cell Polyomavirus Small T Antigen. J Virol. 2015;89(8):4191-200.
267. Johne R, Muller H. Polyomaviruses of birds: etiologic agents of inflammatory diseases in a tumor virus family. J Virol. 2007;81(21):11554-9.
268. Del Valle L, Khalili K. Detection of human polyomavirus proteins, T-antigen and agnoprotein, in human tumor tissue arrays. J Med Virol. 2010;82(5):806-11.
269. Duncavage EJ, Zehnbauer BA, Pfeifer JD. Prevalence of Merkel cell polyomavirus in Merkel cell carcinoma. Mod Pathol. 2009;22(4):516-21.

270. Stakaityte G, Wood JJ, Knight LM, Abdul-Sada H, Adzahar NS, Nwogu N, et al. Merkel cell polyomavirus: molecular insights into the most recently discovered human tumour virus. Cancers (Basel). 2014;6(3):1267-97.
271. Bouvard V, Baan R, Straif K, Grosse Y, Secretan B, El Ghissassi F, et al. A review of human carcinogens--Part B: biological agents. Lancet Oncol. 2009;10(4):321-2.
272. Lanoy E, Costagliola D, Engels EA. Skin cancers associated with HIV infection and solid-organ transplantation among elderly adults. Int J Cancer. 2010;126(7):1724-31.
273. Toker C. Trabecular carcinoma of the skin. Arch Dermatol. 1972;105(1):107-10.
274. Tilling T, Moll I. Which are the cells of origin in merkel cell carcinoma? J Skin Cancer. 2012;2012:680410.
275. Liu W, MacDonald M, You J. Merkel cell polyomavirus infection and Merkel cell carcinoma. Curr Opin Virol. 2016;20:20-7.
276. Verhaegen ME, Mangelberger D, Harms PW, Vozheiko TD, Weick JW, Wilbert DM, et al. Merkel Cell Polyomavirus Small T Antigen Is Oncogenic in Transgenic Mice. J Invest Dermatol. 2014:1-10.
277. Spurgeon ME, Cheng J, Bronson RT, Lambert PF, DeCaprio JA. Tumorigenic activity of merkel cell polyomavirus T antigens expressed in the stratified epithelium of mice. Cancer Res. 2015;75(6):1068-79.
278. Enam S, Del Valle L, Lara C, Gan DD, Ortiz-Hidalgo C, Palazzo JP, et al. Association of human polyomavirus JCV with colon cancer: evidence for interaction of viral T-antigen and beta-catenin. Cancer Res. 2002;62(23):7093-101.
279. Wieland U, Scola N, Stolte B, Stucker M, Silling S, Kreuter A. No evidence for a causal role of Merkel cell polyomavirus in keratoacanthoma. J Am Acad Dermatol. 2012;67(1):41-6.
280. Wetzels CT, Hoefnagel JG, Bakkers JM, Dijkman HB, Blokx WA, Melchers WJ. Ultrastructural proof of polyomavirus in Merkel cell carcinoma tumour cells and its absence in small cell carcinoma of the lung. PLoS One. 2009;4(3):e4958.
281. Bluemn EG, Paulson KG, Higgins EE, Sun Y, Nghiem P, Nelson PS. Merkel cell polyomavirus is not detected in prostate cancers, surrounding stroma, or benign prostate controls. J Clin Virol. 2009;5:5.
282. Sastre-Garau X, Peter M, Avril MF, Laude H, Couturier J, Rozenberg F, 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):48-56.

