Spring 4-17-2013

Telomere-Related Factors and Human Papillomavirus Genome Maintenance

Adam Rogers
University of Nebraska-Lincoln, arogers4@huskers.unl.edu

Telomere-Related Factors and Human Papillomavirus Genome Maintenance

By

Adam J. Rogers

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Doctor of Philosophy

Major: Biological Sciences

Under the Supervision of Professor Peter C. Angeletti

Lincoln, Nebraska

May, 2013

Telomere-Related Factors and Human Papillomavirus Genome Maintenance

Adam Jeffrey Rogers, Ph.D.

University of Nebraska, 2013

Advisor: Peter C. Angeletti

Human papillomaviruses (HPVs) are small DNA tumor viruses identified by their characteristic ability to replicate as a nuclear plasmid in mitotically active basal keratinocytes. A key characteristic of the HPV life cycle is the establishment of a stable maintenance phase wherein the virus replicates at low copy number, which likely occurs in cells expressing little to no E1 and E2. It is thought that HPV16 replicates in a once-per-cell-cycle manner during this portion of its life cycle and presumably interacts with host chromosome replication and maintenance factors to facilitate this replication. The adaptive radiation of papillomaviruses in response to changing host factors was well demonstrated in this work with an examination of the evolution of the Papillomaviruses' E2 proteins and cognate binding sites as the virus has adapted to infect new body tissues. Additionally, the yeast model of HPV replication we utilize in our laboratory showed a varying ability to replicate in S. cerevisiae, again demonstrating that the replication environment plays a significant role in the long-term success of papillomaviruses.

To further investigate these cellular factors, we investigated the role telomeric maintenance factors may play in these processes. We have performed ChIP assays that have shown that components of the telomere maintenance complex (shelterin) can bind to at least four sites in the HPV genome, each of which contain nine-base telomere-repeat sequences

(TTAGGGTTA). We have shown that mutating these sites has a detrimental effect on the virus's ability to replicate under certain conditions. The shelterin complex interacts with a number of important chromosome replication and maintenance proteins with such diverse functions as DNA replication, chromosome segregation, and DNA repair, making it an ideal target for coercion by a DNA virus utilizing a stable low-copy replication strategy. Interaction between Telomer Repeat Binding Factor 2 (TRF2) and Epstein Barr Virus Nuclear Antigen 1 (EBNA1) protein (a structural and functional homologue of E2) is required for replication of plasmids containing the Epstein Barr Virus latent origin of replication. Kaposi's Sarcoma Herpesvirus (KSHV) Latency Associated Nuclear Antigen (LANA) protein (another homologue of E2) also interacts with TRF2. Results from our Far-Western and co-immunoprecipitation assays show that E2 interacts with TRF2 and other shelterin components. In summary, these results suggest that TRF2, TRF1, Rap1, Pot1, and Tin2, plus certain DNA repair proteins, may regulate the maintenance phase of the HPV lifecycle. E2 appears to be capable of mediating these interactions.

Acknowledgements

First, I would like to express my appreciation to my advisor, Dr. Peter Angeletti, for his guidance and support through the years of my doctoral study. I would like to thank Dr. Clinton Jones for inspiring me to enter the field of Virology and for continually providing an example of the kind of scientist I strive to become. I would also like to thank the other members of my Doctoral Committee: Dr. Fernando Osorio, Dr. Luwen Zhang, and Dr. Melanie Simpson for constructive critiques and their guidance throughout my graduate career. I would also like to thank Dr. Daraporn Pittayakhajonwut, my partner in crime, for teaching me innumerable laboratory techniques and for helping to make this work possible. I am grateful to my lab mates Willie Hughes, John Lowe, Greetchen Diaz, Chrispin Chisanga, and many others for providing me with ideas, support, and friendship.

I would like to extend my gratitude to Joe Zhou and Terri Fangman from the UNL Microscopy Core Facility for their patience and invaluable feedback in assisting me with confocal microscopy.

I would also like to express my gratitude to my family for their patience and unwavering support throughout the graduate school process, particularly my fiancée Jen who has been everything I could ask for in a partner. I would also like to thank Charlie, Katie, Scott, Cindy and Trudy for helping to keep me sane through the process of completing the dissertation. And finally, to my Father, for never failing to push me forward to fulfill the potential he's always seen in me.

Yes Dad, the thesis is finally done.

Contents

Chapter 1
Literature Review

Papillomaviruses

Papillomaviruses (PV) are small (55 nm diameter) non-enveloped viruses of icosahedral capsid symmetry that house a single molecule of circular supercoiled double stranded DNA (339). PVs have coevolved extensively with their mammalian hosts, such that there have been human papillomaviruses (HPVs) essentially since the evolutionary emergence of humans (31). This family of viruses infects body surface tissues such as the skin or mucosal surfaces, which include the mouth, airways, and anogenital tissues of vertebrate animals (75). According to the International Committee on Taxonomy of Viruses, the family papillomaviridae has 16 assigned genera (alphapapillomavirus through pipapillomavirus) and one unassigned genus (81). There are over 100 strains of HPV identified at present (141) as well as numerous papillomaviruses that infect mammals, birds, and reptiles. Papillomavirus strains are further classified by differences in the major capsid protein sequence L1. New papillomavirus types are recognized if the L1 gene has 10% difference in DNA sequence as compared to the closest known PV type, a 2-10% difference constitutes a subtype and < 2% difference defines a variant (82, 339). Alphapapillomaviruses are further classified into high and low risk categories by their potential to cause cervical cancer (57, 220, 313).

Mucosal HPV types are the causative agents of cervical cancer as well as some vaginal, anal, and penile cancers (37, 57, 220), typically as a result of genomic integration and resultant overexpression of the viral E6 and E7 oncogenes(306). This was initially described in 1982 with detection of HPV sequences in human tumors (128, 131, 332). High risk HPV DNA is found in greater than 95% of cervical cancers, making it clear that this virus is the causative agent of this disease (221, 307). Of these, HPVs 16 and 18 are the most common causes, with HPV16 being present in over 50%, and

HPV18 10-12% of cervical cancers worldwide (36). HPV18 has been found in certain geographic regions to be more prevalent than HPV16 (224). Additionally, emerging research has implicated HPVs in some head and neck, anogenital, upper respiratory, and even some non-melanoma skin cancers (68, 126, 152, 344).

Life Cycle of Papillomaviruses

The life cycle of papillomavirus is closely associated with the differentiation profile of host keratinocytes. During the course of epithelial cell differentiation, the virus shifts through three replication phases in response to keratinization: establishment, wherein early viral replication occurs; maintenance, where the viral genome is stably maintained episomally by replicating through a theta intermediate; and amplification, where viral replication shifts to a rolling-circle mode and the copy number increases in preparation for encapsidation (112). The virus typically enters the host through wounding of the external epidermis, wherein the virus travels to the mitotically active basal keratinocytes to establish infection. HPV binds to a number of sub-receptors before eventually adhering to the candidate receptor α6 integrin and enters through clathrin-mediated endocytosis (39, 75, 104, 143, 203, 277). Upon nuclear entry, the viral DNA is transferred to the nucleus and promptly targeted to PML bodies, which enhances transcription and are important for successful viral replication (74).

Upon establishing infection in the host, the virus shifts into the maintenance phase of the life cycle. Gene expression of most of the open reading frames besides the viral oncogenes E6 and E7 is repressed, and the virus is maintained at a copy number of less than 20 per cell (105). Studies have demonstrated that the viral replication proteins (E1 and E2) are required for establishment of viral infection, but may be dispensable during the maintenance phase (159, 245). During this phase, the virus replicates in a

once-per-cell-cycle fashion which is reliant on a number of host factors for stability (114). The virus can thus be maintained indefinitely in the basal layers of the epithelium.

As the cells of the basal layer divide, one cell remains attached to the basal layer while the other migrates up through the squamous epithelia and begins to progress through the keratinization and differentiation processes, which drives the HPVs into the vegetative phase of their life-cycle. This phase is characterized by activation of the keratin dependant promoter activity, increased expression of the viral proliferative genes, and a shift from a once-per-mitosis bidirectional theta replication model of viral genome replication to a rolling-circle mechanism along with an increase in expression of the L1 and L2 capsid genes (72, 114).

The HPV Genome

HPV has a small 8kb genome that consists of a long control region (LCR), early gene region and a late gene region. The LCR is ~850 bp long and comprises 10% of the viral genome. The LCR contains the origin of replication (ori) and multiple transcription binding sites, thus controlling the expression of viral genes (141, 339). The early gene region comprises 50% of the viral genome and encodes 6 open reading frames (ORFs). The expression of genes in this and the late region is regulated by alternative splicing due to the compact genome of the virus. The late gene region is the last 40% of the viral genome and is downstream of the early region (339). Early genes are expressed in undifferentiated or newly differentiated keratinocytes and late region genes are expressed in keratinocytes undergoing terminal differentiation (202, 339).The early genes (E1, E2, E6 and E7) are primarily responsible for replication, genome maintenance, and the promotion of cell growth.

The viral E5 protein plays an important role in stimulating cellular proliferation. The protein activates ligand independent activation of the growth factor receptor, down-regulates Major Histocompatibility Complex I (MHCI), and deregulates the Golgi body (11, 94, 334). E6 and E7 are bona fide oncogenes (192, 261). E7 binds to and deactivates the Retinoblastoma (RB) cell-cycle regulation protein as well as other members of the RB family of proteins in order to release E2F and progress the cell cycle. E6 binds to and, with the action of an E3-Ubiquitin Ligase E6 Associated Protein (E6AP), degrades p53 to down-regulate apoptosis. These genes combine to force the cell into an S/G2-like state to promote cell-cycle progression, down-regulate terminal differentiation processes, and promote viral genome replication.

The E4 protein is usually expressed as a spliced mRNA as part of the E1 transcript and is typically expressed in highly-differentiated cells. Its function is to sustain the S-phase like state induced by other viral factors. E4 weakens virus laden cornified envelopes and keratin filaments and alters the host cytoskeletal structure as a means of promoting viral egress (44, 73, 98).

The L1 and L2 ORFs encode the major and minor capsid proteins, and are expressed in the late-stages of the viral lifecycle (62, 243). L1 is the major capsid protein and spontaneously organizes itself into pentameric structures that make up the majority of the viral icosahedral capsid (109, 161, 341). L2 is the minor capsid protein and plays a role in protein packaging, transporting DNA to the nucleus after viral entry, and recruitment of E2 to viral replication foci (74, 76, 335).

E2

The E2 protein serves multiple functions, with its primary role being that of a transcriptional regulator. It is a 350 to 500 amino acid protein expressed from a spliced mRNA transcribed from different promoters. E2 also negatively regulates the expression

of the E6 and E7 oncoproteins(184). E2 polypeptides contain three probable domains: a DNA binding domain (DBD) located at the C-terminus, an N-terminus transactivation domain, and an internal “hinge” domain. Both the C-terminal and N-terminal domains are relatively well conserved within the PVs (254). E2 binds as a dimer to DNA-binding sites through action mediated by the DBD (202). The E2 DBD forms a dimeric β-barrel, with each strand contributing a half-barrel. The dimer interface has a hydrophobic core and uses extensive hydrogen bonding between subunits to maintain tight binding. This β-barrel core contains elaborately packed side chains that contribute to the stability of the dimer. There is a poorly conserved loop connecting β-strands 2 and 3. This loop varies from 6-10 residues. The tertiary structure of characterized E2 DBDs is similar, but there appears to be variation in the orientations of the two subunits (141).

The transcriptional repression activity occurs as a result of E2 binding sites overlapping TATA boxes or Sp1 sites and requires a functional transactivation domain (30, 298). Conversely, E2 has transcriptional activation functions by inserting E2 binding sites (E2BSs) upstream of thymidine kinase promoters in a Chloramphenicol Acetyl Transferase (CAT) assay (52). Much of E2s ability to affect transcription comes from its interaction with the Bromodomain 4 protein (Brd4), an essential cellular protein that binds to acetylated lysines of Histone H3 and H4 (91) and stimulates RNA polymerase II transcription (330, 331). E2 redistributes Brd4 into punctate dots scattered along the chromosome (208) and increases Brd4’s affinity for binding to interphase chromatin (208). Presumably, this is an effect of E2 utilizing Brd4 as a transcription factor (206). There is evidence to suggest that the activation domain mediated oligomerization could influence interaction between E2 molecules bound at distant E2-binding sites forming DNA loops and other DNA structures (8, 141).

Expression levels of E2 in infected cells are a matter of some debate. High levels of E2 expression in mammalian cells leads to activation of caspase-8 dependant

apoptosis (90, 118). While this seems counterintuitive, it is believed this interaction is related to an observed role played by caspase-8 in differentiation of keratinocytes (118). As such, only one cell line has been found which stably expressed E2 with episomal HPV DNA (282) and it has never been observed in raft culture experiments utilizing HPV16 and 18 (28). Observations of HPV infected cervical biopsies have determined that E2 expression is primarily isolated to the more differentiated layers of low-grade Cervical Intraepithelial Neoplasia (CIN) lesions but not in proliferating cells (28). High levels of E2 expression thus appears to be restricted to differentiated cells in the epithelium with high levels of keratin expression, but not in rapidly dividing cells (320).

One of E2's most important functions is to tether the viral genome to replicating chromosomes during mitosis to ensure proper segregation of viral plasmids to daughter cells (201). Initially, this was theorized to be the result of an interaction with the Bromodomain4 protein (BRD4), which is required for plasmid segregation and transcriptional control in Bovine Papillomavirus 1. However, further studies demonstrated that, while all papillomavirus E2 proteins studied thus far utilize Brd4 for transcriptional purposes, it is dispensable for the plasmid maintenance in HPVs (206). The maintenance function of E2 thus remains somewhat cryptic. Although binding of E2 to mitotic chromosomes is consistently observed, the binding location is not conserved amongst PV types (231). Some studies have demonstrated that the beta-papillomavirus HPV8 E2 protein has a binding preference for the rDNA gene region of acrocentric chromosomes (231, 246) which contains a number of binding sites for E2 and, due to its unique chromatin structure, can allow E2 to remain transcriptionally active (67, 119, 164). By contrast, the BPV1 E2 protein is observed during this phase as discreet speckles associated with the papillomavirus genome scattered along the chromosome (274). Alpha-papillomaviruses had a similar binding localization to HPV8, but while other PVs are observed associated with the chromosome throughout mitosis, alpha-

papillomavirus genomes are only observed attached during prophase and telophase (97, 120, 231). This association is observed in other DNA viruses with long-term replication strategies, as Epstein Barr Virus (EBV) Nuclear Antigen 1 (EBNA1) protein binds to a pre-mRNA processing protein and localizes to the nucleolus during interphase ((50, 269) and Kaposi's Sarcoma Herpesvirus (KSHV) Latency Associated Nuclear Antigen (LANA) binds to pericentromeric and telomeric regions of DNA similarly to E2 (231).

The consensus sequence 5'-ACCgNNNNcGGT-3' is recognized by E2, with the position 4 and 9 residues allowing some variability. A number of studies have been performed to examine the binding of E2 protein to its cognate binding site (33, 83, 111, 141, 144, 249, 257, 293). The 4-nucleotide spacer sequence varies by HPV type, and has been identified as being critical for determining E2 binding affinity through indirect readout as well as playing a potential role in gene regulation, despite having no predicted nucleotide-amino acid contacts from crystal structure (33, 83, 141, 144, 328). E2 binds DNA as a homodimer with each monomer supplying an alpha helix to contact two successive major grooves of the target site (83, 141)

Four typical E2 binding sites are conserved in the upstream regulatory region (URR) of most papillomaviruses, numbered according to their distance from the early promoter (202). Each site is differentially regulated and demonstrates variable binding affinity for the E2 protein, resulting in varying replication and transcriptional effects during the viral life cycle (63, 191) presumably as a result of differences in E2 binding affinity (141) due to sequence variation as well as methylation of the E2 binding site (257, 293). These binding sites are typically well conserved across all papillomaviruses. However, in some cases variation in the number and location of some E2 binding sites does exist, including a predicted fifth binding site within the URR of betapapillomaviruses (103) and some alphapapillomaviruses (257) as well as up to 17 sequences with ability to bind E2 from the URR of bovine papillomavirus 1 (22, 71, 170, 249, 273).

HPV Replication

Papillomavirus DNA replication requires primarily cellular factors, recruiting polymerase alpha along with other elements of the cellular replication machinery (240). The accepted model relies on the viral proteins E1 and E2 for genomic amplification (95, 250, 301). The viral origin of replication contains a minimum of one E1 and E2 protein binding site (301). The E1 protein functions as an ATP dependent helicase and recruits DNA polymerase alpha to act as an elongation factor (117, 217, 242). It is regulated by extensive post-translational modification (172). E1 unwinds the viral genome through its ATP-dependant helicase activity.

E1 is loaded onto the origin by the E2 protein, which is localized to the HPV replication foci by the L2 protein (76). The two early proteins bind as dimers cooperatively through an interaction between the N-terminus of E2 and the helicase domain of E1 (259, 265). After loading E1, E2 dissociates from the viral genome while E1 forms into a hexameric helicase ring, similar to that formed by the minichromosome maintenance proteins (258, 262, 264, 296). At this point, E1 recruits Replication Protein A (RPA), topoisomerase I, polymerase α primase, Proliferating Cell Nuclear Antigen (PCNA), Replication factor C, and DNA Polymerase δ. (58, 61, 137, 166, 185, 197, 210, 219). E2 recruits a number of factors important for viral replication including Transcription Factor II β, Transcription Factor IIδ, Activation Domain Modulation Factor-1 (AMF-1), Breast Cancer Associated protein 1 (BRCA1), Poly [ADP Ribose] Polymerase 1 (PARP1), Transcription initiation factor TFIID subunit 1 (TAF1), Topoisomerase I (TopoI), DNA topoisomerase 2-binding protein 1 (TopBPI), and chromatin remodeling P300/CBP-associated factor (p/CAF), transcriptional co-activators p300/CBP, Brahma (BRM), hSNF5 and Nucleosome Assembly Protein (NAP-1), as well as Bromodomain-

Containing Protein 4 (BRD4) (reviewed in (27)). Once replication is completed, it is suggested that an interaction between E2 and ChLR1, a DNA helicase involved in sister chromatid cohesion, plays an important role in ensuring post-replication segregation of viral genomes (239).

Previous research has indicated that papillomavirus genomes can be replicated and maintained stably in the absence of E1 and E2 (7, 158, 336). Silencing mutations of the individual HPV open reading frames have shown that none of the individual ORFs are strictly required for successful genomic replication and maintenance in tissue culture (7). Since it has been demonstrated that replication of papillomavirus genomes occurs in the absence of viral proteins in yeast as well, it is apparent that conserved cellular factors must play a role in replacing E1 function. For instance, the E1 protein forms hexamers which function in a similar manner to cellular helicases such as Werner's (WRN) and Bloom's (BLM) Syndrome Helicases, members of the RecQ family, and the minichromosome maintenance proteins (MCM) (102, 117, 216, 242). It is conceivable, then, that host factors could be adapted to perform similar replication functions in place of viral proteins like E1, suggesting that an E1-independent mode of replication could be relevant during the maintenance phase of the HPV lifecycle.

HPVs Replication in Yeast

Having initially established that HPV genomes can replicate in *Saccaromyces cerevisiae(7, 158),* the HPV/yeast system has proven easy to manipulate to study many aspects of the HPV lifecycle, including transcription, replication and production of virus-like-particles (VLPs). HPVs 6b, 11, 16, 18, and 31 can replicate in short-term assays when transformed into competent yeast (5-7, 158, 336, 337). Furthermore, the Frazier laboratory has reported that BPV1 replicates robustly in yeast (336, 337). Recently the

Khan laboratory has reported that HPV1 can replicate in yeast, but requires a centromere to be maintained stably (51). Kim et. al. mapped both *ARS* and *CEN* replication functions in *S. cerevisiae* to the late region of HPV16 (158). The great degree of homology between the genomic replication mechanisms of yeast and higher organisms creates the possibility that similar mechanisms could be involved in papillomavirus replication in higher eukaryotes, especially during the maintenance phase when replication factor transcription is minimal.

Structure of Telomeres and their Regulation:

Telomeric repeats, telomerase and T and D loops

Telomeres are essential structures that cap and protect ends of linear chromosomes, hiding them from DNA damage sensing mechanisms and repair machinery (77) and solving the problem of sequence loss from the ends of chromosomes as a result of DNA replication through regulation of the telomerase enzyme (reviewed in (60)). The last 50-500 bases of the telomere consist of a single stranded G-rich overhang (77) sequestered into a T-loop lariat structure (79, 135), the end of which is inserted into a complementary C-rich region, resulting in displacement of the G to form a displacement or D-loop.

These protective functions are mediated by telomere associated protein activities as well as formation of T-loops by these proteins (78, 79). Telomeric DNA consists of tandem repeats of (TTAGGG)n synthesized by telomerase (60) an enzyme consisting of the telomerase reverse transcriptase and terc, the RNA template from which the repeats are synthesized (60). Telomerase is a reverse transcriptase similar to that utilized by retrotransposons, which are their potential evolutionary ancestor (101). Telomerase is conserved between vertebrates, invertebrates, plants, fungi and many unicellular

organisms (reviewed in (186). Dipterans, including Drosophila melanogaster, do not use telomerase to maintain chromosome ends, relying instead on retrotransposition (32). Certain immortalized human cell lines utilize alternate lengthening of telomeres (ALT) to replicate telomeres in a telomerase independent manner (187) utilizing telomere-telomere recombination and t-loop-mediated extension.

Telomere-related factors and their functions

The telomere is maintained through the action of a number of proteins combined together into a protein complex called shelterin. The shelterin/telosome complex functions primarily by bringing the three telomeric DNA binding factors (TRF1, TRF2 and Pot1) into the same large complex (78, 181, 237) along with Ras related protein 1 (RAP1) (173), and TRF1-TRF1 Interacting Protein (TIN2) (77). Studies suggest that the shelterin complex binds preferentially to ds/ss-DNA junctions with a Pot1 binding site and at least one Myb-domain (the DNA binding site for TRF proteins) (54). The shelterin complex binds along the length of the telomere repeats and, as the telomere length increases, negatively regulates the activation of the telomerase holoenzyme and, in doing so, regulates the length of telomeres (106) (266). Thus, the shelterin complex functions as a sort of telomerase-length thermostat, down-regulating the activity of the telomerase enzyme as the telomere increases in length, allowing more shelterin to bind.

TRF1 and 2 (telomeric repeat factors 1 and 2) bind duplex telomeric DNA (43), and are almost entirely associated with cellular chromatin (288). TRF1 and 2 share a common architecture defined by two conserved regions: a TRFH domain that mediates homodimerization and a carboxy-terminal DNA binding domain of the SANT/Myb family (43). TRF1 forms long filaments of protein bound along the length of the telomere and negatively regulate telomerase activity (133). TRF2 promotes development of T-loop

structures (134, 283), potentially as a result of positive supercoiling (4). TRF2 also serves to stabilize T-loops (116) through their N-terminal domain's ability to bind ss-DNA in a number of secondary structures. TRF2 also recruits the MRE11 complex, which functions for recombination and repair, as well as WRN and Blm helicases, more DNA repair factors (236), that could facilitate the unwinding step that is involved in t-loop formation. Loss of TRF2 leads to rapid reduction in telomere length, aberrant telomere structural formations, and activation of p53 mediated apoptosis in cells due to the proteins ability to interact with ATM (302, 309). A conditionally activated siRNA against TRF1 and 2 in mouse cells resulted in activation of the non-homologous end joining (NHEJ) system to aberrantly link the chromosomes of cells together into long chains (266).

Pot1 (protection of telomere 1) is the human homologue of the G-overhang DNA binding proteins present at the end of all telomeres, and is a structural homologue of TEBP from various protist species and CDC13 from fusion yeasts. These proteins feature a characteristic oligonucleotide-oligosaccharide binding fold within their DNA binding surfaces that provides high sequence specificity for a minimum of two single-stranded telomeric DNA repeats typically found within 3'-overhang and in D loops (23, 171, 183, 214, 294, 321). Pot1 is localized on the T-loop of the telomere through interphase and only dissociates when DNA replication is occurring, but can also be found bound along duplex regions of T2AG3 repeats away from the single-stranded loop, likely through an interaction with the TRF1/Tankyrase/Tin2 complex (310). The protein has a number of phenotypic effects on the telomeres, including being demonstrated to be both a positive and negative regulator of telomerase (9, 59, 183). Loss of Pot1 due to siRNA, deletion, or expression of a TRF2 dominant negative mutant to strip shelterin from the telomere leads to loss of telomere length, chromosomal abnormality, and eventual induction of senescence and apoptosis (24, 322). One

explanation of the dual functions with regards to telomerase may come from the protein's association with TPP1, which heterodimerizes with Pot1 and regulates recruitment of Pot1 to telomerase (160, 180, 227) and acts synergistically to recruit telomerase (308, 319). *S ceverisiae* utilize Cdc13, a similar ssDNA binding protein, to protect telomeres against exonucleolytic attack and prevents activation of DNA-damage checkpoint by chromosome ends, in place of TRF2 (189)

Additionally, these proteins transition through a number of sub-complexes that do not contain TRF1 or TRF2/RAP1 (47, 48, 145, 181, 237, 288, 327). TRF2 exists in two separate pools of protein on the telomere, one with greater stability than the other (199). Isolated chromatin has been found to contain vast molar excesses of TRF2, TIN2, and RAP1 compared to other shelterin components, implying that they may form a separate complex (288). One of these potentially is a complex dubbed "T2" by Choi et. Al. in in vitro experiments containing TRF2, Pot1, TIN2, TPP1, and RAP1 (54) which seems to be responsible for binding to the ends of telomeres and managing the T and D loops.

Replication Through Telomeres

Replication of telomeric DNA during regular cellular mitosis presents a number of challenges for cells. DNA replication typically initiates from origins located in the sub-telomeric region (311). Telomeric DNA replicates throughout S phase while subtelomeric DNA only replicates at the end of the phase, which suggests that telomeres may have their own separate origin of replication (151, 229, 290, 317). The first problem comes from typical cellular replication of the telomere, known as the end replication problem. All cells' DNA-replication machinery utilizes short RNA primers to initiate DNA replication. In the case of the telomeres, removal of the final primer from the 3' end of the linear chromosome leaves a short, unreplicated segment which cannot be filled in. Over time,

this would lead to gradually increasing loss of telomere sequence and eventually loss of coding DNA [reviewed in (132)]. The cell corrects for this by utilizing the telomerase holoenzyme for the leading strand, paired with synthesis of the lagging strand by small RNA primers utilizing polymerase primase (70, 338). In budding yeast, this occurs at the same time as G-strand synthesis and seems to regulate telomerase activity (93, 226, 312), whereas in humans the two activities are separate (338) and lagging strand synthesis is controlled by activation of cyclin dependant kinase 1 (CDK1) (70). In budding yeast, it has been shown that DNA polymerase α primase is essential for telomerase extension of telomere ends (93) to compensate for the G-overhang. Similar results have been observed in fission yeast (69) and mouse cells (222). The last segment of the telomere thus consists of one of these primers, which is then removed by nucleases, leaving a short single stranded G-overhang, which then form the T-loop.

The second problem has to do with the actual progress of replication forks through telomeres themselves. Progress of the replication loop must deal with a number of uniquely challenging secondary structures when processing through the telomeres, including G-quadruplexes, heterochromatinized DNA, and the t-loop itself. G-quadruplexes are stable intra-molecular structures which occur through the formation of Hoogsteen base-pairs between four guanine residues (142). Organisms utilize the Blm and Wrn helicases to further the replication fork migration, as well as replication protein A (RPA) and Pot1 (256, 333). Most of the telomere consists of regions of the genome which are bound by nucleosome arrays(297) with histones that have been specifically modified to consist of highly repressed structures known as heterochromatin (121, 129), which negatively regulates telomere length (29). The mechanisms by which this is regulated are not well understood, but are predicted to involve complex levels of epigenetic control. Telomeric DNA has a tendency to cause replication forks to slip backwards and generate complicated replication structures like Holliday Junctions or

chickenfoot structures which need to be resolved for successful DNA replication (115). Additionally, the t-loop itself is difficult for cells to resolve and is a site at which supercoiling stress accumulates during the course of DNA replication. Unwinding of this structure is essential to allow efficient DNA replication. TRF1 and fission yeast homologue protein Taz1 promotes efficient replication of telomeric DNA by preventing fork stalling (211, 268), but unfortunately it is not sufficient to prevent all potential replication difficulties. All of these potential replication fork obstacles can result in accumulation of stalled replication forks within the telomere, ultimately leading to activation of ATM or ATR mediated DNA damage responses (194, 304), which explains an observed activation of DNA damage signals during replication of telomeric DNA (304).

When a replication fork stalls during replication at other chromosomal sites, the stall is typically repaired by another replication fork coming from the opposite direction meeting up with the stalled fork, at which point the two strands are joined through recombination. This, of course, is not an option for telomeric sites where replication is unidirectional. Stalls in telomere replication forks initiate a DNA damage response that recruits the MRE11/Rad50/NBS1 complex to the site, because it is detected as a double strand break (304). This complex then activates the Ataxia Telangiectasia Mutant (ATM) and Ataxia Telangiectasia and Rad50 Related (ATR) damage responses. RPA protein binds along the length of any exposed single stranded DNA. MRN, the 911 complex, FEN1, DNA polymerase β, and Rad17 then cooperate to restart the replication fork (reviewed in (304).) Replication through telomeres is thus observed to typically be a two step process as observed by rate of BrdU incroporation: 1) Replication progresses into the telomeric region and stalls and 2) replication fork reinitiates and progresses to the end (304).

RecQ helicases are a conserved family in yeasts and mammalian cells that are essential for maintaining genome integrity. These are critical for telomere replication and resolution of telomeric recombination (10, 14, 66). Patients with defective Wrn helicase specifically lose telomeres replicated by lagging strand synthesis, presumably due to Wrn's ability to unwind G-quadruplexes in a POT1 dependent manner (174, 234). Blm helicase similarly functions to resolve these quadruplexes (234), and mutations to this gene increase levels of sister-chromatid exchange, genomic instability, fragile telomeres, and elevated levels of chromosomal aberrations (122, 123, 268). TRF2 seems to play a role in stimulating Wrn/Blm activity (234). A heterotrimeric complex of proteins known as RPA functions to bind to ssDNA during DNA replication and repair processes (316, 343). Loss of RPA (or its homologues) in budding and fission leads to gradual shortening of the telomeres (233, 276). In vitro models have shown that RPA is capable of stimulating WRN's ability to resolve G quadruplexes, modulates telomerase activity (235, 256, 280), and coats ssDNA during the passage of the replication fork. However, Pot1 has a higher affinity for the G-rich regions of the telomere single-strand overhang (10). An exciting field of research is emerging to study a newly discovered RPA-like heterotrimeric complex, the CST complex. This appears to be a key player in regulating C-strand synthesis (215, 286). In both yeast and human cells, the CST binds to single-stranded DNA and plays a role in mediating C-strand fill-in, regulates Telomerase (both positively and negatively), and prevents excessive G-strand elongation (49, 215).

The role of TRF2 in aiding telomeric DNA replication is essentially through acting as a protein hub (124, 157) and recruiting important replication and DNA maintenance factors to critical sites. For instance, TRF2 recruits proteins like Apollo and Top2α to locations of supercoiling strain to release tension during replication (4, 326). Additionally, TRF2 localizes to sites of recombination and T-loop formation, where it then utilizes factors like BLM and WRN helicase along with the MRE11/Rad50/NBS1 to resolve the

structures efficiently (4). Additionally, TRF2 localizes to the sites of the pre-replication complex (pre-RC), where it interacts with the origin recognition complex (ORC) and facilitates initiation of telomere replication (86, 292).

Telomere maintenance and Segregation

Telomeres play a vital role in ensuring the proper segregation and maintenance of chromosomes during cell division, specifically meiosis. Significant research has been conducted investigating telomere maintenance and meiosis of yeast. At the onset of meiotic prophase I, telomeres attach to the nuclear envelope (NE) and undergo NE-bound motility, attaching to the nucleoplasmic face of the inner nuclear membrane (46). It has been proposed that the telomeres potentially connect to filament bundles that project between the telomere attachment plate through the NE to the cytoplasm (178) The telomeres then move along the inner nuclear membrane to the cytoplasmic microtubule organizing center (MTOC) in animals and fungi (342) or the cell cortex in plants (65)

Disruption of telomere maintenance proteins in eukaryotic cells lead to a number of dysfunctions. *Saccaromyces cerevisiae* meiosis relies on scRap1. Strains with altered telomere sequence or scRap1-binding sites experience defective meiosis (2, 193) Terc -/- mice have reduced telomeric repeat tracks and defective axial element (AE) formation and synapsis (182). Mammalian ring chromosomes, which don't contain telomeres, don't localize to the nuclear periphery of spermatocytes (305).

Information on the proteins required for mediating the attachment to the nuclear envelope is somewhat scarce, but the process seems to rely on the SUN domain proteins (Sad1p, UNC-84 proteins that are important for positioning of the nucleus) which are present in most organisms. SUN domain proteins bridge the gap between the inner and

outer nuclear lamina, connecting the nuclear and cellular cytoskeletons and playing a role in organizing nuclear contents (300) SUN1 (96) and potentially SUN2, given that the two typically have associated protein functions in NE protein attachment and coordination (138).

Unfortunately, information on the role which telomeric proteins play during mitotic chromosomal segregation is incomplete outside of some cursory studies. While it is clear that telomeres play a role in ensuring the success of this process and regulating telomeric recombination during meiotic pairing, a hard mechanism has yet to be defined. TRF1 associated protein Tankyrase 1 (TANK1) is essential for separation of sister chromatids during mitosis (99). Male mice missing the A-type lamin isoform C2, an important protein linking the actin and microtubule skeleton to the nuclear envelope, fail to undergo fine meiotic telomere/NE attachment and clustering (3). Nuclear envelopes of frog oocytes contain a TRF2 homologue (247) that may provide a possible means by which they attach to the envelope during meiosis. Rap1 is dispensable for formation of the “bouquet” structure characteristic of chromosomal organization during cell division in mammalian cells (260) but is required for the same process in Schizosachromyces pombe (53, 154)

Homologous Recombination Based DNA Repair

Five distinct complexes exist to monitor for and repair DNA damage, of which the ATM and ATR system are best characterized. Within the ATM system, DNA damage is detected by the MRN complex, consisting of MRE11, Rad50, and NBS1. Once a break is identified, the complex activates ATM and recruits it to the damage site (167). ATM then phosphorylates a number of down-stream effectors to initiate the repair processes. Over 700 proteins are phosphorylated in response to Ionizing Radiation damage (198).

Double strand breaks in DNA, particularly those caused by the introduction of ionizing radiation, are repaired primarily by one of two mechanisms: the non-homologous end joining (NHEJ) system and Homologous Recombination Repair (HRR) (42, 64, 289). HRR is a primordial mechanism involving a complex series of events including: end resection, Rad51 filament formation, homologous sequence identification, heteroduplex formation, repair synthesis, and heteroduplex resolution [reviewed in (125)]. This repair mechanism can only function during S and G2 phase, when DNA has been replicated and the sister chromatid is available for recombination. Approximately 15% of ionizing-radiation induced DSBs are repaired by HRR (153).

Regulation of the HRR system is complex and redundant. The identification of a DSB in a chromosome is followed immediately by phosphorylation of ATM (17). This phosphorylation is required for recruitment of ATM to the site of the break and is 90% efficient, but is not required for successful repair (315). ATM is thus alternately described as either a director sensor of DSBs and mediator of repair or an indirect sensor that promotes cell survival and repair. ATM then functions to phosphorylate a number of downstream factors at the repair site to activate the repair processes including Chk2, BRCA1, and γH2AX(56, 139). ATM activation is also required for differentiation dependant amplification (218).

The primary indicator of DNA damage in cells is the phosphorylated forms of γ-phosphorylated Histone 2AX (γH2AX) (45, 108, 284). Another factor phosporylated by ATM is MDC1 (mediator of DNA-damage checkpoint 1), which recruits a number of HRR factors including 53BP1 and BRCA1 (113, 228). Another key factor is the NMD complex of proteins, consisting of MRE11, Rad50, and MRN. MRE11 is a single ssDNA endonuclease and possesses 3'-5' endonculease activity (299). Rad50 is an ATPase (148). MRN contains a DNA end binding domain. Together, these proteins rapidly locate

sites of DNA damage independent of ATM activation and helps recruit factors to initiate the repair processes (212). Additionally, the MRN complex actually promotes recruitment of ATM to damage sites and increases its autophosphorylation (108, 167, 168).

DNA Repair Systems Targeted To Telomeres

The interplay between the telomere maintenance proteins and the cellular DNA repair systems is intricate and vital, since loss of these systems leads to rapid activation of NHEJ, telomere loss and/or end-to-end fusion of chromosomes. Formation of the T-Loop at the telomeric end is an important part of this process, as it inhibits the activity of DNA liagase IV, the enzyme responsible for fusing telomeres (77). TRF2 is obviously vital for this, as it is responsible for forming the T and D-loop structures along with POT1. TRF2 is thus required for blocking recognition of telomeres as ds-DNA breaks.

Interestingly, a number of DNA repair factors localize to the telomeres through a direct interaction with TRF2 (278) and a body of evidence is available to demonstrate that the telomere maintenance systems in general and TRF2 in particular play an important positive role in DNA repair. Conditional deletion of TRF1 and 2 to remove the shelterin complex leads to derepression of 6 different repair pathways: ATM, ATR, C-NHEJ, HDR, alt-NHEJ, and 5' resection (267). The basic domain of TRF2 allows the protein to localize to double stranded breaks induced by UV light (40) irrespective of the presence of TTAGGG repeats, as well as Holliday junctions (116). This makes some sense, given that TRF2 is important for initiating a telomerase independent method of telomere maintenance known as ALT which relies heavily on recombination based methods to regulate telomere length (281). This ALT system relies on TRF2 mediated formation of ALT Mediated PML Bodies (APBs) that contain MRE11, NBS1, Rad50, and

PML components. A form of this can be observed in vitro, as it has been found that telomere repeats can recruit PML through a SUMO dependant mechanism and initiate non-replicative DNA synthesis (55).

Indeed, TRF2 appears to play a vital, but as-yet unclear role in activating DNA repair systems. In response to DNA damage such as double strand breaks, PI-3-Kinases like ATM rapidly phosphorylate Thr-188 of TRF2, triggering it's relocation from telomeres to the sites of DNA damage (40, 150, 291). Phosphorylated TRF2 leads to increased survivability of cells after X-Ray induced DNA damage, apparently by initiating the Fast DNA repair response and altering the kinetics of H2AX phosphorylation (150). Contradictorily, TRF2 inhibits ATM activation (155) possibly suggesting a feedback mechanism to regulate ATM activation and TRF2 phosphorylation. TRF2 has been shown to inhibit NHEJ and upregulate the HRR system (195). As such, loss of TRF2 leads to activation of ATM kinase, the p53 tumor suppressor signaling pathway, and the MRE11 double-strand break repair system. This leads to induction of cellular senescence or apoptosis (47, 84, 107), activation of NHEJ, interchromosomal fusions, (302) and can lead to telomeres being fused to other double-stranded breaks at non-telomeric sites. This can ultimately lead to chromosomal translocations (130). Loss of Pot1, additionally, leads to activation of ATR kinase, formation of telomere dysfunction-induced foci, and induction of apoptosis or cell cycle arrest (84, 146). Sudden large deletions of over-elongated yeast telomeres have been observed as a result of resolution of the t-loop as if it were a Holliday Junction, underlining the need for strict regulation of the extent to which t-loops form (190).

Telomere Replication and Maintenance with viruses

Telomere maintenance and DNA repair proteins play an important role in the life-cycle of numerous DNA viruses that utilize a long-term stability strategy. Many DNA viruses utilize similar replication and maintenance strategies to HPV. Epstein Barr Virus (EBV), Kaposi's Sarcoma Herpesvirus (KSHV), Herpesvirus saimiri (HVS), and gamma-herpesvirus68 (MHV-68) all have a chromosome tethering strategy (reviewed in (201).) In addition, a number of other interactions between DNA viruses and the DNA repair systems have been demonstrated. HSV 1, EBV, and SV40 recruit Rad51 to replication compartments (35, 165, 314).

Initially, an interaction between the Epstein Barr virus and the telomere maintenance system was demonstrated by the laboratory of Dr. Paul Lieberman(325). One region of the EBV genome is the plasmid origin of replication, the OriP, which consists of the dyad symmetry region (DS) and the family of repeats (FR). Plasmids which contain the OriP alone are capable of stable replication and maintenance of plasmids in mammalian cells in the presence of the EBV Nuclear Antigen 1 (EBNA1) , a factor which is required for viral genome maintenance during latency (325). EBNA1 as well as KSHV LANA protein are structural and functional homologues of HPV E2, possessing a unique anti-beta barrel DNA binding structure (34, 136). The DS to contains 4 sites of nine-base telomeric repeat DNA in the dyad symmetry region, similar to those found in the late region of HPV. These repeats are required for DNA replication (225) and plasmid maintenance (225, 325). These telomeric repeats also allow binding of TRF2 (87), which is required for plasmid replication and maintenance (87, 323). It's believed that one important function of TRF2's role in OriP replication is recruiting the proteins of the Origin Recognition Complex (ORC). TRF2 interacts with EBNA1 (169),

and this interaction is required for the replication and maintenance effect (85, 87). These processes can be inhibited by overexpression of a TRF2 dominant negative plasmid with the N and C terminal domains deleted (85). Another important role of TRF2 is to recruit the Origin Recognition Complex (ORC) to the DS (13). Other studies demonstrated that the MRN repair complex, which consists of MRE11, Rad50 and NBS1, interact with the OriP plasmids in a TRF2-and-cell-cycle-mediated manner, and this interaction has an effect on OriP plasmid stability (92). Loss of NBS1 or MRE11 leads to loss of episomal maintenance of EBV genomes in certain lymphocyte cell lines (92). Replication at the OriP was shown to involve formation of recombination-like structures similar to Holiday Junctions during S-phase by 2D Gel electrophoresis.

KSHV also encodes a structural and functional homologue of the HPV E2 and EBV EBNA1 proteins, Latency Associated Nuclear Antigen (LANA.) LANA is responsible for latent phase DNA replication, gene expression, and segregation, again similar to the E2 and EBNA1 proteins (20). A study performed by the Renne laboratory demonstrated that, along with thirty other proteins, TRF2 associates with a seventy base pair minimal replication element from the KSHV genome (149). Additionally, a number of DNA repair proteins, including Ku70, PARP-1, and DNA-PK associates with this minimal repeat element (270). LANA co-immunoprecipitates with TRF2, but interestingly only when cotransfected with a plasmid containing an intact, wild type copy of the RE. Unlike results observed in EBV, no nanomer telomere repeats were found in the KSHV RE, and co-transfection of the RE plasmid with the TRF2 dominant negative mutant did not result in a significant change in plasmid replication compared to negative control. The study did not, however, examine what effect loss of TRF2 function would have on plasmid segregation, or if siRNA knockdown of TRF2, rather than overexpression of the dominant negative, would have an effect on plasmid replication.

Some telomeric factors are associated with papillomaviruses. The most obvious example of this is the activation of telomerase expression in cells due to E6 expression (163, 303). One possible explanation of this is that hTert can substitute for E6 in E6/E7 mediated immortalization of human foreskin keratinocyte cells (162), indicating that telomerase activation is an important component of HPV maintenance. ChLR1, a DNA helicase involved in sister chromatid cohesion, is an E2 interaction partner that has been shown to be required for extrachromosomal maintenance of BPV1 genomes (239).

HPV and HRR

Papillomaviruses, like many other low-copy number DNA tumor viruses interact with and utilize components of the DNA repair systems, particularly the homologous recombination repair system. HPV31 was shown by Liu Laimins' laboratory to activate ATM during the course of its replication within cells, partially due to expression of E7 (218, 252, 314). E7 binds to ATM and the MRN complex. ATM activation is also required for genomic amplification but not stable maintenance replication (218). A number of studies have demonstrated that several downstream ATM DNA repair factors are recruited to HPV replication foci (125, 218, 251, 255, 287). HPV31 replication in keratinocytes increases γH2AX and 53BP1 levels and leads to recruitment of these proteins to replication foci. γH2AX specifically binds to the URR of HPV31. A separate study has demonstrated that a similar increase in γH2AX occurs with expression of E1 from HPV18, 16, 11, and 6B, indicating this is a broad feature of HPV replication, and that co-expression of E1 and E2 activates DDR (251). Comet assays have indicated that E1 and E1+E2 complexes induce double stranded DNA breaks. The ATPase and DNA melting functions of E1 are required for this process, while the sequence specific DNA binding function of E1 and E2 are dispensable, suggesting an alternate means of

initiating the DNA damage and implying that E2's primary function in this role is in stabilizing E1. Interestingly, inclusion of an HPV origin-containing plasmid reduced the DNA damaging activity of E1, suggesting that the presence of an HPV origin may alter the E1 protein's function.

pATM, BRACA1, RPA, Rad51, ATRIP, TopBP1, and Chk2 are also recruited to HPV replication sites, and the amount of these factors increases as differentiation progresses in the cells (125, 251). As an increase in differentiation leads to an increase in replication of the HPV genomes, it has been theorized that the DNA repair factors are recruited to the newly synthesized copies of HPV within these foci. This is supported by the gradual increase in levels of activated RPA, which binds to single stranded or newly synthesized DNA. Interestingly, however, activation of ATM and Chk2 signaling appear to be dispensable for transient as well as maintenance replication (218, 251), leaving the question of just what role these processes play in HPV DNA replication.

Specific Aims

The objective of my research was to examine the relationship between human papillomaviruses and their host cells during the maintenance phase of their viral life cycle. During this phase, the virus only expresses E6 and E7 to an appreciable level in infected cells and is highly reliant on the host cell to provide the factors required for viral genome replication and maintenance. DNA viruses which rely on a long-term replication strategy in cells traditionally co-opt some or all elements of the DNA repair and, discovered more recently, the telomere maintenance systems to provide these functions. We thereby utilized bioinformatic, molecular, and cellular methods to investigate the interaction between papillomaviruses and their host cells, particularly through the E2

maintenance protein and its interactions with the telomere maintenance and DNA repair systems, with an aim to better understand the factors which provide long-term stability for HPVs.

Specifically, we utilized bioinformatic analysis to investigate the evolution of the E2 protein and binding sites as papillomaviruses evolved to infect a wider range of host-species and tissue types.

Secondly, through our well-established yeast replication system, we performed experiments to determine which papillomaviruses share the ability to replicate in *Saccharomyces cerevisiae* and to potentially shed light on what common sequence features are important. Through this study, we determined that a number of Papillmavirus species are not capable of replicating stably in yeast, particularly BPV1, which had previously been identified as being capable of replicating short term.

Finally, we investigated the interaction between TRF2 and HPV16 . Upon identifying the presence of nine base telomere repeat sequences in the late region of a number of HPVs, we sought to determine if these were necessary for HPV maintenance. We utilized a Chromosome ImmunoPrecipitation (ChIP) assay to show that these sequences were bound by telomeric and some DNA repair proteins in vitro. We showed that in both yeast and mammalian cells, deletion of these binding sites led to an overall loss in stability ranging from slight shifts in copy number to complete loss of plasmid viability. Through a number of methods, we demonstrated that TRF2 interacts with the viral E2 protein, as do a number of other telomere maintenance proteins. Finally, we took steps to demonstrate the co-localization of the proteins within cells.

Chapter 2

Evolutionary variation of papillomavirus E2 protein and E2 binding sites

Evolutionary Variation of Papillomavirus E2 Protein and E2 Binding Sites

Adam Rogers, Mackenzie Waltke, and Peter C. Angeletti*

Nebraska center for Virology, School of Biological Sciences,

University of Nebraska-Lincoln, Lincoln, NE, 68583-0900, USA

Running title: **Variation of Papillomavirus E2 Protein**

* Corresponding Author Mailing Address: Nebraska Center for Virology, 237 Morrison Center, University of Nebraska-Lincoln, 4240 Fair Street, Lincoln, NE 68583-0900

Tel. 402-472-3986, Fax. 402-472-3323, E-mail: Pangeletti2@unl.edu

KEYWORDS: extrachromosomal DNA, persistent infection, Human papillomavirus, E2 Protein,

DNA binding Domain

Published in: Rogers A, Waltke M, and Angeletti PC. "Evolutionary Variation of Papillomavirus E2 protein and E2 Binding Sites." **Virology Journal.** 8, 379. Aug 1 2011.

Abstract

In an effort to identify the evolutionary changes relevant to E2 function, within and between papillomavirus genera, we evaluated the E2 binding sites (E2BS)s inside the long-control-region (LCR), and throughout the genomes. We identified E2BSs in the six largest genera of papillomaviruses: Alpha, Beta, Gamma, Delta, Lambda, and Xi-papillomaviruses (128 genomes), by comparing the sequences with a model consensus we created from known functional E2BSs (HPV16, HPV18, BPV1). We analyzed the sequence conservation and nucleotide content of the 4-nucleotide spacer within E2BSs. We determined that there is a statistically significant difference in GC content of the four-nucleotide E2BS spacer, between Alpha and Deltapapillomaviruses, as compared to each of the other groups. Additionally, we performed multiple alignments of E2 protein sequences using members of each genus in order to identify evolutionary changes within the E2 protein.

When a phylogenetic tree was generated from E2 amino acid sequences, it was discovered that the alpha-papillomavirus genus segregates into two distinct subgroups (α1 and α2). When these subgroups were individually analyzed, it was determined that the subgroup α1 consensus E2BS favored a spacer of AAAA, whereas subgroup α2 favored the opposite orientation of the same spacer; TTTT. This observation suggests that these conserved inverted linkers could have functional importance.

Introduction

The E2 protein serves many functions with its primary role being that of a transcriptional regulator. E2 is a negative regulator for the expression of the oncogenes E6 and E7 (184). E2 polypeptides contain three probable domains: a DNA binding domain (DBD) located at the C-terminus, an N-terminus transactivation domain, and an internal "hinge" domain. Both the C-terminal and N-terminal domains are relatively well conserved within the PVs (254). E2 binds as a dimer at DNA-binding sites and this action is mediated through the DBD (202). The E2 DBD forms a dimeric β-barrel and each strand contributes a half-barrel. The dimer interface has a hydrophobic core and uses extensive hydrogen bonding between subunits to maintain tight binding. This β-barrel core contains elaborately packed side chains that contribute to the stability of the dimer. There is a poorly conserved loop connecting β-strands 2 and 3. This loop varies from 6-10 residues. The tertiary structure of characterized E2 DBDs is similar, but there appears to be variation in the orientations of the two subunits (141). There is some evidence to suggest that the activation domain mediated oligomerization could influence interaction between E2 molecules bound at distant E2-binding sites forming DNA loops and other DNA structures (8, 141).

The consensus sequence 5'-ACCgNNNNcGGT-3' is recognized by E2, with the position 4 and 9 residues allowing some variability. A number of studies have been performed to examine the binding of E2 protein to its cognate binding site (33, 83, 111, 141, 144, 249, 257, 293). The 4-nucleotide spacer sequence varies by HPV type, and has been identified as being critical for determining E2 binding affinity through indirect readout as well as playing a potential role in gene regulation, despite having no predicted nucleotide-amino acid contacts from crystal structure (33, 83, 141, 144, 328).

E2 binds DNA as a homodimer with each monomer supplying an alpha helix to contact two successive major grooves of the target site (83, 141)

Four typical E2 binding sites are conserved in the upstream regulatory region (URR) of most papillomaviruses numbered according to their distance from the early promoter (202). Each site is differentially regulated and demonstrates variable binding affinity for the E2 protein, resulting in varying replication and transcriptional effects during the viral life cycle (63, 191) presumably as a result of differences in E2 binding affinity (141) due to sequence variation as well as methylation of the E2 binding site (257, 293). These binding sites are typically well conserved across all papillomaviruses. However, in some cases variation in the number and location of some E2 binding sites does exist, including a predicted fifth binding site within the URR of betapapillomaviruses (103) and some alphapapillomaviruses (257) as well as observation of up to 17 sequences with ability to bind E2 from the URR of bovine papillomavirus 1 (249).

In this study, we examined the evolutionary divergence in E2BS recognition by the E2 transcriptional regulatory protein. Currently, the majority of the work performed on the E2 protein function has been performed on domains from a relatively small number of papillomavirus types. To better understand the binding properties of E2 from a wide spectrum of HPV strains, we performed an observational study in which we used bioinformatic tools to generate a list of putative E2BS sequences matching the consensus in all papillomaviruses currently classified by ICTV and analyzed them for variations in binding site number, location, and differences in the 4-nucleotide spacer region between the largest of the HPV genera, the Alpha, Beta, Gamma, Delta, Lambda, and Xipapillomaviruses. We then performed multiple alignment and phylogenetic analysis using the E2 amino acid sequences of these viruses to observe evolutionary patterns from an E2-Centric perspective. Finally, we performed sequence alignment of

the viral E2 protein C-terminal DNA binding domains of each genus and observed that a greater degree of variation is present in the Alphapapillomaviruses compared to Beta. As one of the characteristics associated with the classification of papillomaviruses into their respective genera includes the ability to infect mucosal and cutaneous epithelia as well as fibroblast tissue, we propose that evolution of the E2 protein and its cognate binding site correlate with the adaptive radiation papillomaviruses underwent during the course of evolving to infect new tissue types.

Materials and Methods

Putative E2 Binding Site Identification and Analysis

Initially, we obtained sequences for the confirmed E2 binding sites from three representative, well-characterized papillomavirus species, HPV16, HPV18, and BPV1(141, 176), to create a broad, complete representative training data set. We then utilized Multiple EM Motif Elicitation (MEME) software to use statistic modeling techniques to create a consensus motif sequence for E2 binding sites within the genomes of papillomaviruses(15). This motif was then used to search through all complete papillomavirus sequences (obtained from the Papillomavirus Episteme (PaVE) database containing information from Refseq and Genbank (1, 232, 248)) for all papillomavirus genera containing 5 or more members (HPV 2-40, 42-45, 47-62, 65-78, 80-96, 99, 100, 102, 104-107, 110, 111, FA75/KI88-03, RTRX7, BPV1-9, COPV, DPV, FdPV1, FdPV2, LrPV1, PlpPV1, PcPV1, UuPV1, and MfPV1-10, utilizing the Motif Alignment and Search Tool (MAST)(16). For later phylogenetic analysis of alphapapillomavirus subgroups, we divided our data set to into high and low risk groups and alphaPVs capable of infecting cutaneous keratinocytes. The high risk group

included HPV 16, 18, 26, 31, 33, 35, 39, 45, 52, 56, 58, 59, 67, 73, and 82. The cutaneous subgroup included HPV2, 3, 10, 27, 28, 29, 57, 78, and 94.

E2BS Sequence Analysis

After retrieving the list of putative E2BSs from the ICTVdb papillomavirus sequences, the data was sorted based on a number of criteria. Recovered sequences were manually analyzed from the resultant MAST output to observe the genome location of the identified binding sites as well as the GC content of the four base spacer sequences. Binding sites were classified as either within or outside the LCR according to the criteria of being located between the end of the L1 opening reading frame and the beginning of the E7 open reading frame. Binding sites were similarly separated into their respective papillomavirus genera and the identified E2BSs were analyzed using MEME to generate a SequenceLogo to observe the consensus E2BS sequences for each papillomavirus genus. Similar MEME analysis was performed to compare the E2BSs of low and high risk alphapapillomaviruses.

Protein Sequence Alignment

Amino acid sequences for all known E2 proteins within the papillomaviridae family were acquired from NCBI and sorted into the respective papillomavirus genera analyzed in 2.1 and 2.2. To increase the significance of results, analysis was limited to the alpha and betapapillomavirus genera, as the other genera possess less than ten members each. All E2 sequences were then aligned using Muscle (100). Some sequences (HPV77, 3, and 29) were removed due to long stretches of non-homologous repetitive DNA in the linker region. Alignments were then repeated, focusing specifically on aligning the amino acids located within the C-terminal DNA binding domain of E2.

Weblogo was then used to generate a graphical representation of the sequence alignments.

Phylogenetic Analysis

We performed phylogenetic analysis to examine evolution of papillomavirus E2 amino acid sequences. Complete amino acid sequences were obtained from NCBI for all papillomaviruses analyzed in 2.1 and subjected to multiple alignment using COBALT software (238). The multiple alignment was then used to draw phylogenetic trees using Neighbor Joining and Kimura protocols.

Results

E2BS Identification

To examine the evolution of the E2 DNA binding site (E2BS) sequence, we utilized the sequence motif analysis software Multiple EM Motif Elicitation (MEME) to generate a consensus DNA binding site. Initially, we generated a training set based on the confirmed E2 binding sites from HPV16 and 18 as well as BPV1, as these are well characterized and representative of the papillomavirus family. The resulting binding site motif Sequence Logo is shown in panel 2.1A, demonstrating the high conservation of bases from positions 1-3 and 10-12. As expected, little sequence conservation from the four base spacer region was observed. Genome sequences were collected from ICTVdb (248) and sorted into the various papillomavirus genera. Papillomavirus genera were eliminated from the rest of the analysis if they contained less than five members to improve the statistical significance of results. In total, 68 alpha, 35 beta, 6 delta, 7 gamma, 7 lambda, and 5 xipapillomaviruses were analyzed, totaling 128 papillomaviruses (111 human and 17 animal sequences). These were then used to

identify the location of E2 binding sites utilizing MAST software to identify DNA sequences with high sequence identity to the MEME generated binding site motif.

As expected, the four conserved binding sites located within the URR were identified in the majority of papillomavirus species examined (data not shown.) However, a number of additional potential E2 binding sites were identified within and outside of the URR. The number of E2 binding sites identified averaged between four and six per genome for the alpha, beta, gamma, lambda, and xipapillomaviruses, whereas the delta papillomaviruses averaged eight binding sites per genome, (Fig 2.1B) due in large part to the 14 E2BSs identified in BPV1. The majority of these sequences were located within the URR as expected, averaging approximately 3 for the alpha, beta, gamma, lambda, or xi, and 7 for delta.

E2BS Sequence Analysis

The identified E2 binding sites were then collected and examined to identify the GC content of nucleotides located within their four base spacer regions. G and C nucleotides from the observed E2BSs were counted and tabulated to obtain the average GC content of the four nucleotide spacer. Most cutaneous papillomavirus genera contained approximately 25 to 30% GC content within the spacer region (Fig. 2.1C). Alphapapillomaviruses in general tended to have very low GC content (15%) and deltapapillomaviruses tended to be very high (approximately 50%, indicating no statistical preference for GC versus AT bases.)

When E2BSs were sorted into “within the URR” and “outside the URR” groups, certain trends became apparent. First, alphapapillomaviruses and to a lesser extent xipapillomaviruses seemed to have a unique requirement for AT nucleotide rich spacers within the URR and a much higher GC content in E2BSs located outside. Gamma and

lambdapapillomaviruses seemed to possess an opposite trend, with a 15-18% GC content outside the URR and significantly higher found inside. Deltapapillomaviruses tended to have a much higher GC content within the spacer than the other papillomavirus genera, while the betapapillomaviruses was approximately 30% GC rich.

To further this analysis, we performed MEME analysis on the identified E2BSs for each papillomavirus genus to identify sequence variation binding sites by genus (Fig. 2.1D). As expected, nucleotides 1-3 and 10-12 were well conserved across papillomavirus genera. Some variation was observed in the preference for C and G nucleotides at positions 4 and 9 respectively, particularly in the gamma and delta genera at position 9. The four nucleotide spacer is, as expected, highly variable between papillomavirus genera, however some trends are apparent. Alpha papillomaviruses seemed to have the most consistent sequence conservation, particularly at positions 5-7 where A nucleotides were highly conserved. A and T bases were overrepresented in all papillomavirus genera except deltapapillomaviruses. Despite little evidence of evolution of contact nucleotides, we observed that each of the papillomavirus genera seem to have varying preferences for binding site spacer sequences.

E2 Protein Phylogenetic Analysis

To examine E2 evolution from a protein perspective, we acquired amino acid sequences for all the E2 proteins from papillomaviruses used for the E2BS analysis. The E2 sequences were then analyzed using COBALT software under Neighbor Joining and Kimura protocols. The resultant phylogenetic tree is shown in Figure 2.2A. As shown, when analyzed simply from E2 amino acid sequences, papillomaviruses sort into specific clades matching with the official genera classifications which, as stated previously, were based on L1 amino acid sequences (81)

Three specific clades become apparent based on this analysis: one containing the deltapapillomaviruses, one containing the alphapapillomaviruses, and a third encompassing the other genera analyzed. The delta clade possessed a large degree of evolutionary distance compared to the other clades from the COBALT analysis, implying a significant evolutionary divergence in terms of the E2 protein from the other papillomaviruses. One papillomavirus, FDPV2, did not sort out with the other members of the lambdapapillomavirus genus and did not associate with any of the other clades identified by this analysis.

The alpha clade further subdivides into two subclades, in this study labeled as α1 and α2. When analyzed independently, specific trends were identified for these two subclades. The individual members of the subclades possess specific infectious characteristics (Fig 2.2b). The majority of the human papillomaviruses from subclade α1 are considered to be high risk for progression to cervical cancer. While they do not cluster together particularly well within the subclade, one subclade contains both HPV16 and HPV31, two of the papillomaviruses that are most associated with cervical cancer. Interestingly, subclade α1 also contains a cluster of viruses infecting longtailed and rhesus macaques, which seems to have diverged less than the other members of the subclade (Fig 2.2A). Subclade α2 contains two clusters of alphapapillomaviruses capable of infecting cutaneous keratinocyte cells as well as three clusters associated with large genital warts.

MEME Analysis of Alpha Subclade E2BSs

Given the results of the phylogenic analysis for the alphapapillomavirus genus, we performed MEME analysis on the identified E2BSs for each of the alphapapillomavirus subclades, as well as those papillomaviruses classified as high and

low risk of progression to cancer and the two clusters containing the alphapapillomaviruses capable of infecting cutaneous keratinocytes (Fig. 2.2C.) Given that subclade 1 primarily contains high risk viruses, the consensus motif for alpha subclade1 and high risk alphapapillomaviruses are essentially identical. No significant difference was apparent between the high risk and low risk viruses outside of a slight under-representation of the guanine nucleotide at position 4 which could, potentially, suggest reduced susceptibility at this site for methylation (see discussion). Cutaneous papillomaviruses possessed a significantly reduced preference for A/T nucleotides within the four base spacer. Interestingly, the subclade 2 consensus motif appears to contain a preference for bases within the four base spacer of thymine rather than adenine. Given that the E2BS sequence is a pseudo-palindrome, this implied that the consensus motif for clade 2 is an inversion of the motif from clade 1.

E2 Amino Acid Sequence Conservation

One of the primary differences between alpha papillomaviruses and the other genera is the ability to infect mucosal versus cutaneous keratinocytes. Consequently, we examined whether a similar level of divergence existed in the amino acid sequence of the protein itself. In order to demonstrate evolutionary divergence of human papillomavirus E2 proteins, we first obtained complete amino acid sequences for all the Alphapapillomaviruses and a representative genus of cutaneous papillomaviruses, the Betapapillomaviruses. Other papillomavirus genera were excluded, as these respective groups averaged less than ten members, thus the alignments could not be considered reliable. We initially performed sequence alignments on the entire protein. However, it was determined that the linker region of Alphapapillomavirus sequences, which is not well conserved within the HPV respective types, was skewing the results of the alignments (data not shown). We therefore adjusted our sequences to contain only the

C-terminal 80 amino acids of the E2 protein, which roughly corresponded to the DNA binding domains (DBD) (Figures 2.3A, 2.3B). It is immediately apparent that Alphapapillomaviruses seem to have a great degree of sequence diversity relative to Betapapillomaviruses. A series of representative alignments obtained an average sequence identity of 41% for Alphapapillomaviruses when compared to 65.25% identity for Betapapillomaviruses. The differences were also apparent when the logos representative alignment program was used to generate a consensus sequence (Figures 2.4A, 2.4B) even within the well-conserved region of amino acid sequence that makes direct contact with the nucleotides of the E2BS.

Discussion

The vast majority of papillomaviruses analyzed using MEME and MAST during the course of this study conform to the expected number and location of the four conserved E2BSs within the URR of their genome, with some variations in individual strains. The averages across all the genera were between 4-6 E2BSs, outside of the Deltapapillomavirus genus which was contained significantly more. The majority of the sites identified from the study were located within the URR, as expected, though in some cases sequences that have the potential to be bound by E2 protein were identified within the papillomavirus coding sequences. Whether these putative downstream E2BSs are actually occupied during active infection is an open question but could, presumably, significantly impair the expression of the ORF the binding site is located within by blocking the progress of RNA polymerase.

Papillomaviruses are classified by their tissue tropism, genome organization, and sequence divergence in a conserved region of the L1 open reading frame (82). However, recent phylogenetic analyses have demonstrated that alignment based off of the E1 and E2 protein sequences results in a phylogeny which better clusters

papillomavirus species in terms of their epidemiology and oncogenicity (176). The E2 protein is one of four genes which are present in all known papillomaviruses, but has the highest DN/DS ratio of the four, or ratio between nonsynonymous versus synonymous substitutions (223). This is not surprising because E2 plays numerous functional roles in the cell. For example, E2 regulates transcription, facilitates DNA replication, and regulates viral genome maintenance (202). It is logical, then, that E2 would be under significant evolutionary pressure to adapt to new cellular environments.