283. Cassler NM, Merrill D, Bichakjian CK, Brownell I. Merkel Cell Carcinoma Therapeutic Update. Curr Treat Options Oncol. 2016;17(7):36.
284. Nghiem PT, Bhatia S, Lipson EJ, Kudchadkar RR, Miller NJ, Annamalai L, et al. PD-1 Blockade with Pembrolizumab in Advanced Merkel-Cell Carcinoma. N Engl J Med. 2016;374(26):2542-52.
285. Touze A, Le Bidre E, Laude H, Fleury MJ, Cazal R, Arnold F, et al. High levels of antibodies against merkel cell polyomavirus identify a subset of patients with merkel cell carcinoma with better clinical outcome. J Clin Oncol. 2011;29(12):1612-9.
286. Bouvard V, Baan RA, Grosse Y, Lauby-Secretan B, El Ghissassi F, Benbrahim-Tallaa L, et al. Carcinogenicity of malaria and of some polyomaviruses. Lancet Oncol. 2012;13(4):339-40.
287. Bulut Y, Ozdemir E, Ozercan HI, Etem EO, Aker F, Toraman ZA, et al. Potential relationship between BK virus and renal cell carcinoma. J Med Virol. 2013;85(6):1085-9.
288. zur Hausen H. Red meat consumption and cancer: reasons to suspect involvement of bovine infectious factors in colorectal cancer. Int J Cancer. 2012;130(11):2475-83.
289. Zhang W, Li L, Deng X, Kapusinszky B, Delwart E. What is for dinner? Viral metagenomics of US store bought beef, pork, and chicken. Virology. 2014;468-470:303-10.
290. Peretti A, FitzGerald PC, Bliskovsky V, Buck CB, Pastrana DV. Hamburger polyomaviruses. J Gen Virol. 2015;96(Pt 4):833-9.
291. Imajoh M, Hashida Y, Nakajima H, Sano S, Daibata M. Prevalence and viral DNA loads of three novel human polyomaviruses in skin cancers from Japanese patients. J Dermatol. 2013;40(8):657-60.
292. Scola N, Wieland U, Silling S, Altmeyer P, Stucker M, Kreuter A. Prevalence of human polyomaviruses in common and rare types of non-Merkel cell carcinoma skin cancer. Br J Dermatol. 2012;167(6):1315-20.
293. Antonsson A, Bialasiewicz S, Rockett RJ, Jacob K, Bennett IC, Sloots TP. Exploring the prevalence of ten polyomaviruses and two herpes viruses in breast cancer. PLoS One. 2012;7(8):e39842.
294. Foulongne V, Sauvage V, Hebert C, Dereure O, Cheval J, Gouilh MA, et al. Human skin microbiota: high diversity of DNA viruses identified on the human skin by high throughput sequencing. PLoS One. 2012;7(6):e38499.

295. Bzhalava D, Johansson H, Ekstrom J, Faust H, Moller B, Eklund C, et al. Unbiased approach for virus detection in skin lesions. PLoS One. 2013;8(6):e65953.
296. Kreuter A, Silling S, Dewan M, Stucker M, Wieland U. Evaluation of 4 recently discovered human polyomaviruses in primary cutaneous B-cell and T-cell lymphoma. Arch Dermatol. 2011;147(12):1449-51.
297. Du-Thanh A, Foulongne V, Guillot B, Dereure O. Recently discovered human polyomaviruses in lesional and non-lesional skin of patients with primary cutaneous T-cell lymphomas. J Dermatol Sci. 2013;71(2):140-2.
298. Neu U, Maginnis MS, Palma AS, Stroh LJ, Nelson CDS, Feizi T, et al. Structure-Function Analysis of the Human JC Polyomavirus Establishes the LSTc Pentasaccharide as a Functional Receptor Motif. Cell Host & Microbe. 2010;8(4):309-19.
299. Stroh LJ, Gee GV, Blaum BS, Dugan AS, Feltkamp MC, Atwood WJ, et al. Trichodysplasia spinulosa-Associated Polyomavirus Uses a Displaced Binding Site on VP1 to Engage Sialylated Glycolipids. PLoS Pathog. 2015;11(8):e1005112.
300. Nelson CD, Derdowski A, Maginnis MS, O'Hara BA, Atwood WJ. The VP1 subunit of JC polyomavirus recapitulates early events in viral trafficking and is a novel tool to study polyomavirus entry. Virology. 2012;428(1):30-40.
301. Li TC, Takeda N, Kato K, Nilsson J, Xing L, Haag L, et al. Characterization of self-assembled virus-like particles of human polyomavirus BK generated by recombinant baculoviruses. Virology. 2003;311(1):115-24.
302. Teunissen EA, de Raad M, Mastrobattista E. Production and biomedical applications of virus-like particles derived from polyomaviruses. J Control Release. 2013;172(1):305-21.
303. Norkiene M, Stonyte J, Ziogiene D, Mazeike E, Sasnauskas K, Gedvilaite A. Production of recombinant VP1-derived virus-like particles from novel human polyomaviruses in yeast. BMC Biotechnol. 2015;15:68.
304. Hale AD, Bartkeviciute D, Dargeviciute A, Jin L, Knowles W, Staniulis J, et al. Expression and antigenic characterization of the major capsid proteins of human polyomaviruses BK and JC in Saccharomyces cerevisiae. J Virol Methods. 2002;104(1):93-8.