E2 proteins bind a consensus palindromic sequence ACCgNNNNcGGT through a dynamic, water mediated interface (111, 141). The NNNN central region or "spacer" is conserved in length, but the sequence varies by species and individual binding site position. Hierarchical occupation of the E2BS by the protein has important functional and regulatory consequences for both transcription and replication during infection. Previous studies have shown that A:T-rich spacers have an increased binding affinity in certain papillomavirus species (141, 202). Specifically, while some Alphapapillomaviruses like HPV16 are acutely sensitive to AT concentration in the spacer region, others like BPV1 are essentially insensitive. Hegde et. al. proposed that the reason for this is due to a reduced ability possessed by the E2 protein of some Alphapapillomaviruses, specifically HPV16, to bend DNA into a conformation which fits within the E2 DNA binding pocket (141). Essentially, AT rich nucleotide motifs are intrinsically rigid and pre-bent into a shape that conforms to the E2 protein DNA binding domain. The results of this study support the concept that Alphapapillomavirus E2BSs possess approximately 95% A/T nucleotides within the spacer region, but only 75% in the cutaneous papillomavirus genera, and 50% in Deltapapillomaviruses. With the current limited understanding of nucleotide sequence recognition, specifically for indirect readout which occurs in regions like the E2BS spacer where no direct nucleotide-amino acid contacts are made,

predictions of binding affinity are limited to sophisticated bioinformatic modeling software and empirical data identified using methods like quantitative EMSA. However, regions of increased positive charge tend to correlate favorably with DNA deformation ability, presumably through non-symmetrical charge neutralization by interactions between positively charged amino acid residues and the negatively charged phosphate backbone (285) or by actively attracting the negatively charged DNA to positive residues (156). Observation of alignments of the Alpha and Beta HPV E2 DNA binding domains (Figure 3, 4) appear to support this assertion, as a greater number of conserved positively-charged amino acid residues, both within the nucleotide contact region as well as outside, were present in the Betapapillomaviruses. This observation correlates with the increased presence of GC residues in the spacers of Betapapillomavirus E2BSs. Additionally, a cluster of positively charged residues located c-terminal from the DNA interaction region has been implicated in providing the relative insensitivity to spacer GC content observed with the BPV E2 protein (141). Interestingly, we observed that the consensus MEME motif diverged even within distinct papillomavirus genera. Specifically, the two alpha subclades' consensus binding site possessed an inverted four base spacer. Typically, when the four conserved binding sites are observed individually, the spacer of binding sites 5' of the viral origin of replication tend to be oriented such that the consensus binding site possesses A nucleotides whereas those 3' of the ori contain the inverse, or T nucleotides(257). As a result, given that the E2BS sequence resembles a palindrome, this would likely result in the E2 protein binding in opposite orientation. The functional consequences of this have yet to be fully explored, but may have significant biological implications.

Within the viral genomes of the respective strains, divergence of E2BS locations correlated with tissue type that was infected by the respective strains. This may have

effects on viral transcriptional regulation. There are 4 primary conserved binding sites near the viral origin of replication termed BS1, BS2, BS3 and BS4. E2 binding to the first site (BS1) interferes with TATA box recognition by the TATA binding protein, binding to the second (BS2) and third (BS3) sites causes promoter repression by competition with cellular transcription factors and binding to the fourth site (BS4) up regulates viral early gene expression (141). In addition, binding to BS3 is necessary for DNA replication. When E2 protein concentration is low, the promoter for the E6 and E7 oncogenes is activated and BS4 is occupied. When E2 protein concentration is high, the E6 promoter is repressed and BS1 and BS2 are occupied by E2 (141). Differential affinities for the spacers of these E2BSs have been predicted to play a regulatory role in E2 mediated viral gene transcription (141). The vast differences in number and location of E2BSs identified in this study, however, would seem to suggest that there may be significant differences in this from one virus species to another. Additionally, the E2 proteins of individual papillomaviruses have demonstrated variable ability to tolerate GC content of the four base spacer(141) and binding site methylation(257) may further individualize the specific regulation strategy utilized by each.

All four of the E2BSs in the LCR are almost exclusively AT rich. However, predicted E2BSs outside the LCR generally contained higher levels of GC content. This suggests that these binding sites would tend to have much lower binding affinity for E2. Considering that external binding sites were not conserved between various HPV types and the fact that E2 has numerous functions that are up or down regulated during the course of the viral life cycle, it is difficult to speculate what roles these additional binding sites might play, including remodeling the chromosome structure or potentially blocking the progress of RNA polymerase complexes during the coding process. Further complicating the issue is the fact that, in BPV1, 17 total E2 binding sites have been

previously identified by gel shift assay, many of which had significantly divergent sequences from the consensus (175). However, the study also determined that the binding sites that were more closely related to the consensus generally had the highest binding affinity for E2, so it is likely that the binding sites identified from this study are preferentially filled at more stages of the viral life cycle. This presents a possible regulatory mechanism to control occupation of E2BSs and thus their transcriptional and/or replication effect.

One explanation for the greater degree of variability in mucosal HPVs could stem from the wide tissue types infected by Alphapapillomaviruses. Much of the evolutionary differences observed in the study correlate with differences in preferred infection site. Mucosal epithelia infected by Alphapapillomaviruses ranges from oral to anogenital, all of which could provide a slightly different environment for HPV replication. Additionally, while cutaneous tissue tends to be relatively isolated from the immune system, mucosal epithelia is much more actively surveyed by the immune system and exposed to IgA. This could also potentially serve as a driving force for differentiation in E2 protein function. Previous work has established that differences in tissue type can have significant effect on LCR transcription enhancer activity (213, 253). E2-host co-evolution could then be a potential explanation for the extreme level of tissue specificity exhibited by most members of the papillomaviridae family.

In general, GC content tends to be low in papillomaviruses, presumably as a means of eliminating targets for methylation by the host gene regulation machinery (257). Sanchez et. al. determined that there was an evolutionary selection for CpG methylation sites within the E2BSs of papillomaviruses at positions 4-5 and 9-10(257). Our analysis demonstrated a varying prevalence of G and C nucleotides, respectively, at these sites between the papillomaviruses. Beta and Xipapillomaviruses both possessed a much

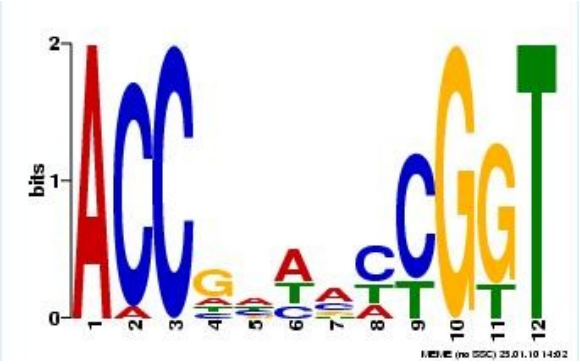
higher prevalence for CpG methylation site at one or more of the potential sites compared to the average for the other genera. Deltapapillomaviruses seemed to favor the presence of a methylation site at the 4-5 position but not at the 9-10 position. For other papillomaviruses, the patterns are somewhat more ambiguous, which is consistent with a previous study (257, 293). As such, if the same holds true for other papillomavirus genera, it is expected that, as our results come from a combination of all the E2BSs, this pattern would be somewhat skewed. A similar effort to examine the individual conserved E2BSs for papillomaviruses beyond the alpha genus would possibly determine if similar methylation patterns exist, but is beyond the scope of this study.

One important observation from this experiment is the large degree of variability between both the proteins and their counterpart DNA binding sites between papillomavirus genera. Deltapapillomaviruses averaged a larger number of E2BSs within the URR (skewed somewhat by the 17 reported E2BSs in BPV1) than any of the other genera examined in this study, and demonstrated a large degree of insensitivity to GC content in the 4 base spacer region. Conversely, the Alphapapillomaviruses showed a preference for A/T nucleotides within the four highly conserved E2BSs in the URR, almost to the point of exclusion at some base positions. The other genera ranged somewhere in between. It's tempting to infer that, as these three groups primarily infect different tissue types (mucosal epithelia for alpha; cutaneous for beta, gamma, lambda, and xi; and fibroblasts for delta) that this in some way represents an element of the adaptive radiation the virus underwent to adopt these infectious substrates. Still, whatever the explanation for this observation, it should remind researchers to be cautious when drawing generalizations between papillomavirus genera E2 proteins, since a particular feature of BPV1 E2 protein may function differently or even be absent

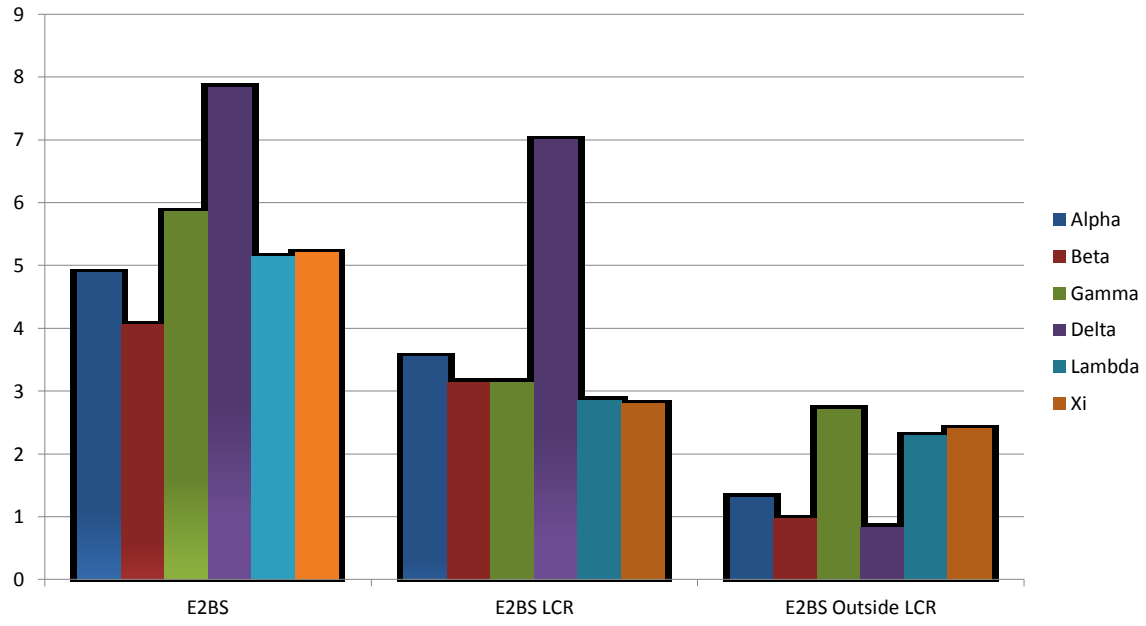
for other PVs, as was eventually discovered to be the case with HPV16 and BPV1's respective utilization of Brd4 for chromosome replication(204).

Figure 2.1: ***Consensus Sequence Analysis of E2BSs Throughout Papillomavirus Genera.*** Well characterized E2BSs from HPV16, 18, and BPV1 were analyzed using MEME software to generate a consensus E2BS motif (Fig 2.1A.) This motif was then utilized by MAST software to search through the complete genomes of 128 papillomaviruses obtained from NCBI and identify sequences with high identity to the consensus. The average number of E2BSs identified per genome were sorted into the six largest papillomavirus genera and were further analyzed to determine if the binding sites were located within or without the upstream regulatory region (URR) of the genome (Fig 2.1B.) Identified E2BSs were then manually analyzed to determine the GC content of their four base spacer regions. Results were again calculated in terms of average GC content of E2BSs for each of the individual papillomavirus genera both inside and outside the URR as well as in total (Fig 2.1C.) Finally, the identified binding sites were used for MEME analysis to identify the consensus E2BS motif for each of the six papillomavirus genera analyzed in this study (Fig 2.1D.)

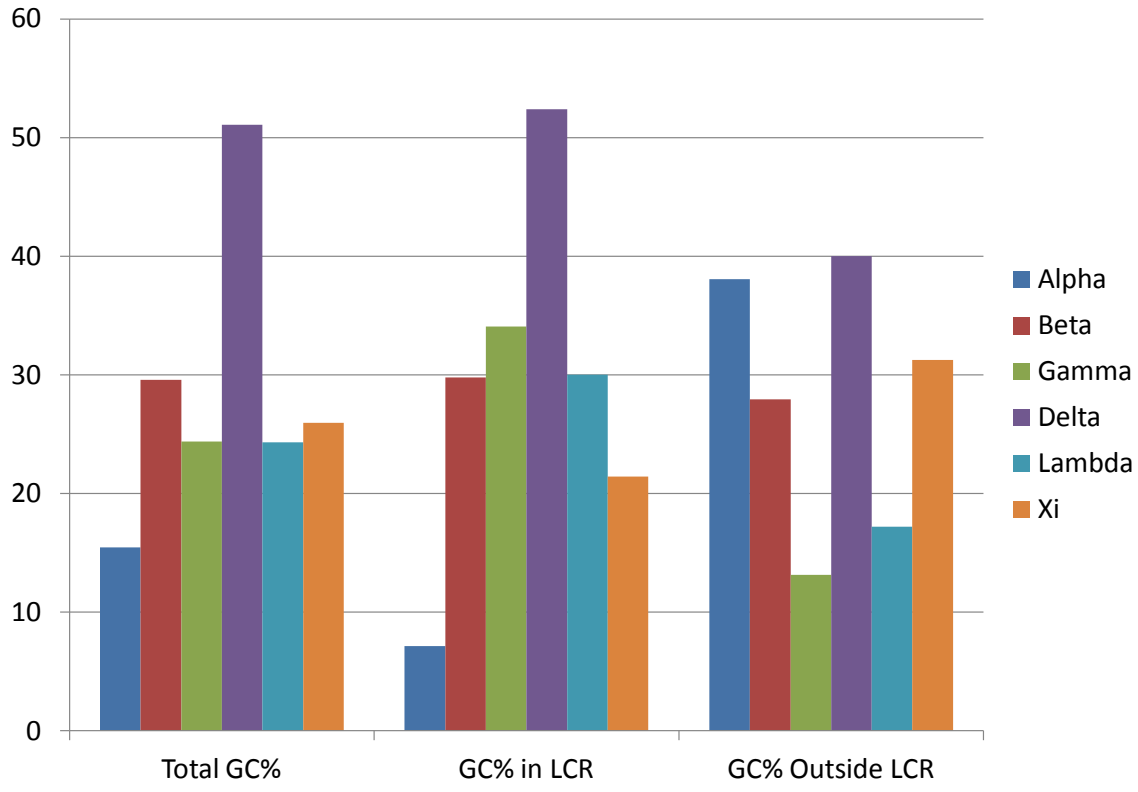
2.1A



2.1B

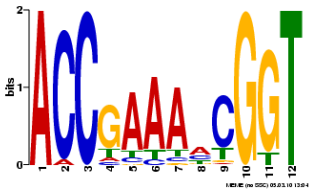


2.1C

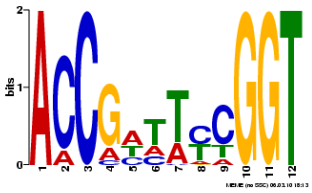


2.1D

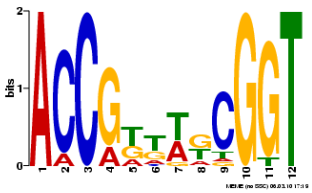
Alpha



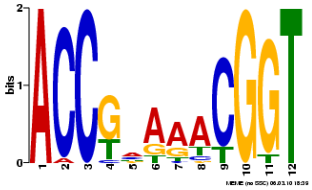
Lambda



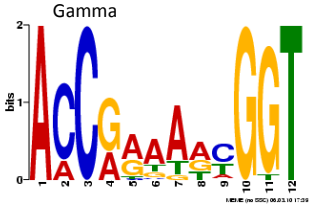
Beta



Xi



Gamma



Delta

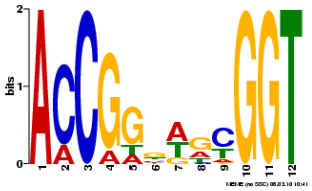
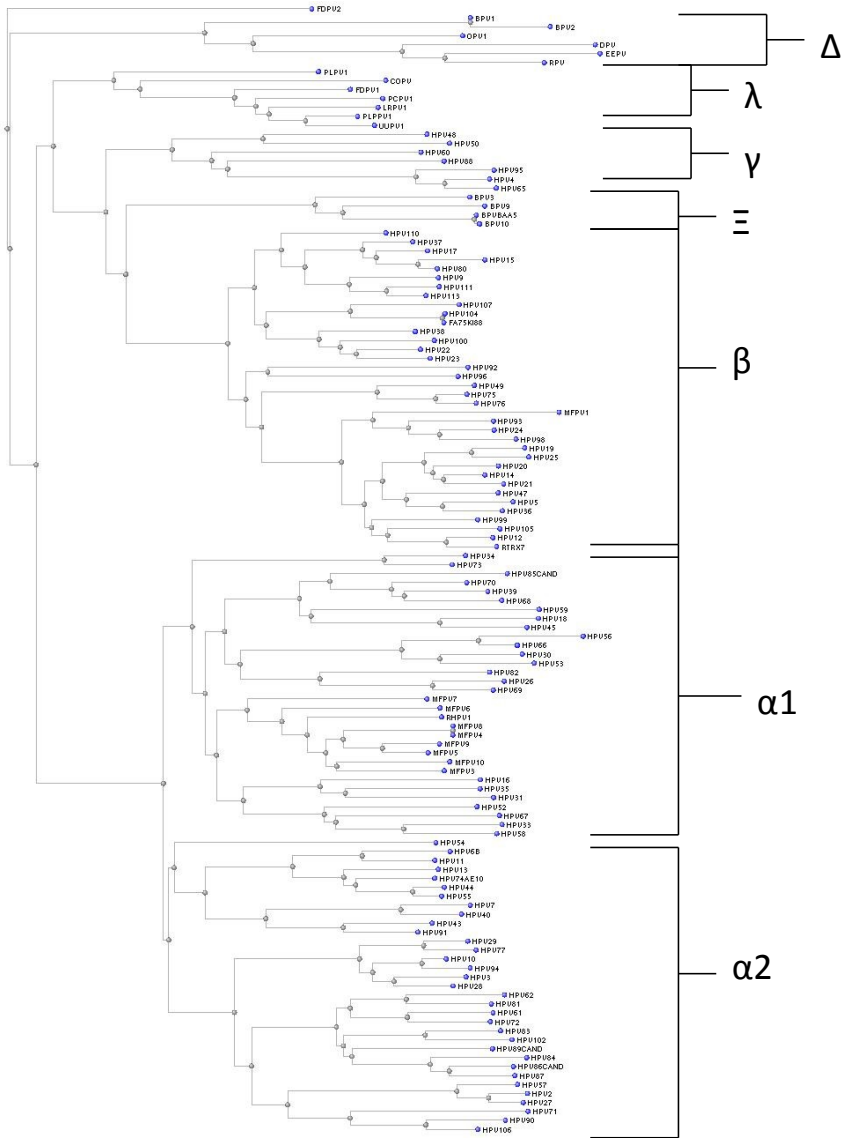
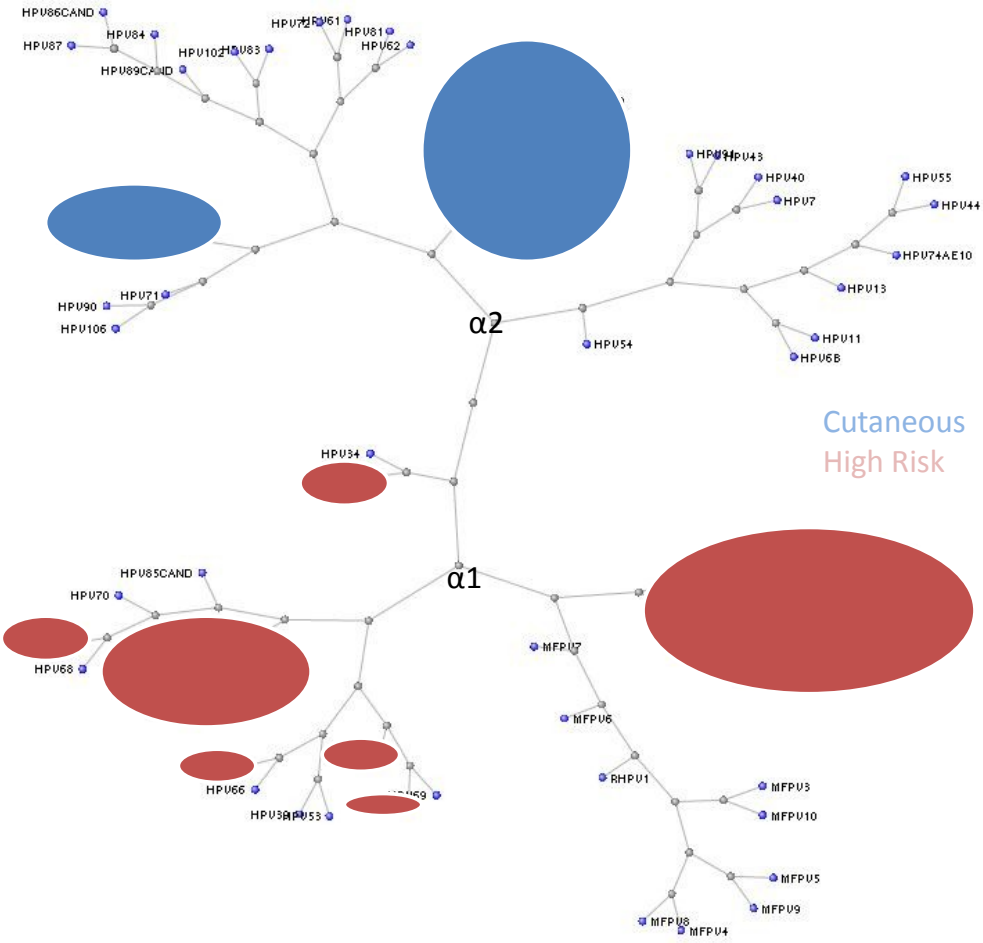


Figure 2.2*: Phylogenetic Analysis of Papillomavirus E2 Protein* E2 protein amino acid sequences for each of the papillomaviruses included in this study were obtained from PDB and used for COBALT analysis. The resulting multiple alignment was then used to generate a phylogenetic tree to analyze papillomavirus evolution in terms of the E2 protein (Fig 2.2A.) Clades were identified corresponding to the classical PV genera and indicated on the tree, as well as two subclades of the alphapapillomavirus genera. These were then expanded and examined individually, and the locations of various types of alphapapillomaviruses (specifically those capable of infecting cutaneous keratinocytes and those possessing a high risk of progression to cervical cancer) were indicated (Fig 2.2B.) HPV E2BSs from part one were then reanalyzed using MEME software to identify a consensus E2BS for the subclades identified in 2b, ie subclade 1 and 2, high and low risk alphapapillomaviruses, as well as those capable of infecting cutaneous keratinocytes tissue (Fig 2.2C.)

2.2A

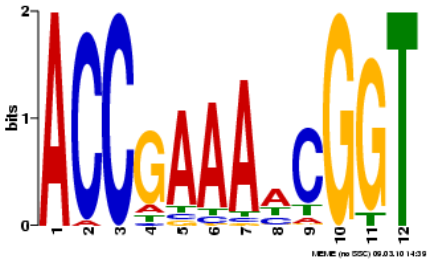


2.2B

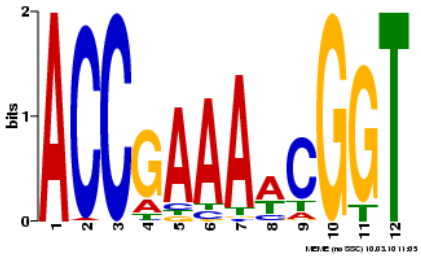


2.2C

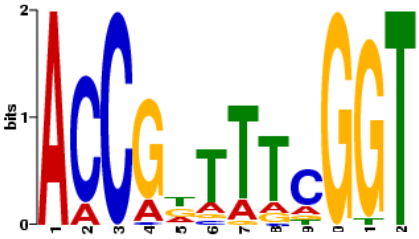
α Subgroup 1



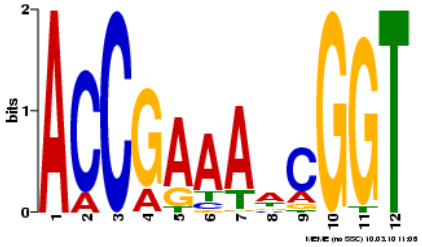
High Risk



α Subgroup 2



α Cutaneous



Low Risk

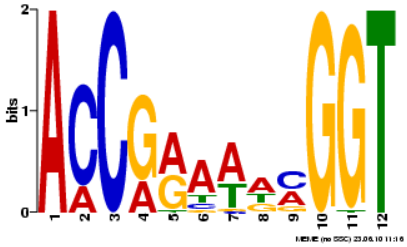
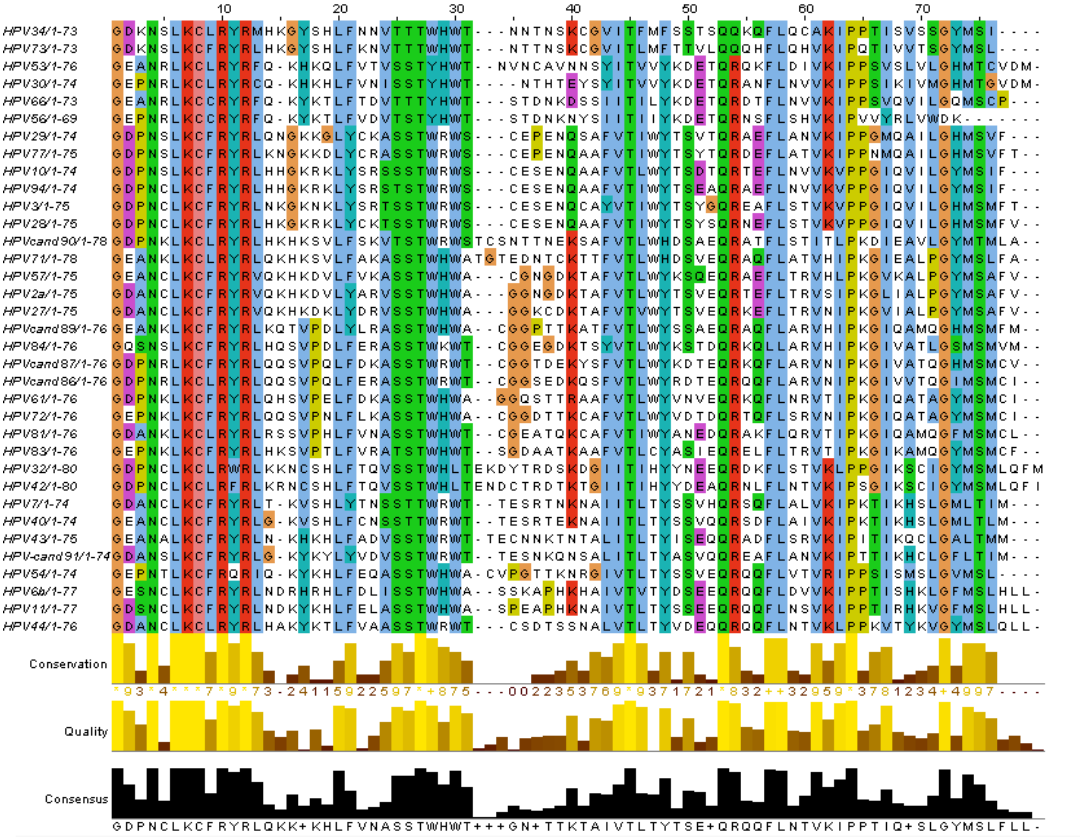


Figure 2.3: ***E2 DNA Binding Domain Protein Alignment*** Amino acid sequence for all known E2 proteins were acquired from NCBI and aligned using Muscle . **(2.3A)** This figure shows the sequence alignment of the Alphapapillomavirus c-terminal DNA binding domain of E2. Colors represent homologous amino acids. **(2.3B)** This figure shows the sequence alignment of the Betapapillomavirus c-terminal DNA binding domain of E2.

2.3A



2.3B

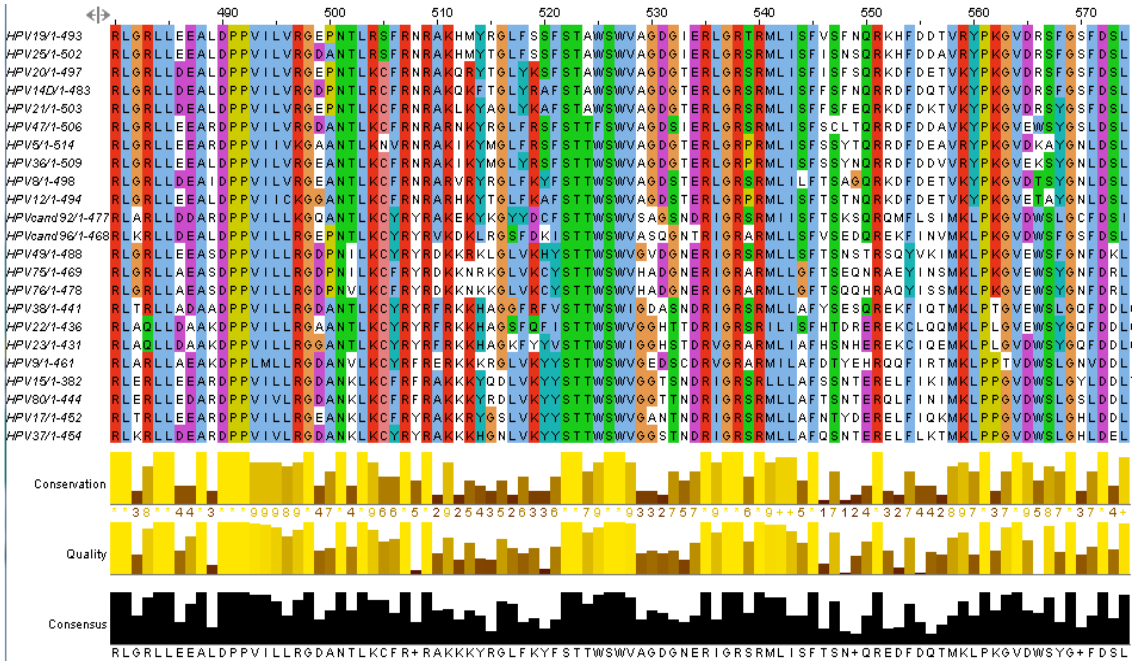
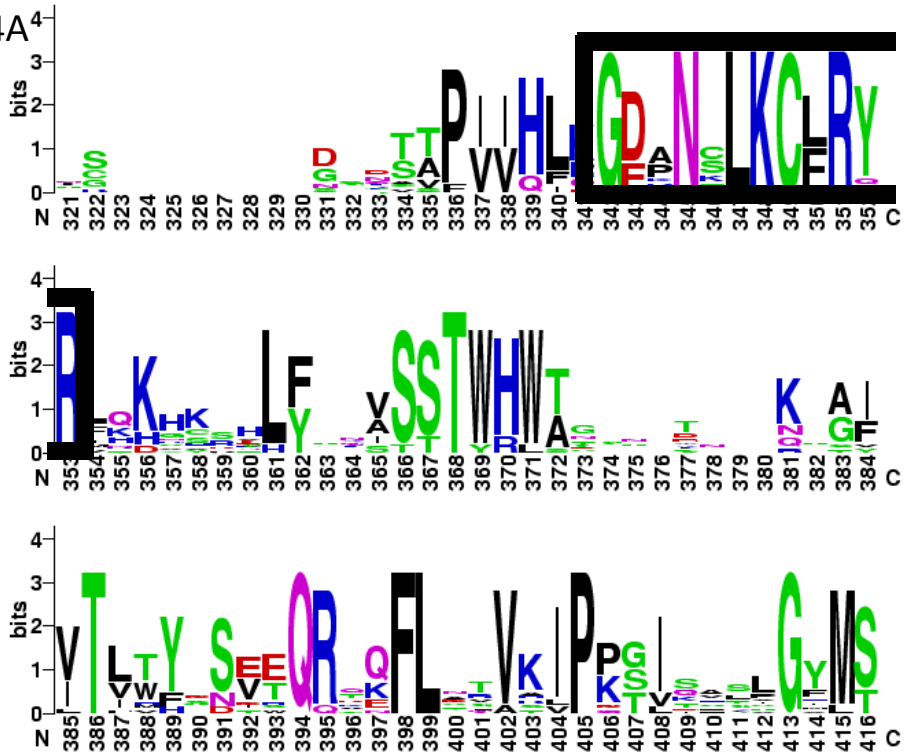
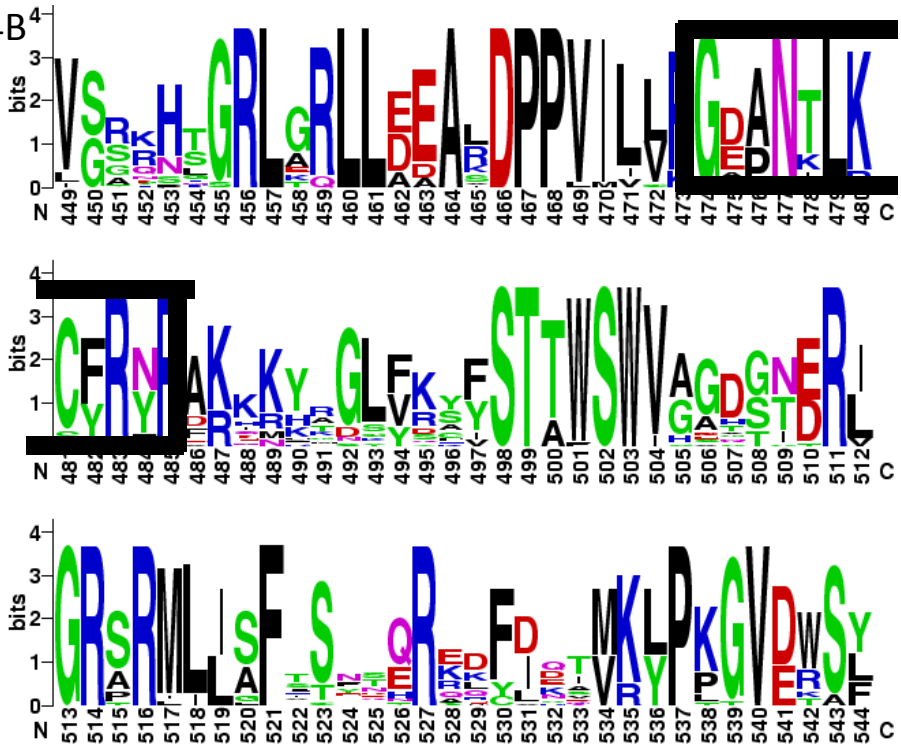


Figure 2.4: *E2 DNA Binding Domain WebLogo* Weblogo was used to generate a graphical representation of the sequence analysis of the c-terminal DNA binding domain of E2. The black box represents the conserved region where E2 protein binds to DNA in the Alphapapillomavirus **(2.4A)** sequence similarity and the Betapapillomaviruses **(2.4B)**.

2.4A



2.4B



Chapter 3

Papillomaviruses Replicate with Varying Success in *Saccharomyces cerevisiae*

Papillomaviruses Replicate with Varying Success in *Saccharomyces cerevisiae*

Published in: Rogers AJ, Loggen M, Lee K, and Angeletti PC. "Papillomaviruses Replicate with Varying Success in *Saccharomyces cerevisiae."* Virology. 381 (1): 6-10. Nov. 10 2008.

Nebraska Center for Virology, School of Biological Sciences

University of Nebraska-Lincoln, Lincoln, NE, 68588-0666, USA

Running title: **Replication of HPVs in Yeast**

* Corresponding Author Mailing Address: Nebraska Center for Virology, E128 Beadle Center, University of Nebraska-Lincoln, 1901 Vine Street, Lincoln, NE 68588-0666

Tel. 402-472-2986, Fax. 402-472-8722, E-mail: Pangeletti2@unl.edu

KEYWORDS: extrachromosal DNA/ persistent infection/ human papillomavirus

ABSTRACT

Human papillomaviruses (HPVs) replicate in mitotically active basal keratinocytes as nuclear plasmids. Two virally encoded proteins, E1, a helicase, and E2, a transcription factor, are important for DNA replication and stable maintenance of HPV episomes in host cells. In previous studies, we have demonstrated that HPV16 can replicate stably in yeast (*Saccharomyces cerevisiae*) (7, 158). In this study, we further demonstrate that multiple HPVs (Types, 6, 16, and 31), when linked to the Ura3 nutritional marker, successfully replicate and are maintained extrachromasomally in yeast. We found differences in replication efficiency; HPV6-Ura3 was the most robust replicator, followed by HPV31-Ura3 and HPV16-Ura3 respectively, while HPV11-Ura3 and HPV18-Ura3 were unable to replicate in the absence of E2 expression. However, we found no evidence that the BPV-Ura3 construct could replicate stably in yeast and the addition of a yeast centromere provided only partial complementation. Together, our studies indicate that there are intrinsic genotype-dependent differences in HPV replication activity in yeast.

INTRODUCTION

Human papillomaviruses (HPVs) are small, double stranded, circular DNA viruses that infect squamous epithelial cells. Two hundred species of HPVs have been identified and are classified into low and high risk categories based on their association with cervical oncogenesis. Integration of the viral genome into host chromosomes is necessary for the development of cervical cancer (306).

The life cycles of papillomaviruses are closely associated to the their host cells. During the course of epithelial cellular differentiation, the virus shifts through three replication phases in response to keratinization: establishment, wherein early viral replication occurs; maintenance, where the viral genome is stably maintained episomally by replicating through a theta intermediate; and amplification, where viral replication shifts to a rolling-circle method and copy number increases in preparation for encapsidation (112). Papillomavirus DNA replication requires primarily cellular factors, recruiting DNA polymerase alpha along with other elements of the cellular replication machinery (240). The accepted model utilizes the viral proteins E1 and E2 for genomic amplification (95, 250, 301).

The E1 protein functions as an ATP dependent helicase and recruits DNA polymerase alpha to act as an elongation factor (117, 217, 242). E2 has multiple functions, including acting as a transcriptional trans-activator, an origin recognition protein, and facilitating binding of E1 to the E1 dependent origin (244). Additionally, E2 serves as a maintenance factor by improving inheritance through mechanisms that are not entirely understood, but may be the result of E2 binding either to chromosomal DNA or mitotic spindles (22, 71, 170, 273). Previous research has indicated that papillomavirus genomes can be replicated and maintained stably in the absence of E1

and E2 (7, 158, 336). Silencing mutations of the individual HPV open reading frames have shown that none of the individual ORFs are required for successful genomic DNA replication and maintenance (7). Since it has been demonstrated that replication of papillomavirus genomes occurs in the absence of viral proteins in yeast, it is apparent then that cellular factors must replace E1 function. For instance, the E1 protein forms hexamers which function in a similar manner to cellular helicases such as Werner's (WRN) and Bloom's (BLM) Syndrome Helicases, members of the RecQ family, and the minichromosome maintenance proteins (MCM) (102, 117, 216, 242). It is conceivable, then, that host factors could be adapted to perform similar replication functions in place of viral proteins like E1, suggesting that an E1-independent mode of replication could be relevant during the maintenance phase of the HPV lifecycle.

Studies have previously indicated that HPV16 is capable of replicating in *S. cerevisiae* in an E1 independent manner (7, 158, 159). Previous efforts by the Lambert laboratory identified regions of the HPV16 genome which are responsible for this trans factor independent replication, showing that portions of the L2 and L1 open reading frames possess replication and maintenance function (158). Plasmids with these cis-acting factors present were capable of long-term, stable, episomal replication independent of viral replication factors.

Having initially established that HPV genomes can replicate in *Saccaromyces cerevisiae,* the HPV/yeast system has proven easy to manipulate for the study of certain aspects of the HPV lifecycle, including transcription, replication and production of virus-like-particles (VLPs). HPVs 6b, 11, 16, 18, and 31 replicate in short-term assays when transformed into competent yeast (5-7, 158, 336, 337). Furthermore, the Frazier laboratory has reported that BPV1 replicates robustly in yeast (336, 337). Recently the Khan laboratory has reported that HPV1 can replicate in yeast, but requires a

centromere to be maintained stably (51). Kim et. al. mapped both *ARS* and *CEN* replication functions in *S. cerevisiae* to the late region of HPV16(158). The great degree of homology between the genomic replication mechanisms of yeast and higher organisms creates the possibility that similar mechanisms could be involved in papillomavirus replication in higher eukaryotes, especially during the maintenance phase when expression of E1 and E2 is minimal.

In this study, we investigate the long term replicative and maintenance competence of five HPV types: HPV6b, 11, 16, 18, and 31 along with BPV1. Replicons containing a Ura3 nutritional marker were created for each papillomavirus, transformed into yeast, and analyzed by Southern Blot to confirm that HPV6b, 16, and 31 genomes were replicating episomally while HPV11 and 18 were not. Notably, BPV1, which had previously been reported to replicate in yeast (336, 337), failed to show significant long term growth in selective media even when complemented with the inclusion of a centromere.

MATERIALS AND METHODS

Yeast Strains, Plasmid Isolation and Transformation Methods. The haploid yeast strain YPH500 (*MAT ura3-53 lys2-801 ade2-101 trp1-63 his3-200 leu2-1*) was used for all described experiments. Yeast was grown on YNB minimal media omitting uracil (Ura) at 30 °C for all transformation and Southern experiments, while YPD complete media was utilized additionally to examine plasmid stability and loss rate. The EZ Yeast Transformation kit (Zymo Research, Orange, Calif.) was used for transformation with the experimental plasmids according to the kit protocol and the Zymoprep kit (Zymo

Research, Orange, Calif.) was used for yeast plasmid minipreps according to company specifications.

Plasmid Constructions Numerous constructs were created using similar methods. The Ura3 gene was ligated into unique restriction sites in either the papillomavirus genome or vector sequence, selected in *E. coli* grown on Luria Broth media containing Ampicilin, and the DNAs were isolated using the Qiaprep Spin Miniprep Kit (Qiagen Sciences, Maryland 20874) in accordance with the manufacturer's instructions. The Ura3 markers were cloned into the various constructs as follows: pPA102 (pGEMT, AgeI), pPA103 (HPV16, XhoI), pPA104 (Puc18, SalI), pPA106 (HPV31, SpeI), pPA112 (PPR 322 with HPV18 ligated into Nco site, AVRII), pPA116 (HPV6, AgeI), pPA117 (HPV11, AgeI), and pPA118 (BPV, Mlu I). Construct PA119 was created by digesting pΔYac with AvrII, releasing the CEN element, and then ligating this fragment into a unique AvrII site in pPA118. PRS 316, an ARS+ CEN+ Ura3+ yeast replicon, was included as a positive control.

DNA Replication in Yeast Two-hundred ng of each plasmid was transformed into YPH500 yeast, plated on Ura- selective agar, and incubated for 3 days at 30°C. Plates were then scored for number of colonies formed and restreaked on selective solid agar, grown for an additional 3 days, and inoculated into 5mL of Ura- liquid media and allowed to grow for approximately 40-50 cell generations. OD_{600} was recorded for each sample to obtain approximate number of cell equivalents per milliliter, and samples were diluted in order to equalize this number, allowing for a determination of copy number by Southern Blot. Low molecular weight DNA was isolated from liquid culture as described above. The DNA was then loaded onto a 1% agarose gel, electrophoresed, transferred to nitrocellulose, and probed with ^{32}P-radiolabeled pPA104. Radiolabeling was

performed using the Amersham Rediprime II kit (GE Healthcare UK, Buckinghamshire) in accordance with the manufacturer's instructions.

Growth Curves 5 milliliters of liquid Ura- media were inoculated to an OD_{600} of 0.15 with each of the papillomavirus genomes transformed into *S. cerevisiae,* as well as PRS316. In addition, 5 mL of YPD media was inoculated to an equivalent OD_{600} with untransformed YPH500 yeast to establish a standard doubling time. OD_{600} for each sample was recorded at time points 0, 3, 6, 8, 24, and 30 hours after inoculation. Growth rate was then determined by calculating doubling time during the mid log growth phase using the formula $T_d=(t_f-t_i)*[\log(2)/\log(q_f/q_i)]$ where T_d is doubling time, t_f and t_i are time final and initial, and q_f and q_i are the OD_{600} values at the t_f and t_i.

Plasmid Loss Rate Assay The plasmid loss rate per cell generation was calculated utilizing a method similar to that described by Marahahrens et. al. (196) Briefly, approximately 100-200 cells were plated onto YPD and YNB-Ura plates after being removed from selection at time points of 0, 4, and 8 hours post inoculation. All sets of plates were then incubated for 3 days and scored for colony growth. Plasmid loss rate was then calculated using the formula $L=P_f-P_i/T$ where L is the percent loss rate per cell generation, P_f is the percentage of growing yeast which contained the plasmid at the final time point, P_i is the percentage of growing yeast with the plasmid at the initial time point, and T is the number of cell generations.

RESULTS

Varying Replicative Efficiencies of Different HPVs in Yeast

In order to investigate the DNA replication capability of varying Papillomaviral species in yeast, we generated plasmid constructs by cloning a Ura3 nutritional marker along with the complete papillomaviral genomes into the multicloning site of Puc18 (Figure 3.1). These were then transfected into YPH500 yeast and observed for ability to generate stable colonies that can be streaked onto new plates and grown in liquid media.

Yeast transformed with the different HPV/Ura3 constructs demonstrated very different growth characteristics when plated on selective media (Figure 3.2). Some, like pPA112 and pPA117, showed very little growth on the plates, with at best only one to two colonies growing (likely the result of recombination or integration of the Ura3 marker into yeast genome). We had previously demonstrated HPV 16-Ura's long term plasmid maintenance in yeast (7), and confirmed that result here. pPA116 and pPA106 transformed yeast seemed to grow as well, or better than pPA103 on selective media. Significantly, BPV1-Ura (pPA118) showed no colony growth on selective media. Inclusion of a yeast centromere in an HPV1 plasmid construct had previously been shown to allow for genomic maintenance (51). Cloning of a *CEN* element into pPA118 resulted in small, slow growing yeast colonies containing the plasmid (Figure 3.2) which were difficult to further propagate.

Southern blotting was performed to confirm that the HPVs which replicated successfully in yeast were being maintained episomally. As the HPV genomes have varying sequences and restriction digest profiles, DNA was run uncut after collection by

yeast mini-prep and will appear on the gel to possess both the low molecular weight supercoiled form along with a number of higher molecular weight nicked-supercoil forms. Expected molecular weights for the products are as follows: PA116 (HPV6b-Ura3) 11,485 bases, PA103 (HPV16-Ura3) 11,395 , PA106 (HPV31-Ura3) 11,401 bases, PRS316 (ARS+CEN+ control) 4887. HPV plasmids replicated at varying copy numbers (Figure 3.3), with HPV6 and 31 appearing to be between 10 and 50 copies per cell while HPV16 was maintained at 1-5 copies per cell, as compared to input standards.

Growth Rates and Plasmid Stability of HPV-Ura Constructs

Experiments with liquid media inoculations allowed for a closer examination of the growth characteristics of each HPV genome transformed yeast culture. By observing changes in OD_{600} over several hours, a doubling time for the mid-log phase of each yeast culture was obtained (Figure 4). HPV6-Ura (pPA116) had the longest doubling time, 7.5 hours, while pPA103 had the shortest at 3.3 hours. All three HPV transformed yeast strains grew more slowly than untransformed YPH500 in YPD complete media or PRS316 positive control transformed yeast.

Additionally, experiments were performed to determine and compare the rate at which plasmid is lost from yeast once removed from selection. All four plasmids which could replicate in yeast were grown first in selective media (-Ura) to mid-log phase and diluted to an OD_{600} of 0.15 into new cultures containing nonselective media (+Ura). The cultures were grown, and at fixed time points, an equivalent number of cells were plated from each culture onto both selective and nonselective media plates, allowing for determination of the ratio of yeast which had maintained the plasmid vs. yeast which had lost it over their individual number of cell divisions. The results are shown in Table 1. All

three HPV constructs replicated with a high degree of stability similar to that of the ARS+ CEN+ positive control, PRS316.

DISCUSSION

Our results indicate that certain HPVs (6, 16 and 31) replicate stably in yeast, while some do not (HPV 18 and 11). The definition of stable replication used in this study is based primarily on the ability to form colonies on selective media which can be restreaked and subcultured onto a new solid media plate as well as liquid media. Previous studies performed by Angeletti et. al. as well as Frazer et. al. have previously established that HPV 11, 18, and BPV1 are capable of replicating successfully in yeast during short term assays utilizing complete, non-selective liquid media (2,24,25). Growth on solid, selective media characteristically includes a 1-3 day growth period wherein yeast colonies are either not visible or very small, resembling petite mutants. Consequently, a replicon which does not replicate stably will not develop to the point of seeing visible colonies and, thus, would be scored as not replicating stably.

The great degree of conservation between the DNA replication systems of yeast and higher eukaryotes suggests the possibility that certain factors may be involved in papillomavirus genome replication and maintenance for both systems. As stated previously, papillomaviruses spend the majority of their life cycle in a maintenance phase, replicating episomally at low copy number, with only minimal expression of the virus's replication factors E1 and E2. We have previously demonstrated that HPV16 is capable of replicating with *cis* acting factors alone in the absence of E1 and E2 (7). Also, Hoffman et. al. have previously demonstrated that HPV16 replicates in a once-per-S-phase manner in certain cell lines, depending on the presence or absence of viral

replication factors, while HPV31 only replicates randomly under the study conditions (147).

All of these results point to a large degree of reliance upon host factors for regulation of viral genome replication and maintenance. The differences in replication and/or maintenance success of different HPV types could represent genotype-dependent differences in cis-acting elements in the late region of the genomes as compared to that of HPV16, which was recently shown to provide both functions by Kim et al. Completing similar mapping experiments of other HPV types should help identify which *cis* acting factors are playing a role in yeast replicative success, but also if the trend for late region replication and maintenance functions observed by Kim, holds true for other HPVs. In addition, the experiments described by Hoffman et al., if performed in yeast, should shed light on the mode of replication being utilized. Furthermore, while all the HPVs in this study replicate successfully in host cells, the difference in replicative success shown in this study may imply a greater or lesser degree of utilization of host replication factors between varying HPV types and differing dependence on E1 and/or E2.

Notably, there seemed to be some degree of difference when comparing the Southern Blot versus the liquid culture growth curves. Specifically HPV16 appeared to have a reduced copy number compared to the other HPV replicons but demonstrated a similar growth rate in transformed yeast. Presumably an increased copy number would lead to increased transcription of the Ura3 nutritional marker and, thus, more growth. The reason for this discrepancy is not currently apparent, but one potential explanation could be an underestimation of HPV16's copy number by the assay. In any case, differences between the growth rates of the HPV replicons as compared to controls were not appreciably different.

Contrary to reports that BPV could replicate in yeast (336), we found that BPV-Ura was unstable during long-term replication. In short-term yeast experiments, we previously showed that BPV could generate Dpn I resistant DNA products, indicating that replication is initiated but that DNAs are most likely not stably maintained (7). However, the addition of a centromere region (pPA119 plasmid), resulted in only partial complementation of maintenance function. This result may be related to the much greater number of E2 binding sites contained within the BPV genome (17 for BPV, as compared to 4 E2BS for HPV16) (177), implying perhaps a greater reliance on E2 tethering. Despite our careful analysis, BPV1 does not appear to be stably maintained in yeast. It is possible that this is the result of differences between the mammalian Brd4 protein and its yeast homologue, Bdf1, which lacks the C-terminal domain that has been shown to interact with E2. Fusion proteins of Bdf1 and the C-terminus of Brd4 restores E2 maintenance function (41), and it is likely that a similar effect would be observed here.

FIGURE 3.1. *Plasmid Maps*. Into each HPV genome, a Ura3 gene cassette was introduced as shown in each of the plasmid maps. The pPA103 vector contains the HPV16 genome with the Ura gene inserted at nt 7309 of the genome, as indicated. PPA106 contains the HPV31 genome, pPA112 contains the HPV18 genome, pPA116 contains the HPV6 genome, pPA117 contains the HPV11 genome, pPA118 contains the BPV genome, and pPA119 contains the BPV genome with a yeast centromere inserted. Control plasmids included pPA104 (puc-Ura3), pGemT-Ura3 as negative controls and pRS316, as a positive control.

Figure 3.1

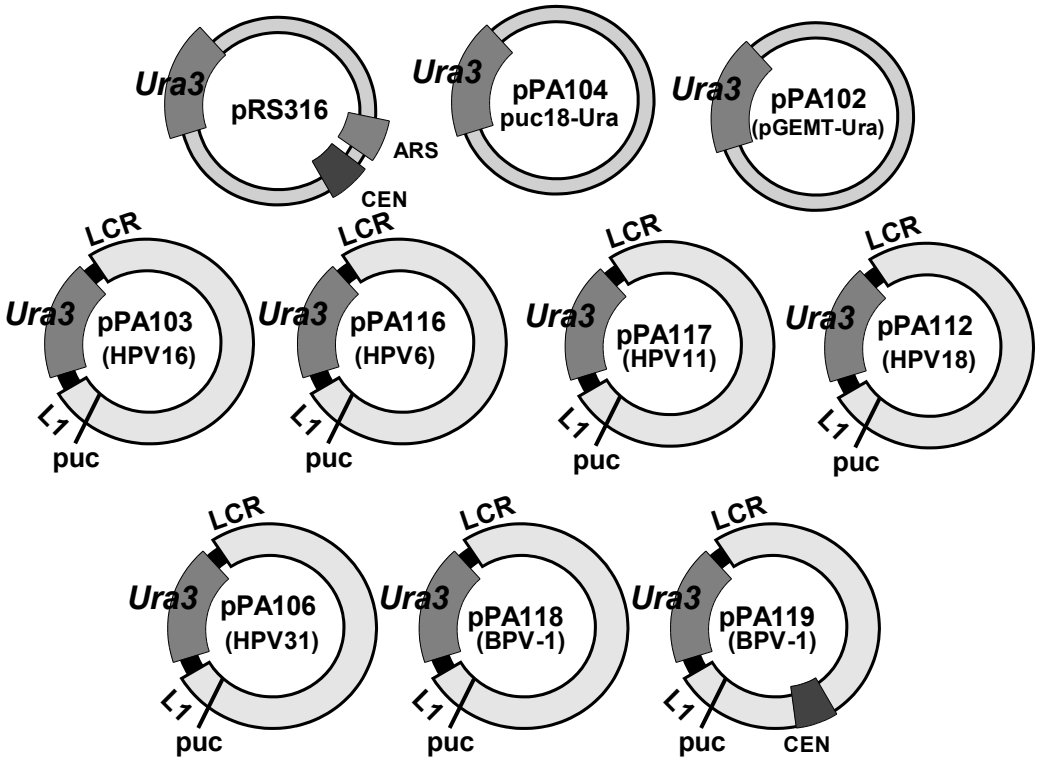


Figure 3.2. ***Yeast Transformations*** Two-hundred ng of each plasmid construct was transformed into YPH500 yeast, plated onto YNB –Ura media, and allowed to grow for 72 hours. Plates were then scanned and scored for growth.

Figure 3.2

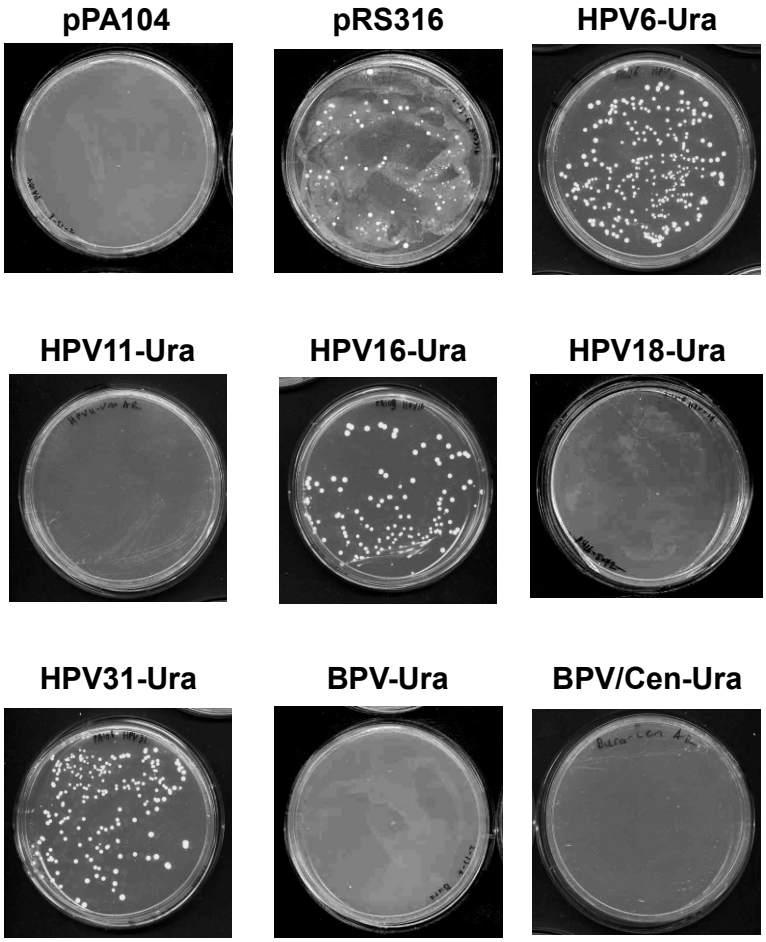
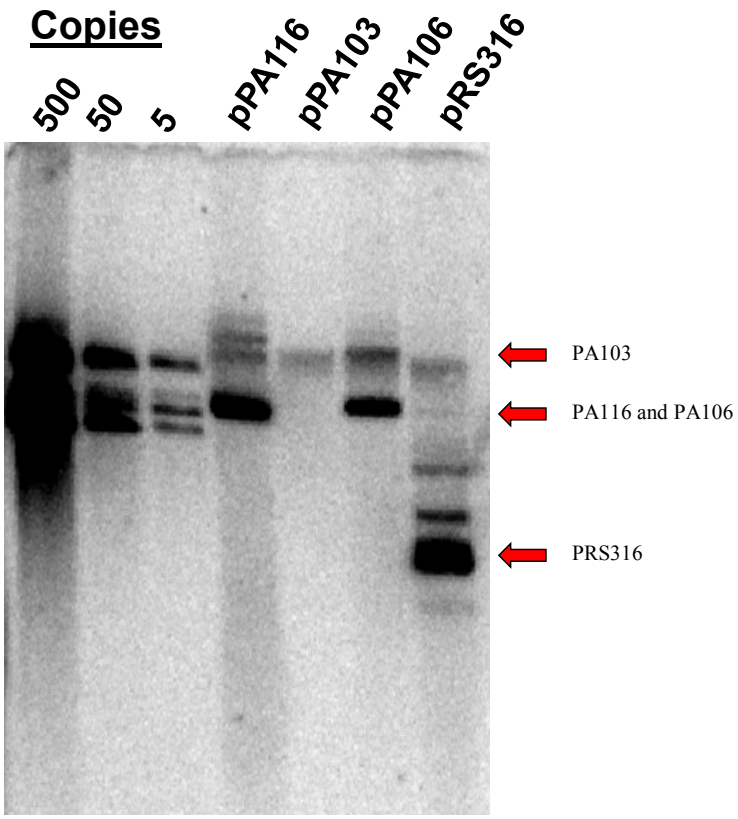


Figure 3.3. ***Southern Blot*** All successfully replicating constructs (HPV6 Ura, HPV16 Ura, HPV31 Ura, and PRS316) were inoculated into 5mL of YNB-Ura media and incubated for 2 days. Small molecular weight DNA was then harvested via yeast miniprep, loaded into 1% agarose gel, and electrophoresed. DNA was then transferred to a nitrocellulose membrane and analyzed via Southern Blot pPA116 (HPV6 Ura), pPA103 (HPV16 Ura), pPA106 (HPV31 Ura), and pPRS 316 were probed utilizing P32 radiolabelled Ura-3. . Control volumes of HPV6 Ura were utilized in the first 3 lanes to provide an idea of relative copy number. Plasmids are run uncut and circular and, as such, will show the predicted lower molecular weight supercoiled form as well as 1-3 nicked supercoil forms which will appear to possess a greater molecular weight. Expected molecular weights are as follows: PA116 (HPV6b-Ura3) 11,485 bases, PA103 (HPV16-Ura3) 11,395 , PA106 (HPV31-Ura3) 11,401 bases, PRS316 (ARS+CEN+ control) 4887. Arrows indicate the supercoiled form visible on the gel.

Figure 3.3

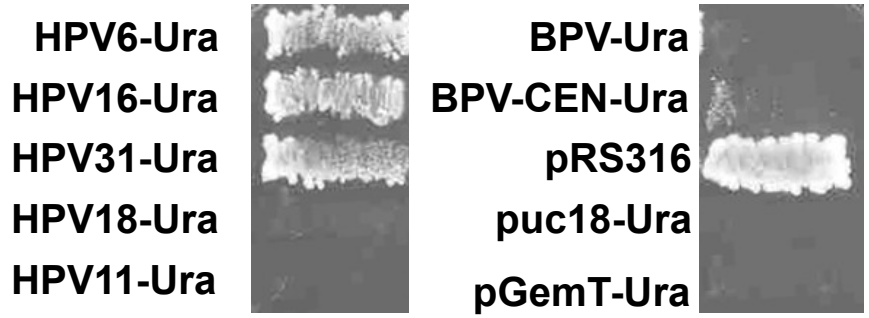


Probe: Ura3

Figure 3.4A) ***Plasmid Construct Master Plate*** After initial growth upon being transformed into yeast, colonies were restreaked onto an additional –Ura plate both to provide for ease of manipulation and also to examine long term maintenance. All plasmids which grew initially continued to grow stably on the master plate. Additionally, BPV-CEN showed signs of limited growth, evident as small colonies. **3.4B) Growth Curve** 5 mL of liquid media was inoculated to an OD600 of .15 with each of the successfully replicating plasmid constructs (HPV6, 16, and 31 Ura as well as p RS316.) Additionally, a similar OD600 was generated with YPH500 alone grown in YPD complete media. OD600 was recorded at time points 0, 3, 6, 8, 24, and 30 hours post inoculation. Growth rate was then determined by calculating doubling time during the mid log growth phase

Figure 3.4

a)



b)

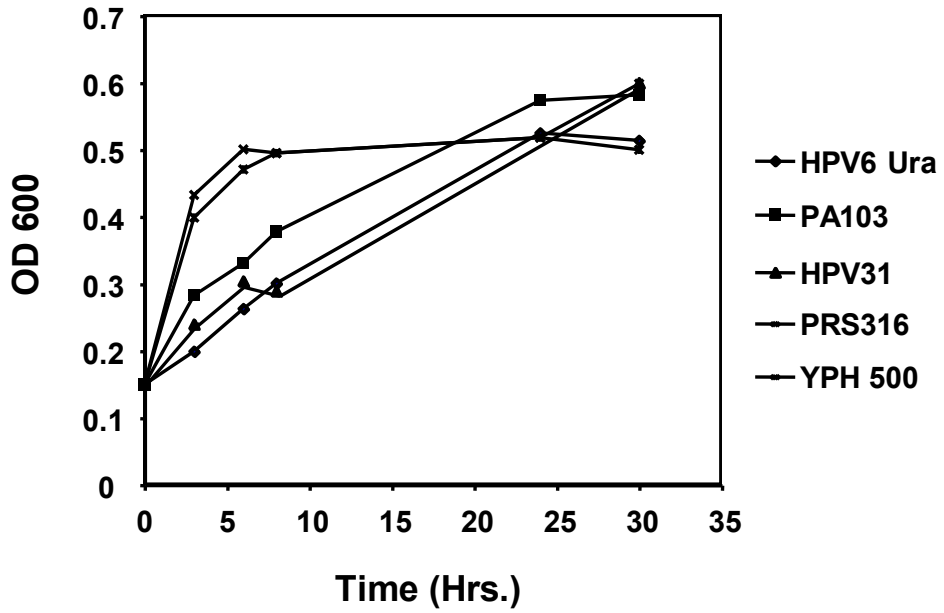


Table 3.1. ***Plasmid Loss Rate*** Yeast transformed with the successfully replicating plasmid constructs were inoculated into YPD complete media. Approximately 100-200 cells were the plated onto YPD and YNB-Ura plates after being removed from selection at time points of 0, 4, and 8 hours post inoculation. All sets of plates were then incubated for 3 days and scored for colony growth. Plasmid loss rate was then calculated using the formula $L=P_f-P_i/T$ where L is the percent loss rate per cell generation, P_f is the percentage of growing yeast which contained the plasmid at the final time point, P_i is the percentage of growing yeast with the plasmid at the initial time point, and T is the number of cell generations.

Table 3.1

Construct	Percent Loss Per CG
pPA116 (HPV6)	2.63
pPA103 (HPV16)	0.64
pPA106 (HPV31)	2.40
pRS316	1.56

Chapter 4

Interaction of TRF2 with HPV16 E2 and Shelterin's Role in HPV16 Plasmid Stability

A Role for Telomere-related Factors in HPV DNA Maintenance

Nebraska Center for Virology, School of Biological Sciences, University of Nebraska-Lincoln, Lincoln, NE, 68583-0900, USA

***Author's Note: Work for figures 2, 3, and 6 was completed by Daraporn Pittayakhajonwut and are included here for completeness.**

ABSTRACT

Papillomaviruses (PV) are small non-enveloped viruses which contain a single molecule of circular supercoiled double stranded DNA (339). The accepted replication model relies on the viral proteins E1 and E2 for genomic amplification (95, 250, 301). Previous research, however, has indicated that papillomavirus genomes can be replicated and maintained stably in the absence of E1 and E2 (7, 158, 336). The observations suggest the maintenance phase of the HPV life cycle may be adapted to extremely low or no E1 and E2.

HPV genomes can persist as episomes in infected individuals for years, which contribute to the development of cancers. Bovine papillomavirus (BPV) E2 tethers newly synthesized genomes to mitotic chromosomes to ensure faithful partitioning of genomes to daughter cells. However, further study demonstrated that, while all papillomavirus E2 proteins studied thus far utilize Brd4 for transcriptional purposes, it is dispensable for plasmid maintenance in HPVs (206). The mechanism of E2-dependent maintenance function thus remains somewhat cryptic. Although binding of E2 to Mitotic Chromosomes is consistently observed, the binding location is not conserved amongst HPV types (231). A recent study suggests that E1 is required for localization of HPV16E2 and viral DNA to fragile sites on Mitotic Chromosomes (255).