305. Barth H, Solis M, Kack-Kack W, Soulier E, Velay A, Fafi-Kremer S. In Vitro and In Vivo Models for the Study of Human Polyomavirus Infection. Viruses. 2016;8(10):1-17.
306. Broekema NM, Imperiale MJ. Efficient propagation of archetype BK and JC polyomaviruses. Virology. 2012;422(2):235-41.
307. Buck CB, Pastrana DV, Lowy DR, Schiller JT. Efficient intracellular assembly of papillomaviral vectors. J Virol. 2004;78(2):751-7.
308. Jeffers-Francis LK, Burger-Calderon R, Webster-Cyriaque J. Effect of Leflunomide, Cidofovir and Ciprofloxacin on replication of BKPyV in a salivary gland in vitro culture system. Antiviral Res. 2015;118:46-55.
309. Hanssen Rinaldo C, Hansen H, Traavik T. Human endothelial cells allow passage of an archetypal BK virus (BKV) strain--a tool for cultivation and functional studies of natural BKV strains. Arch Virol. 2005;150(7):1449-58.
310. Alsanie WF, Niclis JC, Petratos S. Human embryonic stem cell-derived oligodendrocytes: protocols and perspectives. Stem Cells Dev. 2013;22(18):2459-76.
311. Major EO, Miller AE, Mourrain P, Traub RG, de Widt E, Sever J. Establishment of a line of human fetal glial cells that supports JC virus multiplication. Proc Natl Acad Sci U S A. 1985;82(4):1257-61.
312. Henriksen S, Tylden GD, Dumoulin A, Sharma BN, Hirsch HH, Rinaldo CH. The human fetal glial cell line SVG p12 contains infectious BK polyomavirus. J Virol. 2014;88(13):7556-68.
313. Hara K, Sugimoto C, Kitamura T, Aoki N, Taguchi F, Yogo Y. Archetype JC virus efficiently replicates in COS-7 cells, simian cells constitutively expressing simian virus 40 T antigen. J Virol. 1998;72(7):5335-42.
314. Schowalter RM, Pastrana DV, Buck CB. Glycosaminoglycans and sialylated glycans sequentially facilitate Merkel cell polyomavirus infectious entry. PLoS Pathog. 2011;7(7):e1002161.
315. Feng H, Kwun HJ, Liu X, Gjoerup O, Stolz DB, Chang Y, et al. Cellular and viral factors regulating Merkel cell polyomavirus replication. PLoS One. 2011;6(7):e22468.
316. Neumann F, Borchert S, Schmidt C, Reimer R, Hohenberg H, Fischer N, et al. Replication, Gene Expression and Particle Production by a Consensus Merkel Cell Polyomavirus (MCPyV) Genome. PLoS One. 2011;6(12):e29112.

317. Liu W, Yang R, Payne AS, Schowalter RM, Spurgeon ME, Lambert PF, et al. Identifying the Target Cells and Mechanisms of Merkel Cell Polyomavirus Infection. Cell Host Microbe. 2016;19(6):775-87.
318. Israel MA, Chan HW, Hourihan SL, Rowe WP, Martin MA. Biological activity of polyoma viral DNA in mice and hamsters. J Virol. 1979;29(3):990-6.
319. Pastrana DV, Ray U, Magaldi TG, Schowalter RM, Cuburu N, Buck CB. Erratum for Pastrana et al., BK Polyomavirus Genotypes Represent Distinct Serotypes with Distinct Entry Tropism. J Virol. 2016;90(1):624.
320. Khan ZM, Liu Y, Neu U, Gilbert M, Ehlers B, Feizi T, et al. Crystallographic and glycan microarray analysis of human polyomavirus 9 VP1 identifies N-glycolyl neuraminic acid as a receptor candidate. J Virol. 2014;88(11):6100-11.
321. Stroh LJ, Neu U, Blaum BS, Buch MH, Garcea RL, Stehle T. Structure analysis of the major capsid proteins of human polyomaviruses 6 and 7 reveals an obstructed sialic Acid binding site. J Virol. 2014;88(18):10831-9.
322. Akan I, Sariyer IK, Biffi R, Palermo V, Woolridge S, White MK, et al. Human polyomavirus JCV late leader peptide region contains important regulatory elements. Virology. 2006;349(1):66-78.
323. Myhre MR, Olsen GH, Gosert R, Hirsch HH, Rinaldo CH. Clinical polyomavirus BK variants with agnogene deletion are non-functional but rescued by trans-complementation. Virology. 2010;398(1):12-20.
324. Cartharius K, Frech K, Grote K, Klocke B, Haltmeier M, Klingenhoff A, et al. MatInspector and beyond: promoter analysis based on transcription factor binding sites. Bioinformatics. 2005;21(13):2933-42.
325. Ravichandran V, Major EO. DNA-binding transcription factor NF-1A negatively regulates JC virus multiplication. J Gen Virol. 2008;89(Pt 6):1396-401.
326. Ravichandran V, Sabath BF, Jensen PN, Houff SA, Major EO. Interactions between c-Jun, nuclear factor 1, and JC virus promoter sequences: Implications for viral tropism. J Virol. 2006;80(21):10506-13.
327. Bhattacharjee S, Chattaraj S. Entry, infection, replication, and egress of human polyomaviruses: an update. Can J Microbiol. 2017;63(3):193-211.
328. Polyomaviridae Study Group of the International Committee on Taxonomy of V, Calvignac-Spencer S, Feltkamp MC, Daugherty MD, Moens U, Ramqvist T, et al. A taxonomy update for the family Polyomaviridae. Arch Virol. 2016.