HPV16 possesses an E2-independent *cis*-acting maintenance function mapped to the L2 and L1 ORFs (nt 4538-7013)(245). Interestingly, the late region (L2, L1, and the LCR) of HPV16 contains four nine-base DNA sequences corresponding to binding sites for telomere maintenance proteins like TRF2. Further analysis of several HPV genomes revealed that TRF binding sites are relatively common in HPV genomes and are usually found within the late region. Site-directed mutagenesis of these sites in HPV16 resulted in increased plasmid instability, copy number changes, and in some

cases complete loss of plasmid maintenance in human and yeast cell models. Further studies revealed that telomeric protein levels are altered in HPV16 harboring cells, and chromatin immunoprecipitation (ChIP) results indicating TRF2 and other telomere-related proteins such as POT1, TIN2 and BLM helicase are able to bind these telomeric nonamer sequences. Additionally, it was demonstrated that HPV16 E2 protein interacts with the TRF2 scaffold protein, along with a number of other shelterin components. Furthermore, mutating these telomere binding sites to prevent protein binding induced a moderately destabilizing effect on papillomavirus replication in yeast. Long-term replication assays using a plasmid containing an OriP dyad symmetry element and a series of mutants lacking TRF binding sites also induced plasmid instability in transfected mammalian cells, albeit to a lesser extent than that observed in yeast. Since tight regulation of copy number is an important part of the HPV lifecycle, it can be concluded that TRF proteins play an important role in establishing successful DNA maintenance of HPV.

INTRODUCTION

Papillomaviruses (PV) are small (55nm diameter) non-enveloped viruses of icosahedral capsid symmetry that encapsidate a circular supercoiled double stranded DNA genome (339). This family of viruses infects the stratified epithelia of the skin or mucosal surfaces, which include the mouth, airways, and anogenital tissues of vertebrate animals (75). Mucosal HPVs can be further classified into high-risk and low-risk, with the former being the causative agents of cervical cancer as well as some vaginal, anal, and penile cancers (37, 57, 220), typically as a result of genomic integration and resultant overexpression of the viral E6 and E7 oncogenes(306).

Mechanisms ensuring the high-fidelity replication and efficient segregation of the newly replicated viral DNA to the dividing cells are key features of successful persistence of HPV infection. Papillomavirus replication requires recruitment of cellular factors, including DNA polymerase-α along with other elements of the cellular replication machinery (240). The established model relies on the viral proteins E1 and E2 for genomic amplification (95, 250, 301). E1 unwinds the viral genome through its ATP-dependant helicase activity. E1 is loaded onto the origin by the E2 protein, which is colocalized at the HPV replication foci with the L2 protein (76). Previous studies, however, have indicated that papillomavirus genomes can be replicated and maintained stably in the absence of E1 and E2 (7, 158, 336)

An analogy for the possible strategy of stable HPV DNA maintenance comes from similar mechanisms observed in other DNA tumor viruses, such as Epstein-Barr virus (EBV) and Kaposi's sarcoma-associated herpesvirus (KSHV) that initiate similar latency stages during their life cycles (19, 21, 140, 188, 272, 324). "Hitchhiking" by

tethering viral genomes to chromosomal DNA by use of a virally encoded protein as a tethering molecule to allow partitioning and nuclear retention of the viral genomes during mitosis is the common feature among these persistent viruses (38, 110). A single viral protein acts as a molecular linker that establishes a bridge between the viral genome and host mitotic chromosomes. The viral transactivator Epstein Barr Virus Nuclear Antigen 1 (EBNA1) and transcriptional repressor Latency Associated Nuclear Antigen (LANA-1) function as the bridge molecules for the hitchhiking strategy thought to be used by EBV (179) and KSHV (18, 19), respectively.

In the case of papillomaviruses, a similar mode of viral DNA maintenance has been characterized, mainly based on studies related to chromosomal segregation in bovine papillomavirus type 1 (BPV1). E2 polypeptides contain three probable domains: a DNA binding domain (DBD) located at the C-terminus, an N-terminus transactivation domain, and an internal “hinge” domain. Both the C-terminal and N-terminal domains are well conserved within the PVs (254). E2 protein mediates genome maintenance by interaction of its transactivation domain with mitotic chromosomes while the DNA binding domain tethers genomes through the association with multiple E2 binding sites. Further studies have identified that Brd4, a cellular bromodomain protein, is a major component of the tethering complex that attaches the viral genomes to host mitotic chromosomes (25, 200, 207, 329). Both the BPV1 E2 protein and viral genomes colocalize with Brd4 on mitotic chromosomes in punctate spots with no specific attachment sites identified (274). Disruption of the E2-Brd4 interaction dissociates E2 from chromosomes and abolishes viral genome maintenance (200).

Although this interaction is the mechanism of DNA maintenance for BPV1, the interaction between E2 and Brd4 is required for E2-mediated transcriptional regulation but not for genome tethering of other PV subtypes (205, 263, 318). Further analysis has

demonstrated that E2 mutants lacking Brd4 binding remained attached to mitotic chromosomes (230). Therefore, different PVs might utilize different cellular proteins or alternate strategies to maintain the viral genome in the host cells.

In a previous study, our laboratory has mapped maintenance elements that provide plasmid stability in two distinct regions in HPV16 genome (245). These HPV16 subgenomic fragments are located outside of the LCR that contains multiple E2 binding sites (E2BS) and they function as maintenance elements in the absence of any viral protein(245). This finding is consistent with a separate study in yeast in which subgenomic fragments of HPV DNA can autonomously replicate and persist in the absence of E1 and E2 proteins (7, 158). Together, these studies led to the hypothesis that the interaction of viral cis-elements with cellular factors may influence viral DNA maintenance. Sequence analysis identified several possible cellular candidates that can bind to the viral genome and contribute to HPV DNA maintenance. Of those potential candidates, four nine-base DNA sequences of telomeric DNA (TTAGGGTTA) are found within the viral genome, three of which are located in the *cis*-element that is part of the previously identified maintenance element (Figure 1A). TRF binding sites are prevalent among many HPV genomes, whereas there are none in BPV1 (Figure 1B).

Plasmids which contain the Epstein Barr Virus latent origin (OriP) utilize a similar TRF binding site system for long-term plasmid replication and stability. The dyad symmetry region (DS contains 4 sites of nine-base telomeric repeat DNA in the DS region, similar to those found in the late region of HPV. These repeats are required for DNA replication (225) and plasmid maintenance (225, 325). These telomeric repeats allow binding of TRF2 by DNA affinity assay (87), which is required for plasmid replication and maintenance (87, 323). One important function of TRF2's role in OriP replication is recruiting the proteins of the Origin Recognition Complex (ORC). TRF2

interacts with EBNA1 (169), a structural and functional homologue of the HPV E2 protein, and this interaction is required for the replication and maintenance effect (85, 87). These processes are inhibited by overexpression of a TRF2 dominant negative plasmid with the N and C terminal domains deleted (85).

Kaposi's Sarcoma Herpesvirus (KSHV) also encodes a structural and functional homologue of the HPV E2 and EBV EBNA1 protiens, Latency Associated Nuclear Antigen (LANA.) LANA is responsible for latent phase DNA replication, gene expression, and segregation, again similar to the E2 and EBNA1 proteins (20). A study performed by the Renne laboratory demonstrated that, along with thirty other proteins, TRF2 associates with a seventy base pair minimal replication element from the KSHV genome (149). LANA co-immunoprecipitates with TRF2, but interestingly only when cotransfected with a plasmid containing an intact, wild type copy of the viral tandem-repeates (TR). These observations encouraged the hypothesis that the TRF binding sites in *cis* and TRF proteins in *trans* contribute to episomal maintenance of HPV genome.

The telomere is maintained through the action of a number of proteins combining together into a protein complex called shelterin. The shelterin/telosome complex functions primarily by bringing the three telomeric DNA binding factors (TRF1, TRF2 and Pot1) into the same large complex (78, 181, 237) along with Ras related protein 1 (RAP1) (173), and TRF1-Interacting Protein (TIN2) (77). TRF1 and 2 (telomeric repeat factors 1 and 2) bind duplex telomeric DNA (43), almost entirely associated with cellular chromatin (288). TRF1 and 2 share a common architecture defined by two conserved regions: a TRFH domain that mediates homodimerization and a carboxy-terminal DNA binding domain of the SANT/Myb family (43). TRF2 promotes development of T-loop structures (134, 283), potentially as a result of positive supercoiling (4). TRF2 also

serves to stabilize T-loops (116) through their N-terminal domain's ability to bind ss-DNA in a number of secondary structures. TRF2 also interacts extensively with members of the cellular DNA repair system including Ataxia Telangiectasia Mutated (ATM) and the MRE11 recombination-repair complex(40, 80). TRF2 and POT1 proteins physically bind to the RecQ helicase, Werner syndrome ATP-Dependent Helicase (WRN), and then stimulate the helicase activity to unwind duplex telomeric substrate (127). The precise roles of RecQ helicase in telomere maintenance are unclear, however, they likely to function in recombination and/or replication of telomeric ends. TRF2 plays a duel role in DNA repair, inhibiting the action of repair proteins acting on the ends of chromosomes while, at the same time, being phosphorylated by ATM and relocating to sites of DNA breaks to improve homology based recombination repair(275).

In this work, Chromatin Immunoprecipitation (ChIP) assays were used to examine the interaction of TRF2 and its related proteins with the predicted TRF binding site in HPV16 DNA. We also investigated the role of the suspected TRF binding sites and the associated proteins in viral DNA maintenance. It appeared that telomere-related proteins such as TRF2, POT1, TIN2, and Bloom Syndrome Protein (BLM), a RecQ helicase similar to WRN, can bind to the HPV16 genome at the TRF binding site and these interactions contribute to the regulation of the viral genome copy number. Through both bacterial and mammalian expressed proteins, we demonstrated that TRF2 interracts with HPV16 E2, and furthermore that E2 interacts with other shelterin complex proteins. We utilized immunofluorescence to show that E2 interacts with TRF2 in cells outside of E1E2 replication foci. Through site-directed mutagenesis, we demonstrated a copy-number effect was induced when TRF2 binding sites were removed from previously stable plasmids replicating in both yeast and mammalian cells. These results lead us to suggest that TRF2 plays an important role in HPV16 plasmid stability.

MATERIALS AND METHODS

Plasmid Constructs Several yeast constructs were created using similar methods. The negative control plasmid, pPA104 (ARS-, CEN-), was created by cloning the Ura3 gene into the Sal I restriction site of pPuc19. The original HPV16-Ura3 containing plasmid, pPA103 was described in Angeletti et al. (2002) (158). The pPA111 plasmid contains an ARS+ CEN- backbone and a Trp marker which also has the L2 fragment (nt 4538-5072) of HPV16 shown to have maintenance function (described as pPA94;mtc2 library isolate in the original publication) (158). The pPA113 plasmid (pΔyac CEN- L1; nt 6150-6950) was originally described as the pPA94:mtc3 library isolate. pPA103-2 was constructed by performing restriction digest with Spe1 to Brs 361 in order to remove the early region and most of the long control region from PA103 (nucleotides 1462-4337), leaving the late region to the LCR intact (4338-56).

The predicted TRF binding sites were designated A, B, and C by the order of their distributions in HPV16 genome. pPA111 (Mtc2), containing the HPV16 L2 ORF, had the TRF site identified as TRF A. pPA113 (Mtc3), containing the L1 ORF, contained the TRF B and TRF C sites. pPA103-2 was further subjected to site-directed mutagenesis to ablate TRF B and C binding sites.

Plasmid 2380 contains wild-type EBV OriP cloned in pPUR and was a generous gift from Dr. Paul Lambert (University of Wisconsin-Madison). Plasmid 2380.1 was constructed as a derivative of 2380, by inserting the AflIII fragment from pEGFP C1 that contains the expression cassette of enhanced green fluorescent protein (EGFP) gene into the EcoRI site of 2380. The AflIII fragment insert and the EcoRI cut 2380 plasmid were blunt-ended by treatment with DNA polymerase I (klenow) prior to ligation.

Expression of EGFP is driven under the control of the cytomegalovirus (CMV) promoter. Plasmid 2380.2 has a 914 bp deletion between MluI and XhoI from the plasmid 2380.1, resulting in the removal of the family of repeat (FR). A subgenomic fragment of the HPV16 late region containing 3 predicted TRF binding sites (A-C) was then inserted in place of the FR to generate the 2380.5 plasmid. Those individual fragments were cloned into 2380.2 using multiple linkers indicated in Table 4.1 to accommodate incompatible ends at the insertion sites.

To generate TRF mutants, the L1L2GFP construct containing the HPV16 late region that harbors three predicted TRF binding sites was modified by site-directed mutagenesis (Stratagene Quickchange Kit, Stratagene) with 6 bp substitutions converting the site(s) to the MluI restriction site. Initially, sites A, B, and C in plasmid 2380.5 were individually altered to make TRF mutation mutants, referred to as 2380.5ΔA, 2380.5ΔB and 2380.5ΔC. These single-site mutants were then further mutated to obtain double- and triple-binding-site mutation mutants. Constructs with two TRF binding sites changed to MluI were named as 2380.5ΔAB, 2380.5ΔAC, and 2380.5ΔBC, according to the sites which were modified. The triple-binding site mutation mutant was denoted as 2380.5ΔABC.

The E2-His construct was created by cloning the HPV16 E2 ORF into the Bam-HindIII sites of pQE-9 6xHis tag vector and was a generous gift from the laboratory of Lawrence Banks (ICGEB, Trieste Italy). Expression constructs of TRF1, TRF2, Rap1, Tin2, and Pot1 as well as the GST expression empty vector were gifts from Dr. Paul Lieberman's laboratory and were created by PCR amplification of the gene and cloning into pGEX-2T vector (340).

The E2-Flag was generated by cloning the HPV16 genome into the pCMV-Tag2 vector between the BamH1 and EcoR1 restriction sites.

Yeast Strains The haploid yeast strain YPH500 (*MATα ura3-52 lys2-801_amber ade2-101_ochre trp1-Δ63*) was used for all described experiments. Yeast was grown on YNB minimal media omitting uracil (Ura) at 30 °C for all transformation and Southern experiments, while YPD complete media was utilized additionally to examine plasmid stability and loss rate. The EZ Yeast Transformation kit (Zymo Research, Orange, Calif.) was used for transformation with the experimental plasmids according to the kit protocol and the Zymoprep kit (Zymo Research, Orange, Calif.) was used for yeast plasmid minipreps according to manufacturer specifications.

Yeast Plasmid DNA Replication Assay Two-hundred ng of each plasmid was transformed into YPH500 yeast, plated on Ura- selective agar, and incubated for 3 days at 30°C. Plates were then scored for number of colonies formed and restreaked on selective solid agar, grown for an additional 3 days, and inoculated into 5mL of Ura- liquid media and allowed to grow for approximately 40-50 cell generations. OD_{600} was recorded for each sample to obtain approximate number of cell equivalents per milliliter, and samples were diluted in order to equalize the number of cells, thus allowing an accurate determination of copy number by Southern Blot. Low molecular weight DNA was isolated from liquid culture as described above. The DNA was then loaded onto a 1% agarose gel, electrophoresed, transferred to nitrocellulose, and probed with ^{32}P-radiolabeled pPA104. Radiolabeling was performed using the Amersham Rediprime II kit (GE Healthcare UK, Buckinghamshire) in accordance with the manufacturer's instructions.

Mammalian Cell culture and transfection methods Human embryonic kidney 293 cells and two genetic variants stably expressing Epstein Barr nuclear antigen 1, EBNA1, (293E) or the large-T antigen (293T), fibroblast cell line NIH3T3, cervical cancer cell line HeLa, and spontaneously transformed human keratinocyte HaCaT cells were grown in Dulbecco's Modified Eagle's Medium (DMEM, Gibco, USA) supplemented with 10% heat-inactivated fetal bovine serum (FBS, Atlas, USA) and 1 mM sodium pyruvate. Primary cells, neonatal human foreskin keratinocyte (NHFK), human foreskin keratinocyte (HFK) and their derivative cells, HPV16 transformed NHFK (HPV16NHFK) and hTERT-immortalized HFK (hTERT-HFK) were maintained in Keratinocyte Growth Medium-2, KGM-2, supplemented with KGM-2 Bullet Kit (Lonza, USA) and 0.075 mM calcium chloride. All cells were cultured at 37°C in a fully humidified atmosphere of 5%CO_2.

Primary keratinocytes were transfected using Primefect (Lonza, USA). Other cell lines were transfected with Dreamfect Gold according to the manufacturer's protocol. Briefly, cells were plated the day before transfecting at density of 4 x 10^5 per 60 mm dish or 2 x 10^6 per 100 mm dish. Cells were exposed to 2 μg (60 mm dish) or 5 μg (100 mm dish) of DNA in DMEM containing a ratio of 4:1 Dreamfect (μl):DNA (μg) or 10:1 Primefect (μl):DNA (μg).

Nuclear extracts Nuclear extracts were prepared at 4°C by extraction of nuclei with high salt buffer by the following method. Briefly, cells were lysed with NP40 lysis buffer (50 mM NaCl, 10 mM HEPES pH 8.0, 500 mM Sucrose, 1 mM EDTA and 0.5% NP40). Cell nuclei were washed afterward with low salt buffer containing 50 mM NaCl and then extracted with high salt buffer (350 mM NaCl, 10 mM HEPES pH 8.0, 25% glycerol, 0.1

mM EDTA). All buffers were freshly added with 7 mM mercaptoethanol and 1X protease inhibitor cocktail III (Calbiochem) before used.

Nuclear pellet was proceeded with further purification step by sonication in nuclear pallet solubilization (NPS) buffer (10 mM Tris-HCl, pH 8.0, 140 mM NaCl, 1 mM dithiothreitol (DTT), protease inhibitor cocktail III, 1% Triton X-100, 0.1% NaDeoxycholate). Soluble nuclear pellet fraction was collected after centrifugation at 18000 rpm for 30 min. Typically, 1 x 10^7 yielded 100 µl of nuclear extract with 7 mg of protein per ml and 10 µl of soluble nuclear pallet fraction with 10 mg of protein per ml according to the Micro BCA Protein assay kit (Pierce).

In vitro chromatin immunoprecipitation assay (ChIP) One µg of DNA was packed with chromatin by Chromatin Assembly Kit (Active motif) prior to subjection to chromatin immunoprecipiation (ChIP). The ChIP assay was conducted in vitro as described (295) with the following modifications. Briefly, 500 ng of chromatinized pEF399 or the L1L2Sph fragment was incubated with a mixture of 100 µl of nuclear extract and 100 µl of soluble nuclear pellet fraction at 30°C for 30 min. After the incubation, CaCl2 was added to a final concentration of 3 mM. The sample was digested with 5 units of micrococcal nuclease (MNase) (USB, Affymetrix) at room temperature for 15 min before the reaction was terminated by adding a premixed solution containing 50 mM EDTA and 1% Sarkosyl. Formaldehyde was subsequently added for cross-linking and the reaction was allowed to proceed for another 10 min. Then, glycine was added to a final concentration of 0.125 M and the mixture was left on ice for 5 min to stop the cross-linking activity. The reaction was diluted 1:10 with ChIP dilution buffer (16.7 mM Tris-HCl, pH8.0, 167 mM NaCl, 1.2 mM EDTA, 1.1% Triton X-100, 0.01% SDS). A 1/10 volume aliquot was taken as input control and the remaining solution was precleared with Dynabeads protein G (Invitrogen) which was preblocked with 0.5 mg/ml sonicated

bacterial DNA and 1 mg/ml BSA. After 1 h incubation at 4°C, the precleared supernatant was then equally divided and incubated separately in the presence or in the absence of antibodies at 4°C overnight with rotation. Preblocked Dynabeads protein G were then added and the mixture was incubated at 4°C for 1 h to pull down immunoprecipitate complexes. The beads were washed stepwise once each with the following buffers; low salt immune complex wash buffer (20 mM Tris-HCl, pH 8.1, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS); high salt immune complex wash buffer (20 mM Tris-HCl, pH 8.1, 500 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS); LiCl immune complex wash buffer (10 mM Tris-HCl, pH 8.1, 0.25 mM LiCl, 1 mM EDTA, 1% deoxycholic acid, 1% IGEPAL-CA630) and TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA). The DNA-protein complexes were then eluted twice with 250 µl of elution buffer (0.1 M NaHCO3, 1% SDS). Twenty µl of 5 M NaCl were added to input control and the eluted supernatant containing immunoprecipitate complexes and incubated at 65°C overnight to reverse cross-linking. Immunoprecipitated DNA was then purified with phenol:chloroform:isoamyl alcohol (25:24:1), chloroform:isoamyl alcohol (24:1) and precipitated with isopropanol at room temperature for 20 min. After precipitation, DNA was washed once with 70% ethanol, vacuum dried and resuspended in 20 µl of water. The concentration of DNA was measured using nanodrop spectrometer at 260 nm. PCR was performed with primers (Table 1) to individually amplify each predicted TRF binding site. The PCR products were resolved on a 1.5% agarose gel and visualized after ethidium bromide staining. The sizes of PCR products amplified by different pair of primers were indicated in Table 4.1.

Mammalian Cell Long-term Replication Assays Approximately 4 x 10^5 293E cells were transfected with 2 µg of plamids using Dreamfect Gold reagent in medium containing 2% FBS. The media was changed at 4 h posttranfection. On the next day,

the transfected cells were cultured in the presence of 1 μg/ml of puromycin for 4 days to favor the growth of cells containing the plasmids that are resistant to puromycin and select out untransfected cells. After selection, cells were grown for another 14 days. During the experimental period, cells were maintained and passaged in media containing 250 μg/ml G418 to sustain the expression of EBNA1 protein for replication of the plasmids harboring DS sequences. Equal amount of DNA isolated by Hirt extraction at days 0 and 14 after removal of puromycin was subjected to Southern blotting using EGFP as a probe. PhosphorImager analysis was used to quantify plasmid recovery. The quantity of plasmids was converted to copy number by computing with respect to copy number controls that were loaded alongside.

Western blotting Primary antibodies to TRF1 (GeneTex), TRF2 (GeneTex, Santa Cruz), POT1 (Abcam), hRap1 (Lifespan Bioscience), hTERT (Santa Cruz Biotechnology), Tin2 (Santa Cruz Biotechnology), TPP1 (R & D Systems), WRN (Novus Biologicals), BLM (Santa Cruz Biotechnology), histone H3 (Cell Signaling Technology), Flag (Sigma, Santa Cruz), 6xHis (Pierce), and GST (Pierce), were purchased and used according to manufacturer's protocols. Basically, 30 μg nuclear extract were separated through 10% SDS-PAGE gels. Proteins were transferred to Immobilon-P membrane (Millipore). The membrane was blocked for 1 h with Tris-buffered saline/Tween (0.02 M Tris-HCl, pH 7.6, 0.136 M NaCl, 0.1% Tween 20) (TBS-T) containing 5% nonfat dry milk and incubated with primary antibody at 4°C overnight. Primary antibodies were diluted 1:1,000 or 1:2,000 in the blocking solution (TBS-T plus 5% dry milk). Blots were washed three times for 10 min each with TBS-T, incubated with appropriate secondary antibody conjugated with horseradish peroxidase, developed with Amersham ECL reagents and then exposed to X-ray films.

FarWestern blots DH5α bacteria were transformed with His-tagged E2 and GST-tagged telomere maintenance protein expression constructs and grown to an OD600 of .6-.8 in 100mL of LB media with Ampicillin. Protein expression was induced by addition of 1mM IPTG for 1 hour. Cells were then spun down and collected in Sodium Chloride-Tris-EDTA (STE)/lysozyme buffer and incubated on ice for 15 minutes prior to addition of protease inhibitor cocktail and N-Lauroylsarcosine and sonicated on ice until lysate became clear. Cell debris was spun out in a centrifuge at 4°C.

Protein extracts containing His-E2 were run on a 7% polyacrylamide gel and transferred to nitrocellulose membrane as-per standard Western Blotting procedure. Membranes were then probed with either a primary anti-E2 antibody or a solution containing a GST-tagged telomere maintenance/repair protein (TRF1, TRF2, Pot1, as well as GST alone and L2-GST as negative and positive controls) in 5% milk with proteinase inhibitors. These were incubated at 4°C overnight, washed three times in TBS-T, then probed with anti-GST antibody as-per standard Western Blotting protocol.

GST-His Pulldown Bacterial protein extracts were generated in a similar manner to Far Western blot. E2-His protein extracts were then exposed to Ni-NTA beads and tumbled at room temperature for 1 hour to bind His tagged protein. Beads were washed with PBS three times before being added to bacterial protein extracts containing GST-tagged telomere maintenance proteins. The slurries were incubated for an additional hour at room temperature before beads were collected by centrifugation and washed repeatedly with PBS until wash eluates contained the same amount of protein as background when read using a Nanodrop ND-1000 Spectrophotometer device. Protein was then eluted from the beads by boiling in 1x Laemli buffer and analyzed by western blot. Westerns were probed with anti-GST antibody.

Co-Immunoprecipitation Assay Either 293 or NIH3T3 cells were transfected with E2-Flag or Flag empty vector. Cells were incubated for either 2 days with 293 cells or 4 days with NIH3T3 cells prior to collection of protein lysate with lysis buffer (150mM NaCl, 10% NP-40, 50mM Tris pH8.0). Either Santa Cruz anti-TRF2 antibody produced in goat or Invitrogen anti-Flag antibody produced in rabbit were coupled to Dynabeads Protien G. Collected protein lysates were pre-cleared with Dynabeads bound with anti-goat or anti-rabbit antibody and then incubated with the cells and then incubated with antibody-bead complexes for 1 hour prior to pull-down. Proteins were directly eluted into Laemmli buffer and analyzed by Western Blot as previously described.

Cellular Colocalization Assay 293 cells were plated into 6 well plates with sterile 1 dot cover slips placed in the bottom of the wells. 24 hours later they were transfected with plasmid as per previous protocol and allowed to grow for an additional 24 hours to ensure proper gene expression. Cover slips were then removed, washed in PBS-T, fixed in 4% paraformaldehyde, washed, permeabilized with .05% Triton-X 100, washed, and blocked with 3% BSA. Once the cells were blocked, cover slips were placed in a humidification chamber and incubated for 1 hour with anti-Flag antibody and anti-TRF2 antibody in PBS-T with 1%BSA. After another wash, secondary Alexa fluorochromes for anti-Mouse Alexa 488 and anti-Rabbit Alexa 633 Far Red for 1 hour. Cells were finally washed in 1:5000 dilluted Dapi stain in PBS-T, washed again, and then mounted with Electron Microscopy Sciences Fluorogel with Tris Base. Slides were visualized using a Confocal Laser Scanning microscope with Fluoview 500 confocal system on an Olympus IX81 Inverted Scope. Images came from a 60x Oil Immersion lens and visualized with wavelengths of 405, 488, and 633 nm.

RESULTS

Expression profile of telomeric DNA binding proteins and related proteins in various cells

hTERT is up-regulated by the E6 protein in high risk HPV infected cells (26). Subsequent elevated expression of hTERT in these cells imparts a tumorigenic phenotype, presumably through the maintenance of telomere length. The regulation of telomere integrity involves proteins associated with the telomeric repeat unit (TTAGGGTTA) at the chromosome ends. However, expression levels of telomere-related proteins within the context of HPV infection have not been examined.

Therefore, expression levels of telomere binding proteins and associated partners in keratinocytes that were harboring the HPV16 genome or that were transformed by the hTERT gene were compared to parental normal cells. The HPV16 harboring keratinocytes were established by transfection of primary foreskin keratinocytes, NHFK, with the linearized HPV16 genome. The transfected cells that went through crisis (more than 30 passages) and became immortal were selected and referred to as HPV16 immortalized cells (HPV16-NHFK), with HPV16 presence confirmed by Southern Blot. The hTERT-immortalized keratinocytes were a gift from Dr. Al Klingelhutz and are shown with comparison to the parental, non-immortalized HFK cell line. With Western blot analysis, we found that TRF2 was up-regulated in HPV16- and hTERT- immortalized cells compared to their parental normal cells (Figure 4.2). No notable difference in TIN2 and POT1 expression levels in transformed cells were observed, whereas RAP1 was down-regulated in both immortalized cells to differing extents and WRN was elevated in hTERT-immortalized cells.

Interaction of telomere-related proteins with predicted TRF binding sites in the HPV16 genome

Telomere binding factors are comprised of six core proteins such as TRF1, TRF2, POT1, RAP1, TIN2, and TPP1 (279). If TRF binding sites found in the late region of HPV16 recruit shelterin and/or DNA repair factors to be utilized by the viral genome, it is reasonable that these proteins bind to these sites by ChIP assay. For purposes of this study, TRF2, Pot1, Tin2, and the telomere associated BLM helicase were examined for binding to these sites using the ChIP Assay. The results showed that all of the factors bound with varying affinity to the sequences of telomere nonamer sequences from HPV16 (Fig 4.3). Non-specific rabbit antibody and ribosome L30 PCR primer set was used for a negative control to demonstrate that no non-specific chromatin pull-down was occurring.

Mutation of TRF Binding Sites in HPV Late Region Containing Plasmids

Further studies were performed to investigate what role TRF binding sites from the late region of HPV16 played in stable plasmid replication and maintenance. As such, site-directed mutagenesis was used to alter the nucleotide sequence of these binding sites and replace them with an MluI restriction site to prevent protein binding and for identification of successful mutagenesis. Plasmid pPA111 (mtc 2) contains nucleotide 4538-5072 of the L2 open reading frame, a yeast Autonomous Replicating Sequence (ARS) but no centromere (CEN), and the HPV16 TRF binding site labeled as site A. Upon introducing the mutation into pPA111 and transforming YPH500 yeast, a complete loss of successful long-term replication of plasmid as determined by nutritional selection on –Ura media was introduced (Figure 4.4A).

The pPA113 (mtc3) plasmid contains nucleotides 6151-6951 of the HPV16 L1 ORF. It contains the B and C TRF binding sites, and similarly has a yeast ARS sequence but no CEN. EMSA analysis was used to confirm that knockout of the TRF sites successfully abrogated gel shift with nuclear extract protein similar to site A (Figure 4.4B). Unlike pPA111, however, mutation of the TRF binding sites in the pPA113 plasmid did not result in complete loss of stable replication in yeast (Figure 4.4B). Southern analysis of genomes isolated from yeast identified that the pPA113 mutant plasmids were successfully replicating episomally with a slight shift in copy number (Figure 4.4B).

Given the result from pPA111 and the stability results from pPA113, it was desirable to explore whether the TRF binding site maintenance effect represents a separate mechanism outside of or in addition to segregation. As such, the question was posed to examine what effect could be observed with deletion of the TRF binding sites in a plasmid that possesses neither ARS nor CEN sequences and is completely reliant on HPV16 late-region signals for plasmid stability. Towards this end, plasmid pPA103-2, a truncation mutant of HPV16 with the early genes removed by restriction digest with Spe1 to Brs361 (nucleotides 1462-4337), was used to observe what effect mutation of the TRF binding sites would have on plasmid replication and maintenance in yeast. Quikchange XL site-directed mutagenesis was performed to delete B, C, and BC from pPA103-2 and YPH500 yeast were transfected with the resultant mutant plasmids. pPA103-2ΔB transformed yeast replicate robustly and with similar growth and stability characteristics to pPA103-2. However, removal of the C binding site resulted in a plasmid which could support only limited yeast growth, presumably due to an observed dramatic reduction in plasmid copy number as observed by Southern blot of the

extracted circular genomes (Figure 4.5A and 4.5B). The double mutant pPA103-2ΔBC was incapable of replicating stably in yeast and was thus not analyzed by Southern.

Influence of TRF binding sites on HPV16 DNA maintenance in Human Cells

Several studies in EBV have proven that nonamers which resemble TRF binding sequences contribute to the maintenance function of OriP containing plasmids (12, 88). TRF proteins and associated partners can interact with the predicted TRF binding sites in the HPV16 subgenomic region we had mapped for maintenance function (Figure 4.2). To investigate the role of TRF binding sites in the HPV16 genome maintenance in mammalian cells, the well-established EBV replication and maintenance system was used. The minimal replicator for EBV episomal DNA maintenance has been identified in OriP that contains two distinct *cis*-elements responsible for separate processes: replication and partitioning. DS (dyad symmetry) confers EBNA1-dependent replication whereas FR (Family of Repeats) mediates partitioning. For HPV, it was speculated that the late region containing TRF binding sites is required for efficient long-term maintenance. To directly test this, replication and partitioning elements were uncoupled by constructing hybrid origins which contain the DS element of EBV OriP and various TRF mutants in the background of pPur (Fig 4.6A). Plasmid containing hybrid origins should replicate in an EBNA-1-dependent fashion and they have successfully been utilized to verify *cis*-elements of other DNA tumor viruses (271, 272).