329. Cole CN. Polyomavirinae: The Viruses and Their Replication. In: Bernard N. Fields DMK, Peter M. Howley, editor. Fundamental Virology. 3rd Edition ed: Lippincott-Raven Publishers; 1996. p. 917-45.
330. Mitchell PJ, Wang C, Tjian R. Positive and negative regulation of transcription in vitro: enhancer-binding protein AP-2 is inhibited by SV40 T antigen. Cell. 1987;50(6):847-61.
331. Spence SL, Pipas JM. Simian virus 40 large T antigen host range domain functions in virion assembly. J Virol. 1994;68(7):4227-40.
332. Kern FG, Pellegrini S, Cowie A, Basilico C. Regulation of polyomavirus late promoter activity by viral early proteins. J Virol. 1986;60(1):275-85.
333. Harris KF, Chang E, Christensen JB, Imperiale MJ. BK virus as a potential co-factor in human cancer. Dev Biol Stand. 1998;94:81-91.
334. Love TM, de Jesus R, Kean JA, Sheng Q, Leger A, Schaffhausen B. Activation of CREB/ATF sites by polyomavirus large T antigen. J Virol. 2005;79(7):4180-90.
335. Smith RW, Nasheuer HP. Initiation of JC virus DNA replication in vitro by human and mouse DNA polymerase alpha-primase. Eur J Biochem. 2003;270(9):2030-7.
336. Bullock PA. The initiation of simian virus 40 DNA replication in vitro. Crit Rev Biochem Mol Biol. 1997;32(6):503-68.
337. Campbell KS, Mullane KP, Aksoy IA, Stubdal H, Zalvide J, Pipas JM, et al. DnaJ/hsp40 chaperone domain of SV40 large T antigen promotes efficient viral DNA replication. Genes Dev. 1997;11(9):1098-110.
338. Weisshart K, Bradley MK, Weiner BM, Schneider C, Moarefi I, Fanning E, et al. An N-terminal deletion mutant of simian virus 40 (SV40) large T antigen oligomerizes incorrectly on SV40 DNA but retains the ability to bind to DNA polymerase alpha and replicate SV40 DNA in vitro. J Virol. 1996;70(6):3509-16.
339. Deb S, Tsui S, Koff A, DeLucia AL, Parsons R, Tegtmeyer P. The T-antigen-binding domain of the simian virus 40 core origin of replication. J Virol. 1987;61(7):2143-9.
340. Bethge T, Hachemi HA, Manzetti J, Gosert R, Schaffner W, Hirsch HH. Sp1 sites in the noncoding control region of BK polyomavirus are key regulators of bidirectional viral early and late gene expression. J Virol. 2015;89(6):3396-411.
341. DiMaio D, Nathans D. Regulatory mutants of simian virus 40. Effect of mutations at a T antigen binding site on DNA replication and expression of viral genes. J Mol Biol. 1982;156(3):531-48.