293E cells stably expressing EBNA1 were transfected with each hybrid origin and either 2380.1 (intact EBV OriP) or 2380.2 (EBV OriP without FR) as controls. To rule out the potential effects of neighboring sequence and plasmid size on DNA retention and stability, a set of TRF mutants were created by either singly or multiply mutating

TRF binding sites in the plasmid 2380.5 (Fig 4.6A) similar to those created in yeast plasmids. After transfection, cells were grown in the presence of puromycin to select for plasmid containing cells. After 4 days under a selective condition, outgrowing colonies were released from drug selection and grown in the absence of the drug for another 2 weeks. DNAs were Hirt extracted at the first and fourteen days after removal of the drug. Hirt-extracted DNAs were then subjected to DpnI digestion and DpnI-resistant species were detected by Southern blot with a radiolabeled probe specific for GFP sequence (Fig.4.6B). All experiments were performed in triplicate.

All TRF mutants had influence on copy number, as an increased plasmid copy number was observed at day 14 compared to the control which contains intact EBV OriP (Fig 4.6B). However, the effect on the change in copy number by most TRF mutants seems marginal compared to the wild-type DNA harboring intact TRF binding sites. The greatest copy number increases were observed in plasmids with the A and B sites mutated singly and the BC double mutant. A copy number change was observed in a mutant lacking TRF binding sites B and C (about 10 times higher). In summary, these results indicated that mutating two out of three of the TRF binding sites reduced plasmid copy number.

HPV16 E2 Interacts with TRF2 and Other Shelterin Components

Whether or not E2 is expressed during maintenance replication, at some point during the HPV16 lifecycle E2 will be expressed and potentially interact with E2. OriP plasmids, which contain the EBV latent origin of replication, rely on an interaction between EBNA1, a functional homologue of HPV E2, and TRF2 for DNA replication and plasmid segregation. In order to investigate whether HPV16 E2 interacted with TRF2, Far Western blot was used. Purified E2 protein from transformed DH5α bacteria was run on a polyacrylamide gel and transferred to a nitrocellulose membrane prior to probing

with purified individual shelterin proteins (TRF1, TRF2, Pot1) fused with a GST tag. The resulting protein-protein complexes were probed with an anti-GST antibody. As shown in Figure (4.7A), E2 was specifically bound by all three shelterin complex proteins as well as the positive control protein L2, but not the negative control GST alone.

To further explore this, HPV16 E2 protein fused to a 6xHistidine tag was bound to Ni-NTA beads and used to perform a pull-down assay to determine if shelterin complex proteins were capable of interacting with E2. Bacterial lysates were harvested containing GST-tagged shelterin component proteins, cleared with Ni-NTA beads alone, and then mixed with Ni-NTA beads bound with E2. After extensive washing to remove any non-specific binding, bound protein-protein complexes were analyzed by Western Blot to identify proteins capable of binding to HPV16E2. We observed that E2 was capable of pulling down GST-TRF1, TRF2, Rap1, Pot1, and Tin2, as well as positive control L2, but again was not bound by the negative control GST alone (Figure 4.7B). These results indicate that E2 is capable of interacted with several members of the shelterin complex.

To examine if the interaction occurred in mammalian cells, we transformed both 293 cells as well as NIH3T3 fibroblast cells with a Flag-tagged E2 vector. 3T3 cells were used as a means of increasing total E2 levels, as it is one of few cell lines capable of tolerating long-term overexpression of E2 and could thus be passaged under drug selection after transfection. Additional 3T3 cells were co-transfected with either a GFP vector to check for transfection efficiency or a TRF2 overexpression plasmid. After 2 (for 293 cells) or 4 days (NIH3T3) of protein expression, cells were collected and total protein lysate was harvested. Lysates were precleared with anti-mouse and anti-goat antibody bound to Dynabeads protein G prior to being mixed with beads bound to either anti-Flag antibody to pull-down the E2 protein or anti-TRF2 to pull-down the endogenous

TRF2 or, in the appropriate treatments, the overexpressed TRF2 from 3T3 cells. The pull-down results were then analyzed by Western blot and probed with the counterpart antibody to determine whether TRF2 and E2 interact in vitro in mammalian cell extracts. As shown in Figure 8C and D, this appears to be the case, as the proteins are pulled down very specifically through both antibodies. The combination of all three protein-protein interaction results allows us to conclude with certainty that E2 interacts with TRF2 *in vitro* and seems to be capable of interacting with a number of other shelterin components as well.

Finally, to determine whether this interaction occurs *In vivo,* we performed immunofluorescent co-localization studies. Cells were transfected with expression vector for our Flag tagged E2 protein, a combination of E2 and EE tagged E1, or the Flag vector alone, fixed to glass coverslips with paraformaldehyde, and stained with anti-Flag and anti-TRF2 antibody followed by fluorescent secondary antibody (Alexa 488 and 633.) The results are shown in Fig. 4.8E. E2 appeared to consistently co-localize with TRF2. Previously, the McBride laboratory had shown that HPV16 E2 co-localizes to chromosomes more efficiently when E1 is co-transfected into cells, and that the E1-E2 complex induces localized DNA damage at the interaction site(255), leading to activation of ATM and, among numerous other factors, phosphorylation and re-localization of TRF2. We thus wished to observe whether including E1 with the E2 vector would alter localization of the E2-TRF2 complex, and so co-transfected these factors. As shown in Fig. 8E, the interaction with E2 and TRF2 appears to be separate from the E1-E2 nuclear foci.

DISCUSSION

Through this study, TRF2 and its binding partners within the shelterin complex have been demonstrated to promote long-term maintenance strategy of HPV16. Replication of the viral genome within cells leads to an overall up-regulation of telomere maintenance factors in mammalian cells, and some of these factors are in turn capable of binding spontaneously to nine-base sequences of telomere DNA found in the late region of HPV16. Through the use of site-directed mutagenesis, loss of these nine-base repeats induces increased instability with regards to copy-number and overall replicative success in both mammalian cells as well as our well established *S. cerevisiae* based HPV replication model. E2 protein, which plays a critical role in plasmid maintenance for the virus, interacts with TRF2 as well as other members of the shelterin complex, as shown through both in vitro and in vivo assays. Taken together, these results indicate an important role for TRF2 and its subsequent binding partners in the HPV16 life cycle.

Chromatin immunoprecipitation was used to investigate the interaction of TRF2 and its associated proteins with TRF binding sites in HPV16 DNA and investigate whether TRF2 partners that are associated with TRF2 can bind to the TRF binding sites. TRF2 and other telomere-related proteins such as POT1, TIN2, and BLM helicase are able to bind to TRF binding sites. A similar result has been observed with OriP in EBV. Several studies in EBV have demonstrated that TRF2 binds to the TRF binding sites at DS in OriP and contributes to the replication and plasmid maintenance function of OriP. It is likely, then, that the shelterin end-protection complex, responsible for T and D loop formation and protection from the NHEJ system, rather than the TRF1 dependent complex that binds along the length of the telomere, is responsible for any TRF2 dependent viral-maintenance effects.

Separate studies in high risk HPVs (types 31 and 16) have reported that high risk E6 is necessary for episomal maintenance in primary keratinocytes. While genetic analyses have suggested that p53 degradation activity of E6 is necessary for stable maintenance of high-risk HPV genomes (241), it was not clear whether other activity of E6 in telomerase stimulation would provide additional stability to high-risk HPV genomes. Since elevation of telomerase expression has an impact on expression profiles of shelterin complex proteins as well (TRF1, TRF2, POT1, TIN2, TPP1, RAP1 and RecQ helicases), it is attractive to speculate that E6 facilitates viral DNA maintenance by altering expression of these key factors. It was therefore hypothesized that alteration of telomeric associated protein levels would be observed in keratinocytes that have been immortalized by HPV16 DNA. Various transformed/immortalized cells as well as normal primary keratinocytes were also included for comparison. Compared to normal keratinocytes, TRF1, TRF2, and hTERT were expressed at higher levels in most transformed cells. In general, the levels of these proteins were altered in transformed cells to varying extents. Interestingly, we observed elevated expression of TRF2 in all cells that are either virally or spontaneously transformed.

The two regions of HPV16 that contain maintenance signals in yeast (pPA111 and pPA113) coincided with maintenance functions mapped in mammalian cells. If these sites play a role in stable HPV replication and maintenance in either host cell type, then disrupting the sequence of these sites should lead to a reduction in plasmid stability in both models. As such, site-directed mutagenesis was employed to disrupt TRF2 binding. The YPH500 strain of yeast was then transformed with the plasmids in order to test for loss of stability in terms of forming colonies on –Ura plates, growth rate in liquid media, and reduction of copy number. pPA111 lost the ability to replicate stably in yeast with the removal of the TRF A site. This result is different than what would be expected from a

deficiency in segregation, as budding yeast have a closed mitosis with a mother-cell bias and, as such, loss of proper plasmid segregation would result in retention of the plasmid in the original cell, leading to high copy number and growth of a small number of large colonies on the plate. This phenotype was not observed. The presence of an ARS signal within the plasmid rules out a defect in initiation of replication as well. Given that the average loss rate of plasmids with no maintenance signal in yeast is approximately 25% per cell generation, it is more likely that some post-replication defect has been introduced through loss of the A site, leading to loss of the plasmid and cessation of yeast growth after only a few cell generations, explaining the lack of colony formation.

Mutation of the B and C binding sites in pPA113 did not result in the same dramatic decrease in stability seen in pPA111. YPH500 transformed with the mutant plasmids possessed similar growth characteristics to the unmutated plasmid on solid media. All three mutants replicated episomally. There was an increase in the loss rate of plasmids with the B site removed, but not with C. As such, it is difficult to conclude what effect removal of these sites was having on overall plasmid stability. However, when the same mutations are introduced into PA103-2, a plasmid containing only the late region of HPV16 and no ARS or CEN sequences, the results become more pronounced. Mutation of only the B site did not significantly alter plasmid growth or stability, but loss of C resulted in a significant reduction in growth rate and copy number, and the double mutant was completely unstable, much like pPA111ΔA. The loss of growth from the C mutant resulted from a significant reduction in plasmid copy number, leading to insufficient production of the Ura3 gene product to sustain yeast growth. This strongly suggests that the successful growth of pPA113ΔC is due to the ARS+ vector backbone which the plasmid is cloned into and, thus, that the C site may be playing a role in

successful completion of plasmid replication or prevention of integration into host chromosomes, rather than segregation.

To expand these results to the effect of TRF binding sites on the long-term maintenance of HPV DNA in mammalian cells, a heterologous replicon system that has been developed to study the mechanism of stable replication and plasmid maintenance was used (271, 272). The hybrid origin replicon contains the EBV DS that conveys EBNA1 dependent replication while different viral *cis* elements can be replaced at the partitioning part of the EBV OriP. This replicon allows for examination of the activity of the viral *cis* maintenance elements based on the capability to substitute for function of FR. Using EBV-DS/HPV16-maintenance element hybrid origins that contain various truncated derivatives, we observed that the HPV *cis* maintenance element is able to provide DNA stability and slightly increases plasmid copy number.

Mutation of the TRF binding sites to prevent TRF2 binding in these cells had some effect on plasmid maintenance, but did not necessarily coincide with the results observed in yeast. Notably, the PA103-2ΔBC plasmid which was unable to replicate in yeast induces an increase in copy number in mammalian cells. None of the mammalian cell plasmid mutants induced a reduction in copy number compared to the unmutated plasmid. While the ΔA mutation in pPA111 induced a complete loss of plasmid viability in yeast, it yielded only a moderate copy number increase in human cells. It is apparent, then, that some differences exist in the phenotypes of these mutants between the two systems.

While these results do not correlate exactly between the two models, the significance of the results is that alteration of TRF binding sites leads to changes in plasmid copy number. It is likely that, during evolutionary divergence, mammalian cells

have developed complementary DNA maintenance systems which are masking the results observed in yeast. Variations in result between systems can likely be contributed to the presence of multiple binding sites such as MARS, HMG, TopoII, and CENP-B present in the maintenance element in the late region of HPV16 and, thus, a differential ability to bind and utilize these sites between the two model systems as well as differences in the telomere maintenance factor dependent phenotype between these two organisms. It is also possible that the inclusion of OriP and EBNA1 may have reduced the effect of mutating TRF2 sites in our model. The fact remains, however, that alteration of these sites leads to a change in plasmid copy number and, by inference, an overall reduction in plasmid stability.

TRF2 is also important in the life-cycle of other DNA viruses that possess long-term infection strategies comparable to papillomaviruses. Notably, EBV and KSHV proteins EBNA1 and LANA have been shown to interact with TRF2, and evidence shows that these interactions are important for the maintenance replication of plasmids containing the EBV latent origin, OriP(85, 87, 149). HPV E2 proteins, while not possessing sequence identity, possess similar functions to LANA and EBNA1 and consist of the same unique protein fold as these viral factors. This work demonstrates that HPV16 E2 protein interacts with TRF2 through in vitro and in vivo methods, as well as a number of other factors from the telomere maintenance shelterin complex. Further work is necessary to investigate the functional significance of this interaction, but it suggests that HPV16 may be capable of utilizing TRF2 similarly to EBV.

Efficient maintenance is a result of the sum of a virus's interactions with the infected cells' replication, repair, and segregation machineries, as they have been shown to exhibit compensatory effects for viral genome retention (272). Evidence in BPV showed that viral genomes are segregated nonrandomly into both daughter cells and the

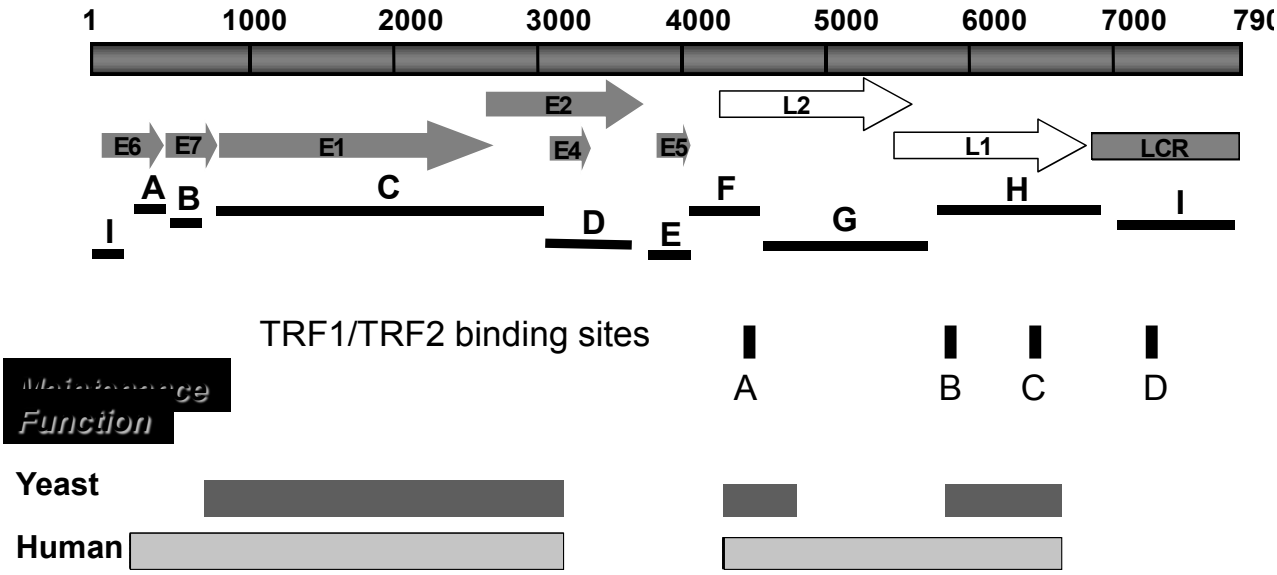
average copy number of viral genomes per infected cell does not fluctuate widely (209). It has been described in several latent DNA viruses such as EBV and KSHV that the viruses adopt similar strategies to maintain their genomes in host cells. These observations imply that a key requirement of papillomavirus stability is regulation of copy number for stable, long-term DNA maintenance in infected cells. The copy number fluctuations observed due to removal of TRF binding sites in this study may reflect a loss of that regulation.

The specific role TRF2 plays in the long-term persistence of HPVs is unclear. While it was initially hypothesized that these binding sites may be important for plasmid segregation, some of our results may implicate an alternate maintenance mechanism. Removal of the binding sites from plasmids with intact origins of replication in both mammalian and yeast plasmid constructs did not result in plasmid behavior consistent with a loss of segregation. The experiments with removal of the B and C binding sites from PA113 in particular do not reflect a loss of plasmid segregation. Rather, the defects could result in overall induction of plasmid instability and copy-number fluctuation through an alternate maintenance mechanism. Recent results from the Lieberman laboratory have indicated that, for EBV, one of TRF2's primary functions is to recruit DNA repair factors responsible for resolving Holiday Junctions to both improve post-replication processing of the viral plasmids as well as improve segregation efficiency(92). Given the observed fluctuations in copy number after induction of TRF binding site mutations, as well as the loss of replicative stability for pPA111ΔA, despite the presence of an ARS sequence, it is possible that similar utilization of the DNA repair systems may be at work here. A number of recent studies have similarly highlighted the importance of the DNA repair systems for papillomaviral replicative success (125, 218,

255), Further work will need to be completed to investigate the possible role TRF2 proteins may play in recruiting and/or utilizing these factors.

Figure 4.1 Telomere Repeats Within the HPV16 Genome A) Predicted telomere-related factor binding sites (A-D) within the HPV16 genome were discovered. Three of these sites (A-C) fall within L1 and L2 regions mapped for maintenance function. The consensus TRF binding site is shown below the map. B) The distribution of predicted TRF binding sites in selected HPV genomes is shown (red boxes). Statistical analysis revealed that greater than 50% of the predicted binding sites were in the late region of HPV genomes (L1 and L2 genes; indicated by blue shading). No identifiable TRF binding sites were detected in the BPV1 genome. **Table 1** lists PCR primers used for mutation of TRF binding sites within subsequent plasmids from later in the work as well as the expected PCR product sizes.

Fig 4.1A



TRF1/TRF2 binding site sequence = TTAGGGTTA

A TTAGGGTTT
B TTAGAGTTA
C TTAAGGTTA
D TTTGGGTTA

Fig 4.1B

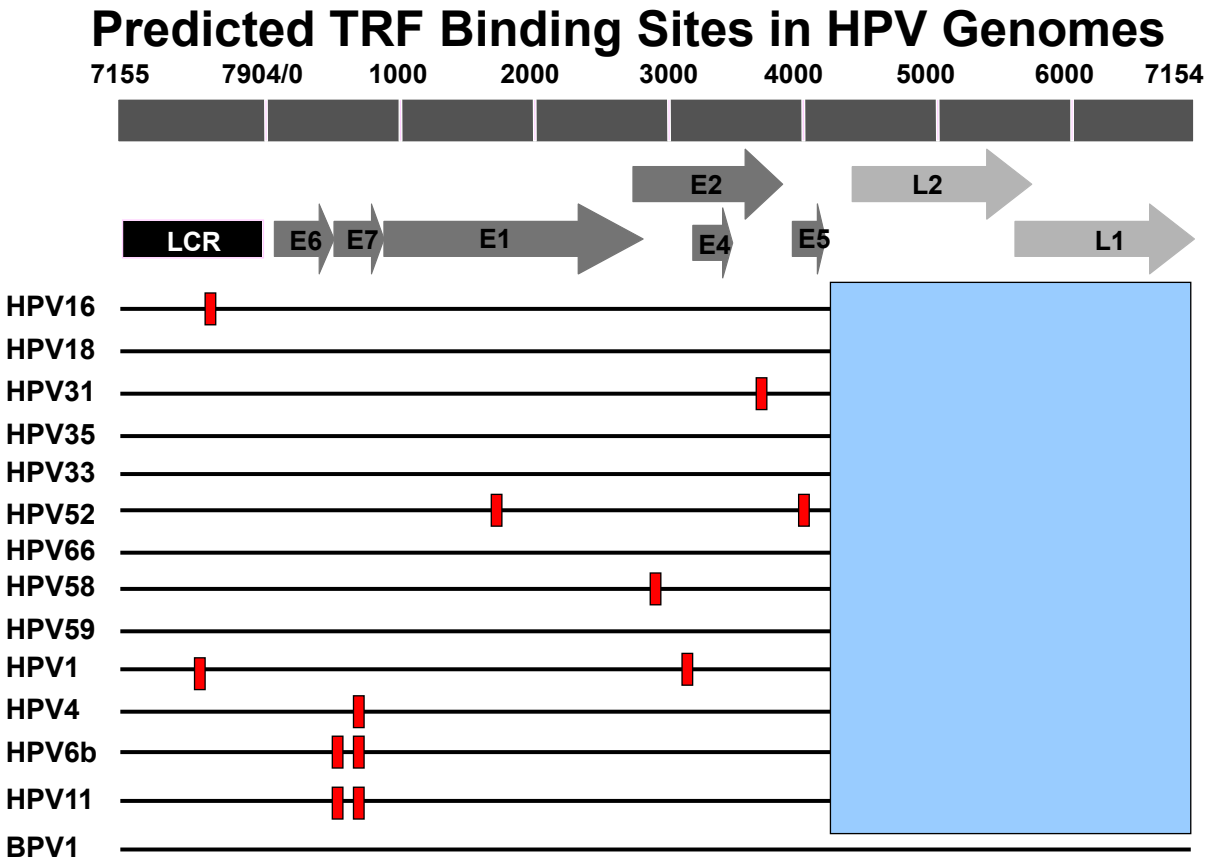


Table 4.1 Primers used to amplify TRF binding consensus sequences.

TRF binding site (nucleotides)	Primer sequences	PCR product sizes (bp)
A (4849-4855)	5'ACCCATCTGTATTGCAGCCTC3' 5'AACTTGTTGTGTTGTGCGAC3'	192
B (6198-6204)	5'ATGCAGCAAATGCAGGTGTG3' 5'CACCTGGATTTACTGCAACATTGG3'	123
C (6778-6784)	5'AGCACAGGGCCACAATAATG3' 5'TCTTCTAGTGTGCCTCCTGG3'	283
D (7863-7869)	5'CTAAGGCCAACTAAATGTCACC3' 5'CGATTTCGGTTACGCCCTTAG3'	186
Neg	5'??3' 5'?3'	200

Figure 4.2 Western blot analysis for levels of telomere-related proteins in normal keratinocytes, those harboring HPV16, or transformed with hTert. Whole cell extracts were prepared from 4 day-old cells using SDS lysis buffer. Equal amount of protein extract were subjected to SDS-PAGE electrophoresis and then hybridized with specific antibody against protein of interest. Expression levels of telomere-related protein (TRF2, TIN2, POT1, RAP1, and WRN) in transformed cell compared to parental normal counterpart. HPV16-NHFK (HPV16 transformed neonatal human foreskin keratinocyte cells), and hTERT-HFK (hTERT immortalized human foreskin keratinocyte cells).

Fig 4.2

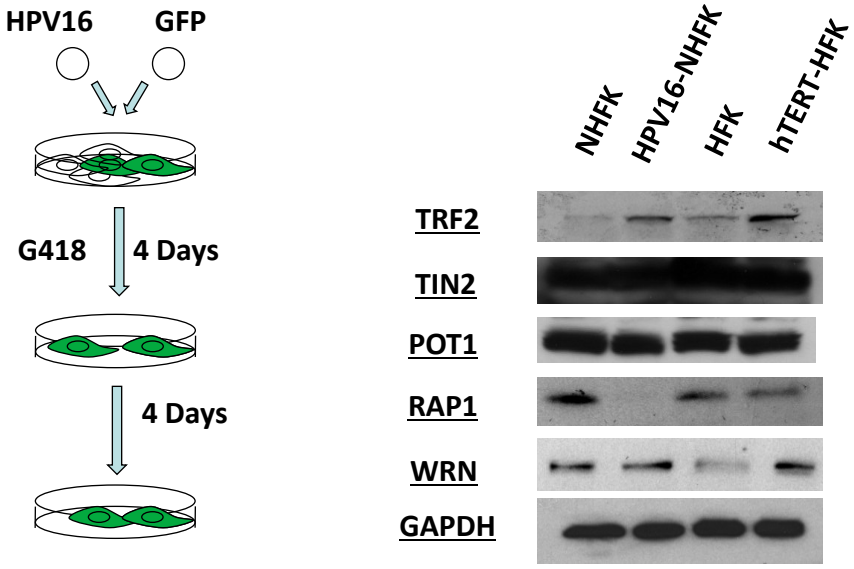


Figure 4.3 Analysis of telomere-related proteins binding to the TRF binding sites in the viral genome by in vitro ChIP assay. ChIP was performed *in vitro* using plasmid DNA with the HPV16 subgenomic fragment containing TRF binding sites incubated with 293 nuclear extract and soluble nuclear pallet. The DNA-protein complexes were then crosslinked and processed as described in the materials and methods section. To determine proteins bound at TRF binding sites, immunoprecipitations were performed using no antibody (No Ab), anti-TRF2, anti-POT1, anti-TIN2, or anti-BLM. A portion of material prior to subjection to immunoprecipitation was saved to serve as a control for the amount of DNA added to each immunoprecipitate (Input). PCRs were conducted using primers specific for the TRF binding site, indicated by A-D as in Fig 4.1 (primer sequences are shown in the Material and Methods section).

Fig 4.3

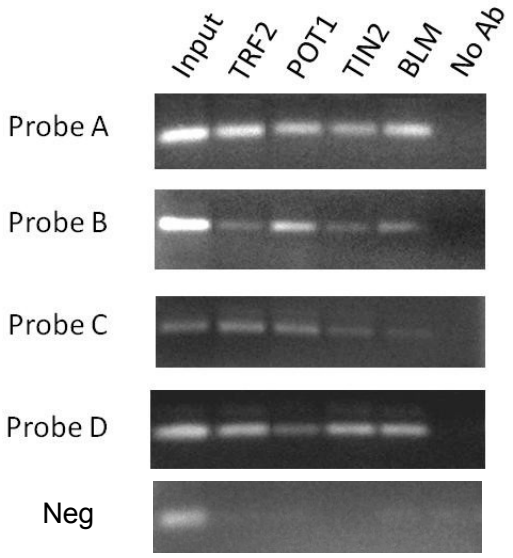
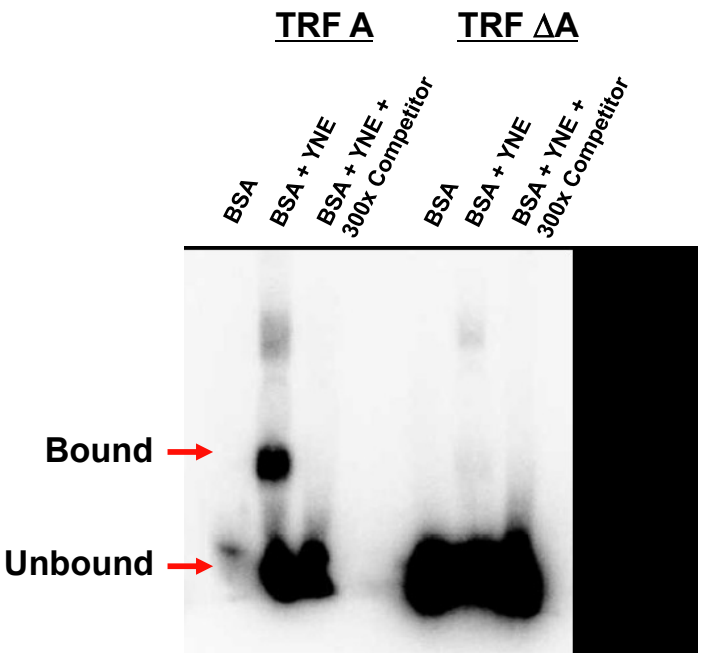


Figure 4.4 Mutation of TRF binding sites disrupt plasmid stability in yeast. (A). A mutation in TRF binding site A was analyzed for binding to telomeric proteins in yeast nuclear extract by gelshift (Left panel). Double stranded competitor TRF binding site was used at 300X molar concentration to demonstrate specificity. The effect of TRF binding site A disruption on episomal maintenance of an ARS^{+} plasmid (pPA111) in yeast is shown in the (Right panel)**.** TRF binding site A is essential for establishment of stable episomes. Viable yeast containing the remaining mutant combinations were analyzed by Southern analysis **(B)**. Mutation of TRF binding site B consistently led to about a 50 percent increase in copy number, whereas the TRF binding site C mutation led to a detectable decrease in copy number. A combination mutant in binding site B and C revealed an intermediate phenotype. Triplicate quantification results are shown in the right panel. EMSA comparing gel shift of WT vs. Mutated TRF B binding site is shown below. A summary of TRF binding site mutant effects on yeast growth and plasmid copy number is shown **(C)**. Each mutation is indicated by an x below the affected binding site. The arrows indicate the orientation of each of the TRF binding sites. The + and – signs provide a qualitative summary of mutant effects.

Fig 4.4A



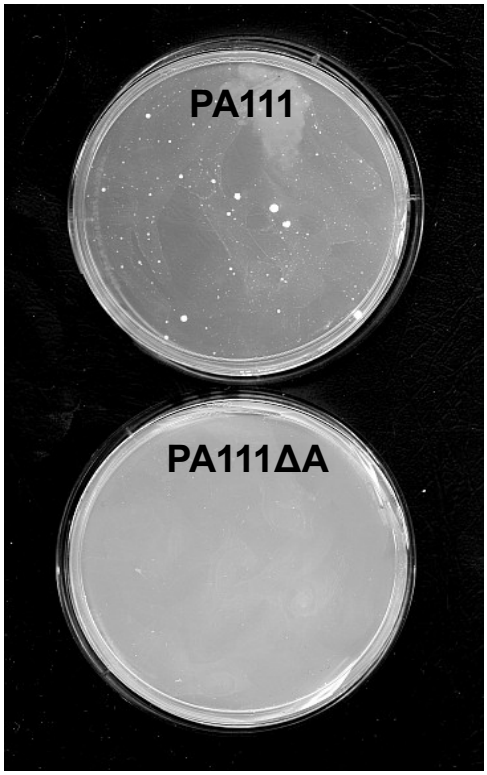
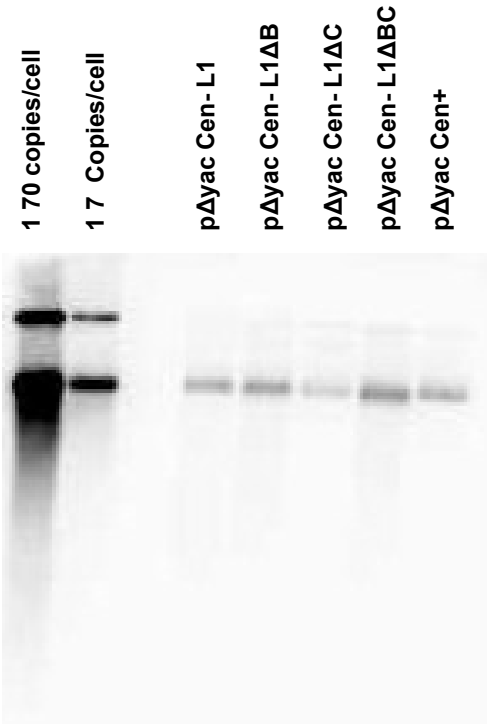
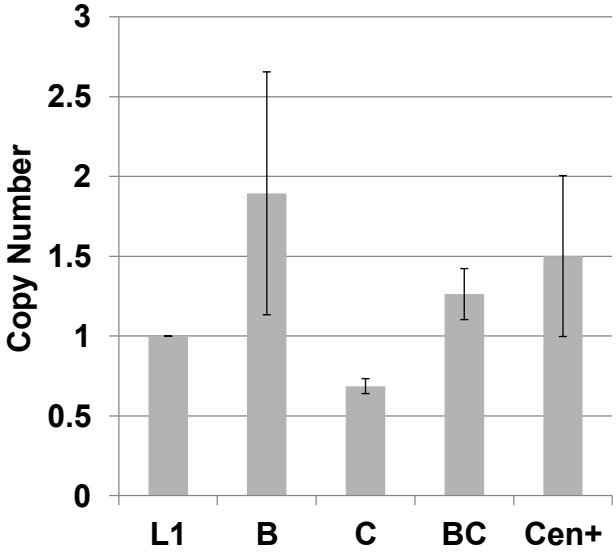


Fig 4.4B





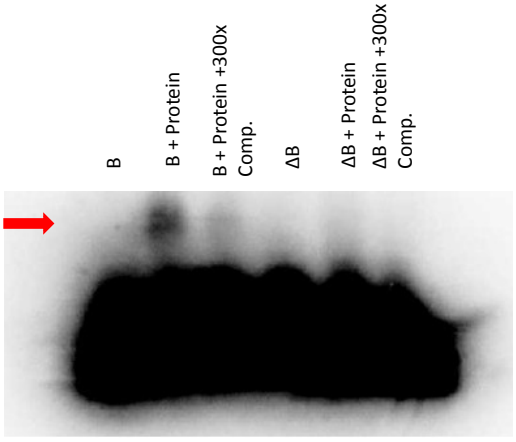


Fig 4.4 C

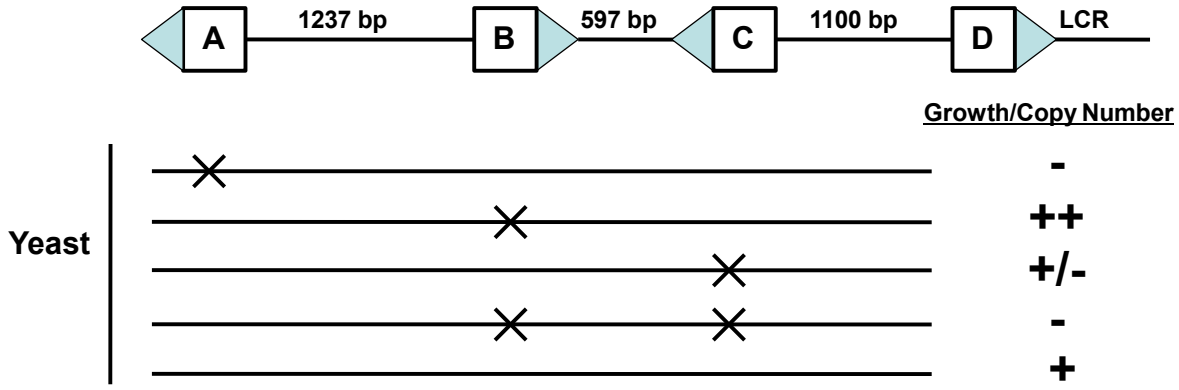


Figure 4.5 Mutation of the TRF binding sites B and C influence growth phenotypes and plasmid copy number. A fragment of the HPV16 genome spanning nt 4334 to 56 was shown to replicate autonomously when linked to a Ura3 selectable marker (pPA103-2) **(A)**. Two predicted TRF binding sites (B and C) were mutated and each of the resultant constructs (pPA103-2ΔB, pPA103-2ΔC, pPA103-2ΔBC), were analyzed for viability and growth effects in yeast when plated on media lacking uracil. Viable mutants were analyzed by Southern blot **(B)**. While pPA103-2ΔB seemed to have a slightly increased copy number, the ΔC mutant had a significantly reduced copy number. ΔBC was incapable of sustaining long-term replication in yeast. DNA is run uncut and, thus, bands for supercoiled will appear to have lower molecular weight while nicked supercoiled will appear higher. Arrow indicates the expected supercoiled molecular weight.