342. Lashgari MS, Tada H, Amini S, Khalili K. Regulation of JCVL promoter function: transactivation of JCVL promoter by JCV and SV40 early proteins. Virology. 1989;170(1):292-5.
343. Wormke M, Stoner M, Saville B, Walker K, Abdelrahim M, Burghardt R, et al. The aryl hydrocarbon receptor mediates degradation of estrogen receptor alpha through activation of proteasomes. Mol Cell Biol. 2003;23(6):1843-55.
344. Briggs MR, Kadonaga JT, Bell SP, Tjian R. Purification and biochemical characterization of the promoter-specific transcription factor, Sp1. Science. 1986;234(4772):47-52.
345. Simmons DT. SV40 large T antigen functions in DNA replication and transformation. Adv Virus Res. 2000;55:75-134.
346. DeCaprio JA, Imperiale MJ, Major EO. Polyomaviruses. In: Knipe DM, Howley P, editors. Fields Virology, 6th Edition. 6th Edition ed2013. p. 1633-61.
347. Schowalter RM, Reinhold WC, Buck CB. Entry tropism of BK and Merkel cell polyomaviruses in cell culture. PLoS One. 2012;7(7):e42181.
348. Shuda M, Feng H, Kwun HJ, Rosen ST, Gjoerup O, Moore PS, 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):16272-7.

9 GLOSSARY

Antibody: A protein secreted by mainly plasma cells for the neutralization of foreign agents, such as bacteria and virus.

Acute infection: A relatively rapid disease onset, brief symptoms and disease cleared within days.

Antigen: Any foreign molecule that can induce an immune response, especially producing antibodies.

B recognition element: Is a cis-regulatory element immediately located upstream of TATA-box.

Cancer: Abnormal cell growth with the ability to spread or invade other compartment of the body.

Capsid: A protein shell for protecting a virus, consisting of small units called capsomeres

Core promoter: It is portion of the proximal sequence upstream of a gene harboring the transcriptional start sites (TSS). It also contains the following; regulatory factors, RNA polymerase binding motif and approximately –34 base pair upstream from the TSS.

Core promoter elements: These are sequence motifs such as initiation elements, TATA-box, downstream promoter elements (DPEs) and B recognition element (BRE) within the core promoter of a gene vital for gene transcription.

Disease: a disorder that affects part or the whole body of an organism associated with signs and symptoms.

Deoxyribonucleic acid: Is a macromolecule that contains the genetic information vital in the development, growth, reproduction and functioning of all living organisms including viruses.

Downstream promoter element: Is one of the core promoter elements vital for gene transcription by polymerase II, located downstream of the initiation element.

Developmental or regulated promoter: Is a promoter with a focused TSS, TATA-box or downstream promoter element and activated by a stimulus.

Dyad symmetry: These are two regions of a DNA sequence with inverted repeats of each other.

Encapsidation: It is a process of enclosing a viral genome in a capsid.

ERAD pathway: This is a cellular pathway targeting misfolded protein from endoplasmic reticulum for proteasomal degradation prior ubiquitination.

Housekeeping or tissue-restricted promoter: Is promoter with dispersed TSS, enriched with upstream BRE, contain CpG-island and constitutively active.

Infection: This is the invasion of the tissues of an organism by a disease-causing agent.

Initiation element: It is a core promoter element located downstream of a TATA box and upstream of a DPE, which initiate transcript even without a functioning TATA-box.

Natural killer (NK) cells: These are immune cells of the innate immune system exhibiting various activating and inhibitory surface receptors controlling their ability of kill target cells, activation and produce cytokines.

Nuclear localization signals: These are short amino acid sequences required for nuclear translocation of a protein.

Open reading frame: It is a continuous region of a codon with start and stop codons that can encode potential proteins.

Palindrome sequence: A sequence on DNA/RNA where a 5´to 3´forward sequence matches a sequence 5´to´3 reverse on a double helix.

Pathogen: A biological agent that can cause disease.

Persistent infection: This is a life-long infection which occur when a primary infection cannot be cleared completely by the host immune system.

Promoter: DNA regions that control regulation and the initiation of transcription.

Protein: A biomolecule consisting of a chain or chains of amino acids.

Ribonucleic acid: It is a biomolecule with long chains of nucleotides for protein synthesis and containing genetic information like DNA.

Recombination: A process of recombining genetic material by genetic engineering.

Seroprevalence: Is the number of individuals in a population who are positive for a particular disease based on antibody presence in the blood against that specific disease.

TATA-box: It is an AT-rich sequence located in the promoter of genes upstream of the TSS.