Fig 4.5A

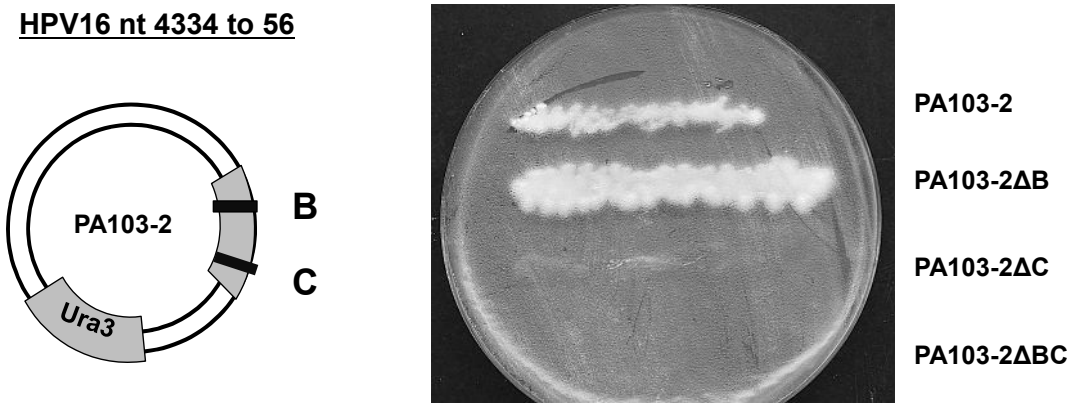


Fig 4.5B

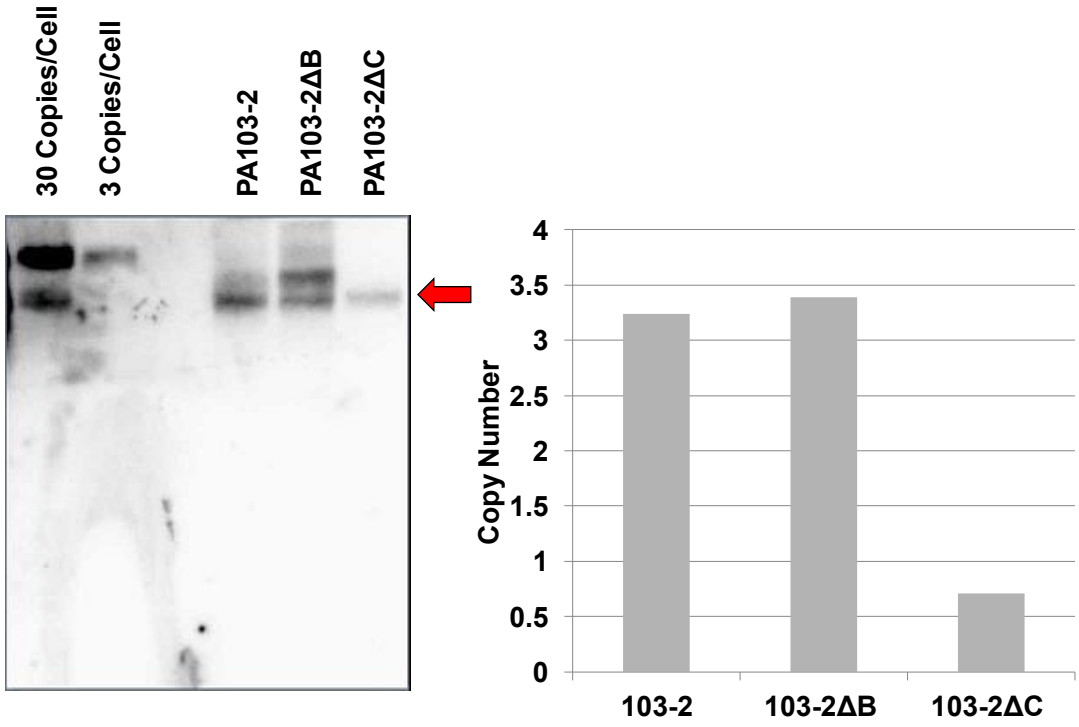


Fig 4.5C

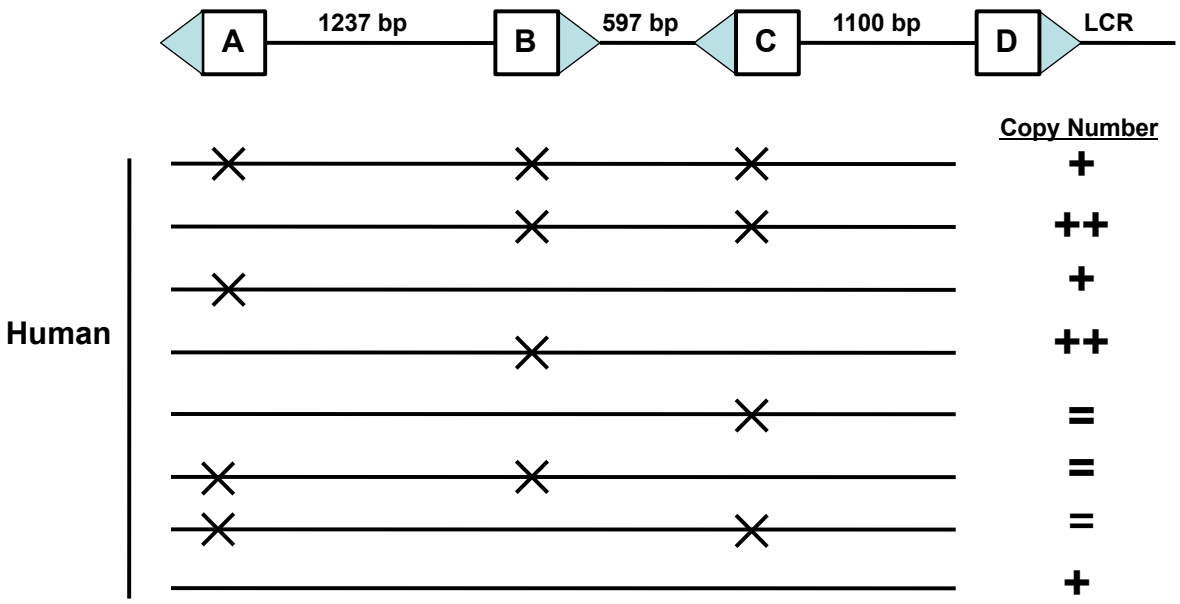


Figure 4.6 Map of constructs containing mutated TRF binding site(s). The plasmid 2380.5 carries the HPV16 sequence (nt. 4466-7154) with three TRF binding sites, dyad symmetry of EBV ori, and an EGFP gene. A series of mutant constructs, derived from 2380.5, contain single, double or triple mutations at TRF binding site as indicated in diagram **(A)**. These constructs were transfected into 293E cells and selected with puromycin for 4 days. Each mutation is indicated by an x below the affected binding site. The arrows indicate the orientation of each of the TRF binding sites. The + and – signs provide a qualitative summary of mutant effects on copy number. **(B)** After release from the drug, the transfected cells were grown under non-selective condition for another 14 days Hirt extraction was performed to collect low molecular weight DNA 14 days (day 14) after cells were released from the drug. Hirt-DNAs were then DpnI digested and subjected to Southern analysis using EGFP gene as a probe. Experiments were performed in triplicate. The band intensity was determined by densitometry. The plasmid stability was shown as relative change in copy number. A summary of TRF binding site mutant effects on copy number of 2380.5 plasmids containing the HPV16 late region sequences

Fig 4.6A

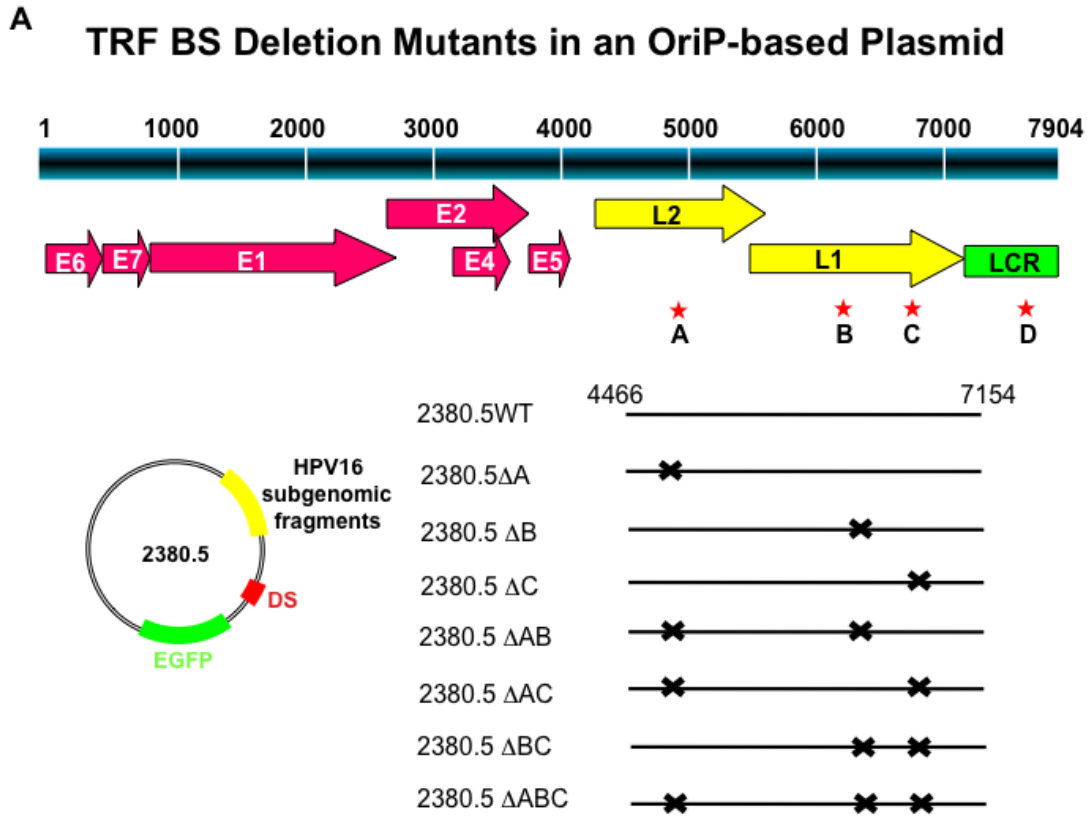


Fig 4.6B

B

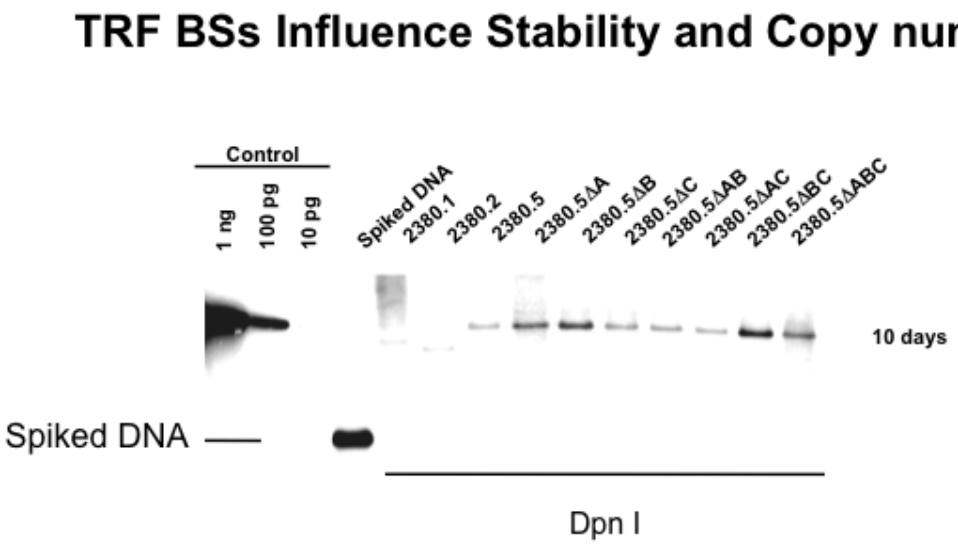


Figure 4.7 E2 interacts with TRF2 and other shelterin components. A) Purified His-tagged E2 or negative bacterial extract was separated on an SDS PAGE gel and transferred to nitrocellulose. Each blot was then probed with either GST-TRF1, GST-TRF2, GST-Pot1 or GST-HPV16L2 as a positive control. GST alone was used as a negative control. Each blot was reacted with primary monoclonal antibody to GST and secondary polyclonal antibody to mouse IgG, followed by development with ECL. An interaction with the purified E2 protein is shown for TRF1, TRF2, Pot1 and for the positive control HPV16 L2. GST alone shows no interaction with E2. B) His-tagged E2 protein was bound to Ni-NTA beads and used for pull-down in lysate containing either GST tagged TRF1, TRF2, Pot1, Rap1, Tin2, HPV16 L2, or GST alone. Resulting protein eluates were analyzed by western blot and probed for either GST (top) or His (bottom).

Fig 4.7A

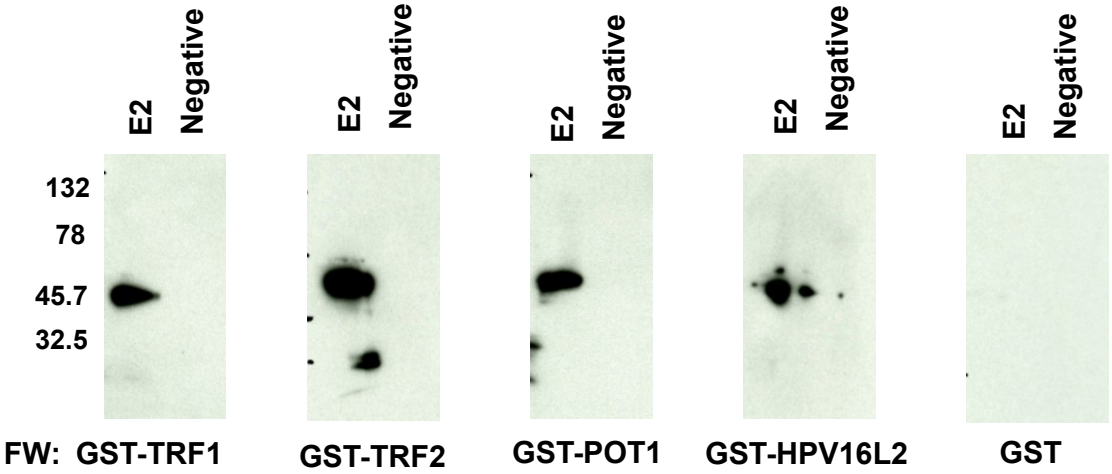


Fig 4.7B

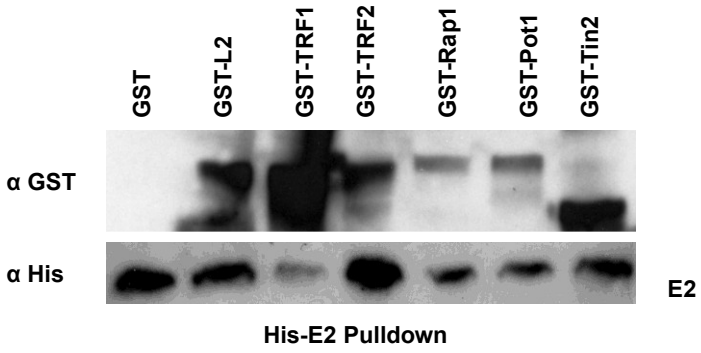


Figure 4.8 E2 TRF2 Interaction in Mammalian Cells A) Western blot showing crude expression of Flag-E2 from expression vectors from 2 transfections in NIH3T3 cells. Reblotting of the same membrane to demonstrate TRF2 levels is shown in B) C) Co-immunoprecipitation of E2-Flag by pull-down of TRF2. NIH3T3 cell lysate from 4 days post-transfection with either E2-Flag or empty vector were mixed with Dynabeads Protein G coupled to TRF2 antibody. Beads were collected, washed, and eluted directly into Laemmli buffer and analyzed by western blot. D) 293 cells were transfected with the same plasmid constructs as C). Lysates were collected and pull-down was performed with Dynabeads bound with anti-Flag antibody. Western blot shown is probed with anti-TRF2. E) 293 cells were transfected with Flag, E2-Flag, or E2 Flag with E1 plasmids plasmid and fixed to glass coverslips prior to treatment with primary and fluorescent secondary antibodies and visualization by confocal microscopy. Flag antibody is visualized in red and endogenous TRF2 is shown in green. E2 co-localizes in diffuse nuclear structures with TRF2 associated with DNA.

Fig 4.8

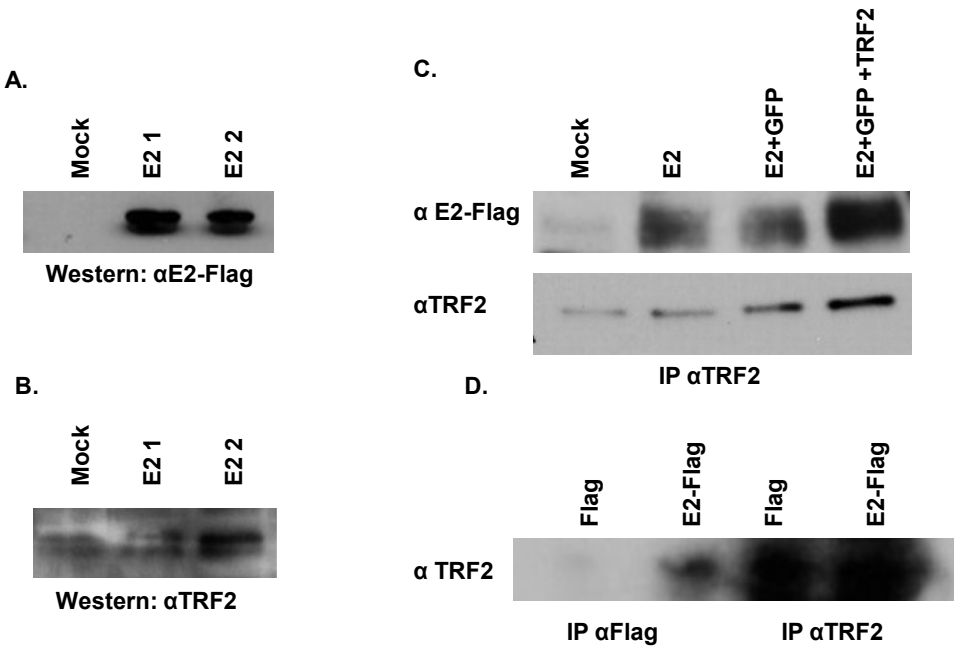
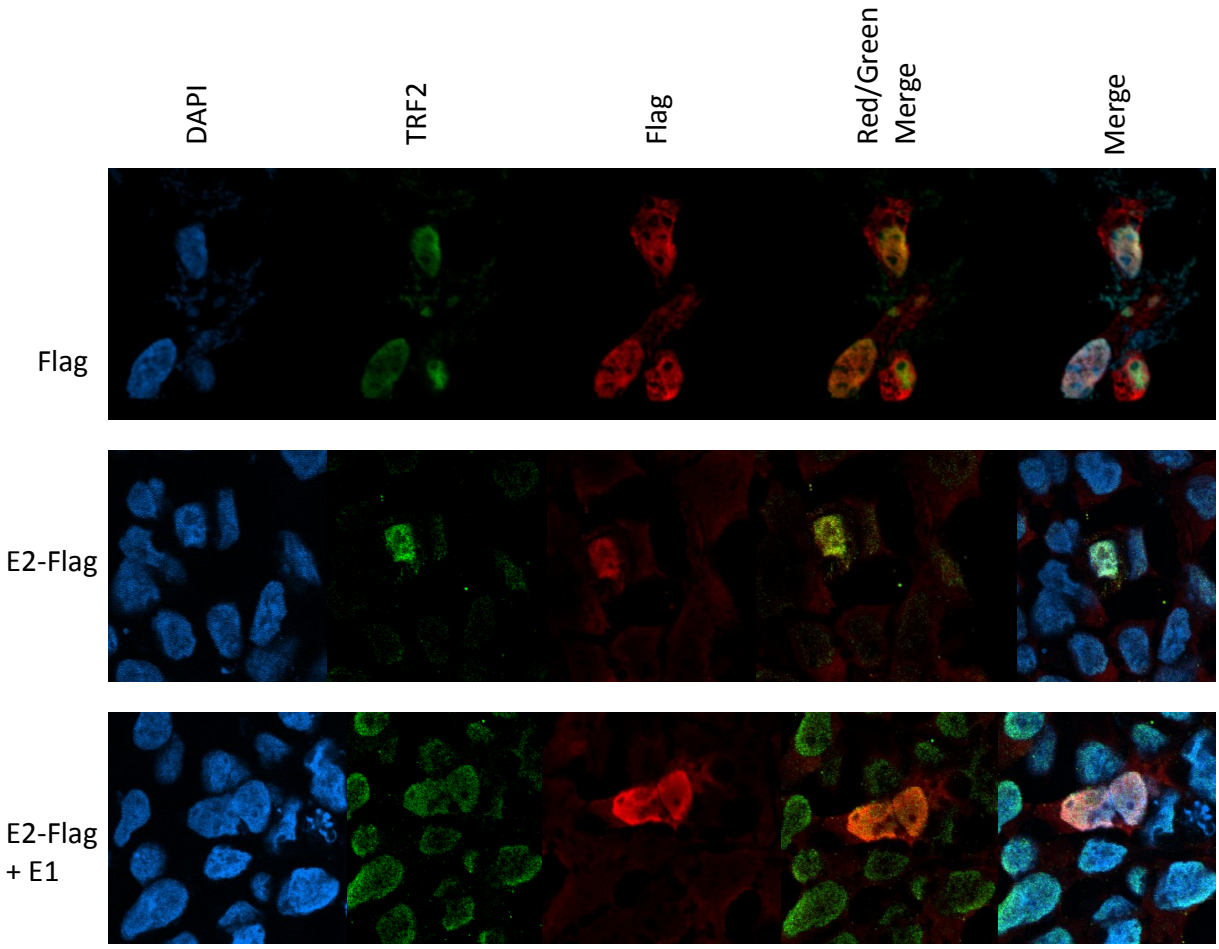


Fig 4.8E



Chapter 5

General Conclusions

Several tumor viruses interact with elements of the DNA Repair and Telomere Maintenance systems [reviewed in (89, 201). The DNA repair systems are co-opted by a number of other DNA viruses that utilize a long-term persistence strategy for survival inside the nucleus. Additionally, Epstein Barr-Virus and Kaposi's Sarcoma Herpesvirus utilize factors from the telomere maintenance system for their own persistence: for example, the TRF2 protein as a means of attracting elements of the homologous recombination repair system to improve segregation and recruitment of the Origin Recognition Complex (ORC) (85, 87). HPVs are also implicated in co-opting elements of the DNA repair system as a means of initiating replication. This is done through the E1 protein triggering double-strand breaks, leading to recruit of ATM and a number of downstream factors which, ultimately, are required for efficient replication.

The objective of this work was to identify cellular factors that HPVs utilize for their DNA replication and maintenance of the genome. During the maintenance phase, the virus expresses only low levels of its own replication proteins and DNA replicates once per cell cycle. This mechanism of viral DNA replication is in contrast to the rolling-circle mechanism observed during the vegetative phase, when high levels of viral DNA replication proteins are expressed. It is likely, then, that the virus utilizes certain cellular proteins that are necessary for cellular replication.

At some point during the viral lifecycle, E2 begins to be expressed and initiates its functions in DNA replication and maintenance. E2 binds with high affinity to a DNA site with sequence matching ATTg-N4-cAAT, with a high degree of sequence variability in the four nucleotide internal spacer. While it was previously surmised that no specific sequences were required in this spacer, my results have indicated that the specificity

varies by papillomavirus genus. The Deltapapillomaviruses, which include the BPV1 virus that was used for much of the early Papillomavirus E2 research, seem to be essentially insensitive, having a GC percentage of around 50%. Alphapapillomaviruses, by contrast, were observed to have a significantly lower GC content, which agrees with previous research demonstrating that HPV16 requires AT nucleotides in their spacer to maintain proper binding affinity. This is theorized to be due to BPV E2's greater ability to distort the DNA helix's shape to accommodate the E2 binding pocket compared to HPV16's, a trend that appears to remain consistent throughout their respective genera. The other papillomavirus genera seemed to all fall into the middle of these two extremes, containing an average GC content of approximately 30%. Interestingly, these trends further correlate with adaptive radiation to infect different cell types. The majority of PV's infect cutaneous keratinocytes. The Alpha genus contains members which are capable of infecting both cutaneous and mucosal keratinocytes, and the Delta genus contains members which are capable of infecting fibroblasts.

While newly discovered papillomaviruses are classified based on sequences from the L1 Orf, phylogenetic analysis based on different HPV open reading frames allows investigators to group Papillomaviral species by varying characteristics of that gene. As such, we performed phylogenetic analysis of the E2 ORF of all papillomaviruses, demonstrating that, while all the viruses sorted into their clades based on their respective genera, some differences were apparent from L1 phylogenetic trees. The genus which showed the most evolutionary distance from the others was the Deltapapillomavirus genus. This correlates with previous results indicating the greater numbers of E2BSs in Deltapapillomaviral genomes and their greater ability to deform DNA structure with their E2 protein as compared to the other Papillomaviral genera (177). Additionally, the Alpha PVs can be divided into two sub-clades within their genus,

one of which contained all of the Alphapapillomaviruses which are classified as “high risk” for progression to cervical cancer. The second group contained the majority of the Alphapapillomaviruses which were still capable of infecting cutaneous keratinocytes. As expected, the first group also showed a much higher reliance on low GC content of their binding sites compared to the second. This coupled with the E2BS spacer data makes it tempting to associate the sites of infection with reliance of E2 on GC content of spacers, particularly given that the differential binding affinity of the E2 binding sites affects the order in which they are occupied and, thus, the levels of gene expression during the infection. These differences could theoretically allow for precision regulation of gene expression and plasmid replication.

The yeast system allows for simplification and dissection of replication and maintenance functions for PVs. It has previously been demonstrated that HPVs are capable of replicating in yeast in the absence of any HPV specific genes. It is likely, then, that the host factors which play a key role in replication in mammalian cells are similarly involved in replication in yeast. Surprisingly, we discovered that a subset of HPVs are capable of replicating in this system. HPVs 16, 31, and 6 all replicate robustly, while 11, 18, and BPV1 do not. This failure to initiate stable replication does not match up with any specific phylogenetic groupings, including those performed in this report, but it does correlate with an observed lack of nine-base telomeric repeat sites in the late regions of the non-replicators. BPV1, in particular, comes from the very divergent Deltapapillomavirus genera and relies on Brd4 for segregation, unlike other HPVs that only utilize Brd4 for transcription. Working from the hypothesis that these binding sites were involved in yeast replication, we designed plasmid constructs whereby we could remove these nine-base repeats by site-directed mutagenesis. The results showed that, while the results for removing particular sites did not correlate with a loss of segregation

effect, mutation of TRF binding sites had an overall disruptive effect on plasmid stability, particularly when multiple mutations were introduced into the same plasmid. This instability was reflected by changes in DNA copy number.

The role of TRF2 in EBV replication involves an interaction between TRF2 and the EBV EBNA1 protein, which is a structural and functional homologue of the HPV E2 protein. To investigate if a similar interaction could be observed with E2, Far Western blots, bacterial GST-His pull-downs, and co-immunoprecipitation of mammalian proteins were performed. These studies demonstrated that E2 and TRF2 interact. Additionally, E2 interacted with other telomere maintenance and DNA repair proteins, namely TRF1, Pot1, Rap1, and Tin2. This agreed with results from CHIP assay demonstrating that similar DNA repair/telomere maintenance proteins interact with HPV16 telomere repeat sequences and that some of these factors are upregulated in cells after immortalization with HPV16 or hTert. Additionally, E2 co-localizes with TRF2 in the nucleus of cells independent of the E1-E2 complexes that were observed. Collectively, this evidence suggests that TRF2/E2 interactions are similar to EBNA1/TRF2 complex, where the purpose facilitates binding of repair factors to the DNA binding site in order to improve plasmid stability and segregation fidelity after DNA replication. Given the growing evidence that HPV16 DNA replication is initiated through E1 protein inducing site-specific double strand breaks in host chromosomes and, in doing so, initiates an activation of the host DNA repair response, which is necessary for Papillomavirus replication. Thus, one of TRF2's important functions may be to prevent integration of the newly replicated HPV16 DNA into the host chromosome during the repair response, a function that is crucial for stability of the viral chromosome. Another possibility is that the proteins that facilitate post-replication processing of replication products are required for

disentangling plasmids after once-per-cell-cycle replication and processing concatamers after rolling circle replication.

Our proposed model for HPV utilization of TRF2 is similar to that which has been proposed for EBV and KSHV. One of TRF2's important functions in cells is to loop the end of the chromosome back into a D-loop structure which protects the telomere ends from attack by the Non-Homologous End Joining System. In EBV and KSHV, the proteins similarly bind latent viral origins and reshape the viral chromosomes into a higher order chromatin structure. This would increase the frequency of initiation of DNA replication at these sites as well as down-stream maintenance of the DNA replication products. The interaction between TRF2 proteins with EBNA1 or LANA is critical for this process, particularly in KSHV, which does not contain TRF binding sites within the terminal repeat. We propose that HPV16 utilizes a similar loop-remodeling activity to improve its own long-term plasmid stability, as evidenced by the interaction of E2 with shelterin proteins and the subsequent plasmid instability if TRF binding sites are mutated. Further work will be required to elucidate the precise mechanisms of these processes.