Transcription start site: It is the region of the 3' end of a gene where transcription starts.

Transcription factors: These are proteins that regulate the rate of transcription of DNA to mRNA, by specifically binding to a recognition DNA sequence.

Viral reactivation: Active viral replication after a period of viral dormancy.

Viremia: Presence of viral DNA in the blood.

Viruria: Presence of viral DNA in the urine.

Virus: A small-infectious agent that can only replicate inside a living cell.

10 ACKNOWLEDGEMENTS

Firstly, I would like to thank Prof. Hans H. Hirsch who gave me the opportunity to do a PhD with him. More so, I am grateful for his scientific supports and supervision which has taught me a lot not only on scientific issues but also on social issues.

Secondly, sincere thanks go to my Faculty representative, Prof. Marcel Tanner and my co-referee Prof. Volker Thiel for being part of my PhD committee and for their scientific in puts during my committee meetings, and also Prof. Hans-Peter Beck for accepting to be the chair of my PhD defense.

I would like to thank all my colleagues for the helpful discussions during our lab meetings and of course the Norwegian research group of Prof. Rinaldo Christine for their nice discussions during our journal clubs, which has helped me a lot with respect to my scientific works.

Immerse thanks go to my wife Dr Nicole and her family for their supports throughout the period of my PhD. Special thanks go to my mom Grace for all the nice foods she sent from Cameroon to keep me strong and all my siblings back home especially Sylvester for their encouragements and of course my friends and relatives in Switzerland, Mathias, Calista, Clovis, Brenda, Vero, Nelvis for all the nice and relaxing moments.

Last but not the least, I thank the almighty God for the strength to accomplish this work.

11 CURRICULUM VITAE

Name:	Ajuh Elvis Tasih
Date and place of birth:	Born 13. October 1986, in Cameroon
Address:	Schleifenbergstrasse 35, 4058 Basel
Nationality:	Cameroonian
Marital status:	Married
Phone contact:	+41779696222
E-mail (personal):	tasihe@yahoo.com
E-mail (professional):	elvis.ajuh@unibas.ch

Educational qualifications

Oct. 2013-Sept. 2017	PhD in Microbiology, department of biomedicine, University of Basel, Switzerland
Sept. 2011-Feb. 2013	Master of Science (MSc) in Infection Biology, Swiss Tropical and Public Health Institute (SwissTPH), University of Basel, Switzerland
Sep. 2007- Jul. 2010	Bachelor of Science (BSc) in Microbiology, University of Buea, Cameroon
2005-2007	G.C.E advanced level, Bilingual Grammar School Molyko, Buea, Cameroon
2000-2005	G.C.E Ordinary level, Government Secondary School Bombe Bakundu, Cameroon

Conferences presented

1. 25th annual meeting of the society for Virology, Bern Switzerland-Poster presentation
2. 74th annual meeting of Swiss Society for Microbiology-Bochum, Germany
3. 6th European Congress of Virology- Hamburg, Germany-Poster presentation
4. 1st Symposium of the Geneva Center for Emergence Viral Diseases-Geneva, Switzerland
5. Joint Annual Meeting SSM, SSI, SSHH, SSTMP, SSTTM, 2017, Basel, Switzerland-Poster presentation

Publications

Elvis T. Ajuh, Tobias Bethge, Emma Kraus, Nicole Fischer, Fabian H. Weissbach, Rainer Gosert, Hans H. Hirsch*. *Novel Human Polyomavirus non-coding control regions differ in bi-directional gene expression according to host cell, large T-antigen expression and clinically occurring rearrangements* (Manuscript under review).

Tobias Bethge, **Elvis T. Ajuh,** Hans H. Hirsch* (2016) *Imperfect Symmetry of Sp1 and Core Promoter Sequences Regulates Early and Late Viral Gene Expression of the Bi-Directional BK Polyomavirus Non-Coding Control Region* J. Virol 90: 10101.

David C. Müller, Maarit Rämö, Christian Wetterauer, Valeria Perrina, Luca Quagliata, Tatjana Vlajnic, Christian Ruiz, Beate Balitzki, Rainer Grobholz, Rainer Gosert, **Elvis T. Ajuh**, Hans H. Hirsch, Lukas Bubendorf, Cyrill A. Rentsch. *Donor-derived, metastatic urothelial cancer after kidney transplantation associated with BK polyomavirus bearing a critical deletion in the viral non-coding control region* (Manuscript under revision).