References

1. 2002. Chapter 17, The Reference Sequence (RefSeq) Project., The NCBI handbook. National Library of Medicine (US), National Center for Biotechnology Information, Bethesda.
2. **Alexander, M. K., and V. A. Zakian.** 2003. Rap1p telomere association is not required for mitotic stability of a C3TA2 telomere in yeast. Embo J **22:**1688-1696.
3. **Alsheimer, M., B. Liebe, L. Sewell, C. L. Stewart, H. Scherthan, and R. Benavente.** 2004. Disruption of spermatogenesis in mice lacking A-type lamins. Journal of Cell Science **117:**1173-1178.
4. **Amiard, S., M. Doudeau, S. Pinte, A. Poulet, C. Lenain, C. Faivre-Moskalenko, D. Angelov, N. Hug, A. Vindigni, P. Bouvet, J. Paoletti, E. Gilson, and M.-J. Giraud-Panis.** 2007. A topological mechanism for TRF2-enhanced strand invasion. Nat Struct Mol Biol **14:**147-154.
5. **Angeletti A.K., P. D., Rogers, A., Lee. K., and Angeletti, P. C.** 2008. Human papillomavirus E2-dependent genome amplification, maintenance, and transcription, recapitulated in Saccharomyces cerevisiae. In preparation.
6. **Angeletti, P. C.** 2006. Human papillomavirus E2-dependent genome amplification, maintenance, and transcription, recapitulated in Saccharomyces cerevisiae. In preparation.
7. **Angeletti, P. C., K. Kim, F. J. Fernandes, and P. F. Lambert.** 2002. Stable replication of papillomavirus genomes in Saccharomyces cerevisiae. J Virol **76:**3350-8.
8. **Antson, A. A., J. E. Burns, O. V. Moroz, D. J. Scott, C. M. Sanders, I. B. Bronstein, G. G. Dodson, K. S. Wilson, and N. J. Maitland.** 2000. Structure of the intact transactivation domain of the human papillomavirus E2 protein. Nature **403:**805-9.
9. **Armbruster, B. N., C. M. Linardic, T. Veldman, N. P. Bansal, D. L. Downie, and C. M. Counter.** 2004. Rescue of an hTERT Mutant Defective in Telomere Elongation by Fusion with hPot1. Molecular and Cellular Biology **24:**3552-3561.
10. **Arnoult, N., C. Saintome, I. Ourliac-Garnier, J.-F. o. Riou, and A. LondoÃ±o-Vallejo.** 2009. Human POT1 is required for efficient telomere C-rich strand replication in the absence of WRN. Genes & Development **23:**2915-2924.
11. **Ashrafi, G. H., D. R. Brown, K. H. Fife, and M. S. Campo.** 2006. Down-regulation of MHC class I is a property common to papillomavirus E5 proteins. Virus Research **120:**208-211.
12. **Atanasiu, C., Z. Deng, A. Wiedmer, J. Norseen, and P. M. Lieberman.** 2006. ORC binding to TRF2 stimulates OriP replication. EMBO Rep **7:**716-21.
13. **Atanasiu, C., Z. Deng, A. Wiedmer, J. Norseen, and P. M. Lieberman.** 2006. ORC binding to TRF2 stimulates OriP replication. EMBO Rep **7:**716-721.
14. **Bai, Y., and J. P. Murnane.** 2003. Telomere instability in a human tumor cell line expressing a dominant-negative WRN protein. Human Genetics **113:**337-347.
15. **Bailey, T. L., and C. Elkan.** 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers, p. 28-36, Proceedings of the Second International

Conference on Intelligent Systems for Molecular Biology. AAAI Press, Menlo Park, California.
16. **Bailey, T. L., and M. Gribskov.** 1998. Combining evidence using p-values: application to sequence homology searches. Bioinformatics **14:**48-54.
17. **Bakkenist, C. J., and M. B. Kastan.** 2003. DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. Nature **421:**499-506.
18. **Ballestas, M. E., P. A. Chatis, and K. M. Kaye.** 1999. Efficient persistence of extrachromosomal KSHV DNA mediated by latency-associated nuclear antigen. Science **284:**641-644.
19. **Ballestas, M. E., and K. M. Kaye.** 2001. Kaposi's sarcoma-associated herpesvirus latency-associated nuclear antigen 1 mediates episome persistence through cis-acting terminal repeat (TR) sequence and specifically binds TR DNA. J Virol **75:**3250-3258.
20. **Ballestas, M. E., and K. M. Kaye.** 2011. The latency-associated nuclear antigen, a multifunctional protein central to Kaposi's sarcoma-associated herpesvirus latency. Future Microbiol **6:**1399-413.
21. **Barbera, A. J., J. V. Chodaparambil, B. Kelley-Clarke, K. Luger, and K. M. Kaye.** 2006. Kaposi's sarcoma-associated herpesvirus LANA hitches a ride on the chromosome. Cell Cycle **5:**1048-52.
22. **Bastien, N., and A. A. McBride.** 2000. Interaction of the papillomavirus E2 protein with mitotic chromosomes. Virology **270:**124-34.
23. **Baumann, P., and T. R. Cech.** 2001. Pot1, the Putative Telomere End-Binding Protein in Fission Yeast and Humans. Science **292:**1171-1175.
24. **Baumann, P., E. Podell, and T. R. Cech.** 2002. Human Pot1 (Protection of Telomeres) Protein: Cytolocalization, Gene Structure, and Alternative Splicing. Molecular and Cellular Biology **22:**8079-8087.
25. **Baxter, M. K., M. G. McPhillips, K. Ozato, and A. A. McBride.** 2005. The mitotic chromosome binding activity of the papillomavirus E2 protein correlates with interaction with the cellular chromosomal protein, Brd4. J Virol **79:**4806-18.
26. **Bedard, K. M., M. P. Underbrink, H. L. Howie, and D. A. Galloway.** 2008. The E6 oncoproteins from human betapapillomaviruses differentially activate telomerase through an E6AP-dependent mechanism and prolong the lifespan of primary keratinocytes. J Virol **82:**3894-902.
27. **Bellanger, S., C. Ling Tan, S. Teissier, and F. Thierry.** 2011. Tumor suppressor or oncogene? A critical role of the human papillomavirus (HPV) E2 protein in cervical cancer progression. Am J Cancer Res. **1:**373-389.
28. **Bellanger, S., C. Ling Tan, Y. Zhen Xue, S. Teissier, and F. Thierry.** 2011. Tumor suppressor or oncogene? A critical role of the human papillomavirus (HPV) E2 protein in cervical cancer progression. Am J Cancer Res **1:**373-389.
29. **Benetti, R., M. Garcia-Cao, and M. A. Blasco.** 2007. Telomere length regulates the epigenetic status of mammalian telomeres and subtelomeres. Nat Genet **39:**243-250.
30. **Bernard, B. A., C. Bailly, M. C. Lenoir, M. Darmon, F. Thierry, and M. Yaniv.** 1989. The human papillomavirus type 18 (HPV18) E2 gene product is a repressor of the HPV18 regulatory region in human keratinocytes. Journal of Virology **63:**4317-4324.
31. **Bernard, H.-U., I. E. Calleja-Macias, and S. T. Dunn.** 2006. Genome variation of human papillomavirus types: Phylogenetic and medical implications. International Journal of Cancer **118:**1071-1076.
32. **Biessmann, H., and J. M. Mason.** 1997. <b><i>Chromosoma Focus</i></b&gt. Chromosoma **106:**63-69.

33. **Blakaj, D. M., C. Kattamuri, S. Khrapunov, R. S. Hegde, and M. Brenowitz.** 2006. Indirect Readout of DNA Sequence by Papillomavirus E2 Proteins Depends Upon Net Cation Uptake. J Mol Biol **358:**224-240.
34. **Bochkarev, A., J. A. Barwell, R. A. Pfuetzner, W. Furey Jr, A. M. Edwards, and L. Frappier.** 1995. Crystal structure of the DNA-binding domain of the Epstein-Barr virus origin-binding protein EBNA1. Cell **83:**39-46.
35. **Boichuk, S., L. Hu, J. Hein, and O. V. Gjoerup.** 2010. Multiple DNA Damage Signaling and Repair Pathways Deregulated by Simian Virus 40 Large T Antigen. Journal of Virology **84:**8007-8020.
36. **Bosch, F. X., and S. de SanjosÃ©.** 2003. Chapter 1: Human Papillomavirus and Cervical Cancerâ€"Burden and Assessment of Causality. JNCI Monographs **2003:**3-13.
37. **Bosch, F. X., S. Sanjose, X. Castellsague, V. Moreno, and N. and Munoz.** 2006. Papillomavirus Research: From Natural History to Vaccines and Beyond., p. 19-40. *In* M. In Saveria Campo (ed.). Caisreir Academic Press, Wymondham.
38. **Botchan, M.** 2004. Hitchhiking without covalent integration. Cell **117:**280-1.
39. **Bousarghin, L., A. TouzÃ©, P.-Y. Sizaret, and P. Coursaget.** 2003. Human Papillomavirus Types 16, 31, and 58 Use Different Endocytosis Pathways To Enter Cells. Journal of Virology **77:**3846-3850.
40. **Bradshaw, P. S., D. J. Stavropoulos, and M. S. Meyn.** 2005. Human telomeric protein TRF2 associates with genomic double-strand breaks as an early response to DNA damage. Nat Genet **37:**193-197.
41. **Brannon, A. R., J. A. Maresca, J. D. Boeke, M. A. Basrai, and A. A. McBride.** 2005. Reconstitution of papillomavirus E2-mediated plasmid maintenance in Saccharomyces cerevisiae by the Brd4 bromodomain protein. Proc Natl Acad Sci U S A **102:**2998-3003.
42. **Branzei, D., and M. Foiani.** 2008. Regulation of DNA repair throughout the cell cycle. Nat Rev Mol Cell Biol **9:**297-308.
43. **Broccoli, D., A. Smogorzewska, L. Chong, and T. de Lange.** 1997. Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2. Nat Genet **17:**231-235.
44. **Brown, D. R., D. Kitchin, B. Qadadri, N. Neptune, T. Batteiger, and A. Ermel.** 2006. The human papillomavirus type 11 E1^E4 protein is a transglutaminase 3 substrate and induces abnormalities of the cornified cell envelope. Virology **345:**290-298.
45. **Burma, S., B. P. Chen, M. Murphy, A. Kurimasa, and D. J. Chen.** 2001. ATM Phosphorylates Histone H2AX in Response to DNA Double-strand Breaks. Journal of Biological Chemistry **276:**42462-42467.
46. **Byers, B., and G. L.** 1975. Electron Microscopic Observations on the Meiotic Karyotype of Diploid and Tetraploid Saccharomyces Cerevisiae. Proc Natl Acad Sci U S A **72:**5056-5060.
47. **Celli, G. B., and T. de Lange.** 2005. DNA processing is not required for ATM-mediated telomere damage response after TRF2 deletion. Nat Cell Biol **7:**712-718.
48. **Celli, G. B., E. L. Denchi, and T. de Lange.** 2006. Ku70 stimulates fusion of dysfunctional telomeres yet protects chromosome ends from homologous recombination. Nat Cell Biol **8:**885-890.
49. **Chandra, A., T. R. Hughes, C. I. Nugent, and V. Lundblad.** 2001. Cdc13 both positively and negatively regulates telomere replication. Genes & Development **15:**404-414.
50. **Chatterjee, A., J. W. Freeman, and H. Busch.** 1987. Identification and Partial Characterization of a Mr 40,000 Nucleolar Antigen Associated with Cell Proliferation. Cancer Res **47:**1123-1129.

51. **Chattopadhyay, A., M. C. Schmidt, and S. A. Khan.** 2005. Identification of a 450-bp region of human papillomavirus type 1 that promotes episomal replication in Saccharomyces cerevisiae. Virology **340:**133-142.
52. **Chiang, C. M., G. Dong, T. R. Broker, and L. T. Chow.** 1992. Control of human papillomavirus type 11 origin of replication by the E2 family of transcription regulatory proteins. Journal of Virology **66:**5224-5231.
53. **Chikashige, Y., and Y. Hiraoka.** 2001. Telomere binding of the Rap1 protein is required for meiosis in fission yeast. Current Biology **11:**1618-1623.
54. **Choi, K. H., A. S. Farrell, A. S. Lakamp, and M. M. Ouellette.** Characterization of the DNA binding specificity of Shelterin complexes. Nucleic Acids Research **39:**9206-9223.
55. **Chung, I., H. Leonhardt, and K. Rippe.** De novo assembly of a PML nuclear subcompartment occurs through multiple pathways and induces telomere elongation. Journal of Cell Science **124:**3603-3618.
56. **Ciccia, A., and S. J. Elledge.** 2010. The DNA Damage Response: Making It Safe to Play with Knives. Molecular Cell **40:**179-204.
57. **Clifford, G. M., J. S. Smith, M. Plummer, N. Munoz, and S. Franceschi.** 2003. Human papillomavirus types in invasive cervical cancer worldwide: a meta-analysis. Br J Cancer **88:**63-73.
58. **Clower, R. V., J. C. Fisk, and T. Melendy.** 2006. Papillomavirus E1 Protein Binds to and Stimulates Human Topoisomerase I. Journal of Virology **80:**1584-1587.
59. **Colgin, L. M., K. Baran, P. Baumann, T. R. Cech, and R. R. Reddel.** 2003. Human POT1 Facilitates Telomere Elongation by Telomerase. Current Biology **13:**942-946.
60. **Collins, K.** 2006. The biogenesis and regulation of telomerase holoenzymes. Nat Rev Mol Cell Biol **7:**484-494.
61. **Conger, K. L., J.-S. Liu, S.-R. Kuo, L. T. Chow, and T. S.-F. Wang.** 1999. Human Papillomavirus DNA Replication. Journal of Biological Chemistry **274:**2696-2705.
62. **Conway, M. J., and C. Meyers.** 2009. Replication and assembly of human papillomaviruses. J Dent Res **88:**307-17.
63. **Corbach, G. S. a. S.** 1997. Dose-dependent regulation of the early promoter of human papillomavirus type 18 by the viral E2 protein. J Virol **71:**50-58.
64. **CouÃ«del, C., K. D. Mills, M. Barchi, L. Shen, A. Olshen, R. D. Johnson, A. Nussenzweig, J. Essers, R. Kanaar, G. C. Li, F. W. Alt, and M. Jasin.** 2004. Collaboration of homologous recombination and nonhomologous end-joining factors for the survival and integrity of mice and cells. Genes & Development **18:**1293-1304.
65. **Cowan, C. R., P. M. Carlton, and W. Z. Cande.** 2002. Reorganization and polarization of the meiotic bouquet-stage cell can be uncoupled from telomere clustering. Journal of Cell Science **115:**3757-3766.
66. **Crabbe, L., R. E. Verdun, C. I. Haggblom, and J. Karlseder.** 2004. Defective Telomere Lagging Strand Synthesis in Cells Lacking WRN Helicase Activity. Science **306:**1951-1953.
67. **D'Amours, D., F. Stegmeier, and A. Amon.** 2004. Cdc14 and Condensin Control the Dissolution of Cohesin-Independent Chromosome Linkages at Repeated DNA. Cell **117:**455-469.
68. **D'Souza, G., A. R. Kreimer, R. Viscidi, M. Pawlita, C. Fakhry, W. M. Koch, W. H. Westra, and M. L. Gillison.** 2007. Case-Control Study of Human Papillomavirus and Oropharyngeal Cancer. N Engl J Med **356:**1944-1956.
69. **DahlÃ©n, M., P. Sunnerhagen, and T. S.-F. Wang.** 2003. Replication Proteins Influence the Maintenance of Telomere Length and Telomerase Protein Stability. Molecular and Cellular Biology **23:**3031-3042.

70. **Dai, X., C. Huang, A. Bhusari, S. Sampathi, K. Schubert, and W. Chai.** Molecular steps of G-overhang generation at human telomeres and its function in chromosome end protection. Embo J **29:**2788-2801.
71. **Dao, L. D., Duffy, A., Van Tine, B.A., Wu, S., Chiang, C., Broker, T.R., and L.T. Chow.** 2006. Dynamic Localization of the Human Papillomavirus Type 11 Origin Binding Protein E2 through Mitosis While in Association with the Spindle Apparatus. Journal of Virology **80:**4792-4800.
72. **Dasgupta, S., J. Zabielski, M. Simonsson, and S. Burnett.** 1992. Rolling-circle replication of a high-copy BPV-1 plasmid. Journal of Molecular Biology **228:**1-6.
73. **Davy, C. E., D. J. Jackson, Q. Wang, K. Raj, P. J. Masterson, N. F. Fenner, S. Southern, S. Cuthill, J. B. A. Millar, and J. Doorbar.** 2002. Identification of a G2 Arrest Domain in the E1âˆ§E4 Protein of Human Papillomavirus Type 16. Journal of Virology **76:**9806-9818.
74. **Day, P. M., C. C. Baker, D. R. Lowy, and J. T. Schiller.** 2004. Establishment of papillomavirus infection is enhanced by promyelocytic leukemia protein (PML) expression. Proceedings of the National Academy of Sciences of the United States of America **101:**14252-14257.
75. **Day, P. M., D. R. Lowy, and J. T. Schiller.** 2003. Papillomaviruses infect cells via a clathrin-dependent pathway. Virology **307:**1-11.
76. **Day, P. M., R. B. S. Roden, D. R. Lowy, and J. T. Schiller.** 1998. The Papillomavirus Minor Capsid Protein, L2, Induces Localization of the Major Capsid Protein, L1, and the Viral Transcription/Replication Protein, E2, to PML Oncogenic Domains. Journal of Virology **72:**142-150.
77. **De Lange, T.** 2002. Protection of mammalian telomeres. Oncogene **21:**532-540.
78. **de Lange, T.** 2005. Shelterin: the protein complex that shapes and safeguards human telomeres. Genes & Development **19:**2100-2110.
79. **de Lange, T.** 2004. T-loops and the origin of telomeres. Nat Rev Mol Cell Biol **5:**323-329.
80. **DE LANGE, T., and J. H. J. PETRINI.** 2000. A New Connection at Human Telomeres: Association of the Mre11 Complex with TRF2. Cold Spring Harbor Symposia on Quantitative Biology **65:**265-274.
81. **de Villiers, E.-M., C. Fauquet, T. R. Broker, H.-U. Bernard, and H. zur Hausen.** 2004. Classification of papillomaviruses. Virology **324:**17-27.
82. **de Villiers, E. M., C. Fauquet, T. R. Broker, H. U. Bernard, and H. zur Hausen.** 2004. Classification of papillomaviruses. Virology **324:**17-27.
83. **Dell, G., K. W. Wilkinson, R. Tranter, J. Parish, R. Leo Brady, and K. Gaston.** 2003. Comparison of the Structure and DNA-binding Properties of the E2 Proteins from an Oncogenic and a Non-oncogenic Human Papillomavirus. J Mol Biol **334:**979-991.
84. **Denchi, E. L., and T. de Lange.** 2007. Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1. Nature **448:**1068-1071.
85. **Deng, Z., C. Atanasiu, J. S. Burg, D. Broccoli, and P. M. Lieberman.** 2003. Telomere Repeat Binding Factors TRF1, TRF2, and hRAP1 Modulate Replication of Epstein-Barr Virus OriP. J. Virol. **77:**11992-12001.
86. **Deng, Z., J. Dheekollu, D. Broccoli, A. Dutta, and P. M. Lieberman.** 2007. The Origin Recognition Complex Localizes to Telomere Repeats and Prevents Telomere-Circle Formation. Current Biology **17:**1989-1995.
87. **Deng, Z., L. Lezina, C.-J. Chen, S. Shtivelband, W. So, and P. M. Lieberman.** 2002. Telomeric Proteins Regulate Episomal Maintenance of Epstein-Barr Virus Origin of Plasmid Replication. Molecular Cell **9:**493-503.

88. **Deng, Z., L. Lezina, C. J. Chen, S. Shtivelband, W. So, and P. M. Lieberman.** 2002. Telomeric proteins regulate episomal maintenance of Epstein-Barr virus origin of plasmid replication. Mol Cell **9:**493-503.
89. **Deng, Z., Z. Wang, and P. M. Lieberman.** 2012. Telomeres and viruses: common themes of genome maintenance. Front Oncol **2:**201.
90. **Desaintes, C., C. Demeret, S. Goyat, M. Yaniv, and F. Thierry.** 1997. Expression of the papillomavirus E2 protein in HeLa cells leads to apoptosis. Embo J **16:**504-514.
91. **Dey, A., F. Chitsaz, A. Abbasi, T. Misteli, and K. Ozato.** 2003. The double bromodomain protein Brd4 binds to acetylated chromatin during interphase and mitosis. Proceedings of the National Academy of Sciences **100:**8758-8763.
92. **Dheekollu, J., Z. Deng, A. Wiedmer, M. D. Weitzman, and P. M. Lieberman.** 2007. A Role for MRE11, NBS1, and Recombination Junctions in Replication and Stable Maintenance of EBV Episomes. PLoS ONE **2:**e1257.
93. **Diede, S. J., and D. E. Gottschling.** 1999. Telomerase-Mediated Telomere Addition In Vivo Requires DNA Primase and DNA Polymerases Î± and Î´. Cell **99:**723-733.
94. **Dimaio, D., and D. Mattoon.** 2001. Mechanisms of cell transformation by papillomavirus E5 proteins. Oncogene **20:**7866-7873.
95. **DiMaio, D., and J. Settleman.** 1988. Bovine papillomavirus mutant temperature sensitive for transformation, replication and transactivation. Embo J **7:**1197-204.
96. **Ding, X., R. Xu, J. Yu, T. Xu, Y. Zhuang, and M. Han.** 2007. SUN1 Is Required for Telomere Attachment to Nuclear Envelope and Gametogenesis in Mice. Developmental Cell **12:**863-872.
97. **Donaldson, M. M., W. Boner, and I. M. Morgan.** 2007. TopBP1 Regulates Human Papillomavirus Type 16 E2 Interaction with Chromatin. Journal of Virology **81:**4338-4342.
98. **Doorbar, J., S. Ely, J. Sterling, C. McLean, and L. Crawford.** 1991. Specific interaction between HPV-16 E1-E4 and cytokeratins results in collapse of the epithelial cell intermediate filament network. Nature **352:**824-827.
99. **Dynek, J. N., and S. Smith.** 2004. Resolution of Sister Telomere Association Is Required for Progression Through Mitosis. Science **304:**97-100.
100. **Edgar, R. C.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res **32:**1792-7.
101. **Eickbush, T. H.** 1997. Telomerase and Retrotransposons: Which Came First? Science **277:**911-912.
102. **Ellis, N. A., J. Groden, T. Z. Ye, J. Straughen, D. J. Lennon, S. Ciocci, M. Proytcheva, and J. German.** 1995. The Bloom's syndrome gene product is homologous to RecQ helicases. Cell **83:**655-66.
103. **Ensser, A., and H. Pfister.** 1990. Epidermodysplasia verruciformis associated human papilomaviruses present a subgenus-specific organization of the regulatory genome region. Nucl. Acids Res. **18:**3919-3922.
104. **Evander, M., I. H. Frazer, E. Payne, Y. M. Qi, K. Hengst, and N. A. McMillan.** 1997. Identification of the alpha6 integrin as a candidate receptor for papillomaviruses. Journal of Virology **71:**2449-56.
105. **Evans, M., H. Aliesky, and K. Cooper.** 2003. Optimization of biotinyl-tyramide-based in situ hybridization for sensitive background-free applications on formalin-fixed, paraffin-embedded tissue specimens. BMC Clinical Pathology **3:**2.
106. **Evans, S. K., and V. Lundblad.** 2000. Positive and negative regulation of telomerase access to the telomere. Journal of Cell Science **113:**3357-3364.

107. **Fagagna, F. d. A. d., P. M. Reaper, L. Clay-Farrace, H. Fiegler, P. Carr, T. von Zglinicki, G. Saretzki, N. P. Carter, and S. P. Jackson.** 2003. A DNA damage checkpoint response in telomere-initiated senescence. Nature **426:**194-198.
108. **Falck, J., J. Coates, and S. P. Jackson.** 2005. Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage. Nature **434:**605-611.
109. **Favre, M., F. Breitburd, O. Croissant, and G. Orth.** 1977. Chromatin-like structures obtained after alkaline disruption of bovine and human papillomaviruses. Journal of Virology **21:**1205-1209.
110. **Feeney, K. M., and J. L. Parish.** 2009. Targeting mitotic chromosomes: a conserved mechanism to ensure viral genome persistence. Proc Biol Sci **276:**1535-44.
111. **Ferriero, D. U., M. Delarole, A. D. Nadra, and G. De Prat Gay.** 2005. Free Energy Contributions to Direct Readout of a DNA Sequence. J Biol Chem**:**32480-32484.
112. **Fields, B. N., D. M. Knipe, and P. M. Howley.** 1996. Fields Virology, third edition ed, vol. 2. Lippincott-Raven, Philadelphia.
113. **Fitzgerald, J. E., M. Grenon, and N. F. Lowndes.** 2009. 53BP1: Function and Mechanisms of Focal Recruitment. Biochemical Society Transactions **37:**897-904.
114. **Flores, E. R., and P. F. Lambert.** 1997. Evidence for a switch in the mode of human papillomavirus type 16 DNA replication during the viral life cycle. Journal of Virology **71:**7167-79.
115. **FouchÃ©, N., S. Ã–zgÃ¼r, D. Roy, and J. D. Griffith.** 2006. Replication fork regression in repetitive DNAs. Nucleic Acids Research **34:**6044-6050.
116. **FouchÃ©, N., A. J. Cesare, S. Willcox, S. Ã–zgÃ¼r, S. A. Compton, and J. D. Griffith.** 2006. The Basic Domain of TRF2 Directs Binding to DNA Junctions Irrespective of the Presence of TTAGGG Repeats. Journal of Biological Chemistry **281:**37486-37495.
117. **Fouts, E. T., X. Yu, E. H. Egelman, and M. R. Botchan.** 1999. Biochemical and electron microscopic image analysis of the hexameric E1 helicase. J Biol Chem **274:**4447-58.
118. **Frattini, M., S. Hurst, H. Lim, S. Swaminathan, and L. A. Laimins.** 1997. Abrogation of a mitotic checkpoint by E2 proteins from oncogenic human papillomaviruses correlates with increased turnover of the p53 tumor suppressor protein. Embo J **16:**318-331.
119. **Freeman, L., L. Aragon-Alcaide, and A. Strunnikov.** 2000. The Condensin Complex Governs Chromosome Condensation and Mitotic Transmission of Rdna. The Journal of Cell Biology **149:**811-824.
120. **Gammoh, N., H. S. Grm, P. Massimi, and L. Banks.** 2006. Regulation of Human Papillomavirus Type 16 E7 Activity through Direct Protein Interaction with the E2 Transcriptional Activator. Journal of Virology **80:**1787-1797.
121. **Garcia-Cao, M., R. O'Sullivan, A. H. F. M. Peters, T. Jenuwein, and M. A. Blasco.** 2004. Epigenetic regulation of telomere length in mammalian cells by the Suv39h1 and Suv39h2 histone methyltransferases. Nat Genet **36:**94-99.
122. **German, J., B. D, P. E, F. K, G. RM, K. I, L. Z, L. C, L. S, and Wahrman.** 1977. Bloom's syndrome. VI. The disorder in Israel and an estimation of the gene frequency in Ashkenazim. Am J Hum Genet. **29:**553-562.
123. **German, J., B. D., and P. E.** 1977. Bloom's Syndrome. V. Survellance for cancer in affected families. Clin Genet. **1977:**3.
124. **Giannone, R. J., H. W. McDonald, G. B. Hurst, R.-F. Shen, Y. Wang, and Y. Liu.** The Protein Network Surrounding the Human Telomere Repeat Binding Factors TRF1, TRF2, and POT1. PLoS ONE **5:**e12407.

125. **Gillespie, K. A., K. P. Mehta, L. A. Laimins, and C. A. Moody.** 2012. Human Papillomaviruses Recruit Cellular DNA Repair and Homologous Recombination Factors to Viral Replication Centers. Journal of Virology **86:**9520-9526.
126. **Gillison, M. L., and D. R. Lowy.** 2004. A causal role for human papillomavirus in head and neck cancer. The Lancet **363:**1488-1489.
127. **Gilson, E., and V. Geli.** 2007. How telomeres are replicated. Nat Rev Mol Cell Biol **8:**825-38.
128. **Gissmann, L., V. Diehl, H. J. Schultz-Coulon, and H. zur Hausen.** 1982. Molecular cloning and characterization of human papilloma virus DNA derived from a laryngeal papilloma. Journal of Virology **44:**393-400.
129. **Gonzalo, S., M. Garcia-Cao, M. F. Fraga, G. Schotta, A. H. F. M. Peters, S. E. Cotter, R. Eguia, D. C. Dean, M. Esteller, T. Jenuwein, and M. A. Blasco.** 2005. Role of the RB1 family in stabilizing histone methylation at constitutive heterochromatin. Nat Cell Biol **7:**420-428.
130. **Goytisolo, F. n. A., E. Samper, J. MartÃ-n-Caballero, P. Finnon, E. s. Herrera, J. M. Flores, S. D. Bouffler, and M. a. A. Blasco.** 2000. Short Telomeres Result in Organismal Hypersensitivity to Ionizing Radiation in Mammals. The Journal of Experimental Medicine **192:**1625-1636.
131. **Green, M., K. H. Brackmann, P. R. Sanders, P. M. Loewenstein, J. H. Freel, M. Eisinger, and S. A. Switlyk.** 1982. Isolation of a human papillomavirus from a patient with epidermodysplasia verruciformis: presence of related viral DNA genomes in human urogenital tumors. Proceedings of the National Academy of Sciences **79:**4437-4441.
132. **Greider, C. W.** 1996. Telomere Length Regulation. Annual Review of Biochemistry **65:**337-365.
133. **Griffith, J., A. Bianchi, and T. de Lange.** 1998. TRF1 promotes parallel pairing of telomeric tracts in vitro. Journal of Molecular Biology **278:**79-88.
134. **Griffith, J. D.** 1999. Mammalian telomeres end in a large duplex loop. Cell **97:**503-514.
135. **Griffith, J. D., L. Comeau, S. Rosenfield, R. M. Stansel, A. Bianchi, H. Moss, and T. de Lange.** 1999. Mammalian Telomeres End in a Large Duplex Loop. Cell **97:**503-514.
136. **Grundhoff, A., and D. Ganem.** 2003. The Latency-Associated Nuclear Antigen of Kaposi's Sarcoma-Associated Herpesvirus Permits Replication of Terminal Repeat-Containing Plasmids. Journal of Virology **77:**2779-2783.
137. **Han, Y., Y.-M. Loo, K. T. Militello, and T. Melendy.** 1999. Interactions of the Papovavirus DNA Replication Initiator Proteins, Bovine Papillomavirus Type 1 E1 and Simian Virus 40 Large T Antigen, with Human Replication Protein A. Journal of Virology **73:**4899-4907.
138. **Haque, F., D. Mazzeo, J. T. Patel, D. T. Smallwood, J. A. Ellis, C. M. Shanahan, and S. Shackleton.** Mammalian SUN Protein Interaction Networks at the Inner Nuclear Membrane and Their Role in Laminopathy Disease Processes. Journal of Biological Chemistry **285:**3487-3498.
139. **Harper, J. W., and S. J. Elledge.** 2007. The DNA Damage Response: Ten Years After. Molecular Cell **28:**739-745.
140. **Harrison, S., K. Fisenne, and J. Hearing.** 1994. Sequence requirements of the Epstein-Barr virus latent origin of DNA replication. J Virol **68:**1913-1925.
141. **Hegde, R. S.** 2002. The papillomavirus E2 proteins: structure, function, and biology. Annu Rev Biophys Biomol Struct **31:**343-60.
142. **Henderson, E., C. C. Hardin, S. K. Walk, I. Tinoco, and E. H. Blackburn.** 1987. Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine·guanine base pairs. Cell **51:**899-908.

143. **Hindmarsh, P., and L. Laimins.** 2007. Mechanisms regulating expression of the HPV 31 L1 and L2 capsid proteins and pseudovirion entry. Virology Journal **4:**19.
144. **Hines, C. S., C. Meghoo, S. Shetty, M. Biburger, M. Brenowitz, and R. S. Hegde.** 1998. DNA structure and flexibility in the sequence-specific binding of papillomavirus E2 proteins. J Mol Biol **276:**809-18.
145. **Hockemeyer, D., W. Palm, T. Else, J.-P. Daniels, K. K. Takai, J. Z. S. Ye, C. E. Keegan, T. de Lange, and G. D. Hammer.** 2007. Telomere protection by mammalian Pot1 requires interaction with Tpp1. Nat Struct Mol Biol **14:**754-761.
146. **Hockemeyer, D., A. J. Sfeir, J. W. Shay, W. E. Wright, and T. de Lange.** 2005. POT1 protects telomeres from a transient DNA damage response and determines how human chromosomes end. EMBO J. **24:**2667-2678.
147. **Hoffmann, R., B. Hirt, V. Bechtold, P. Beard, and K. Raj.** 2006. Different modes of human papillomavirus DNA replication during maintenance. J Virol **80:**4431-9.
148. **Hopfner, K.-P., L. Craig, G. Moncalian, R. A. Zinkel, T. Usui, B. A. L. Owen, A. Karcher, B. Henderson, J.-L. Bodmer, C. T. McMurray, J. P. Carney, J. H. J. Petrini, and J. A. Tainer.** 2002. The Rad50 zinc-hook is a structure joining Mre11 complexes in DNA recombination and repair. Nature **418:**564-566.
149. **Hu, J., E. Liu, and R. Renne.** 2009. Involvement of SSRP1 in Latent Replication of Kaposi's Sarcoma-Associated Herpesvirus. J. Virol. **83:**11051-11063.
150. **Huda, N., H. Tanaka, M. S. Mendonca, and D. Gilley.** 2009. DNA Damage-Induced Phosphorylation of TRF2 Is Required for the Fast Pathway of DNA Double-Strand Break Repair. Molecular and Cellular Biology **29:**3597-3604.
151. **Hultdin, M., E. Gronlund, K. F. Norrback, T. Just, K. Taneja, and G. Roos.** 2001. Replication timing of human telomeric DNA and other repetitive sequences analyzed by fluorescence in situ hybridization and flow cytometry. Exp Cell Res **271:**223-9.
152. **Iarc.** 2007. Human papillomaviruses. IARC Monogr Eval Carcinog Risks Hum **90:**1-636.
153. **KÃ¶cher, S., T. Rieckmann, G. Rohaly, W. Y. Mansour, E. Dikomey, I. Dornreiter, and J. Dahm-Daphi.** 2012. Radiation-induced double-strand breaks require ATM but not Artemis for homologous recombination during S-phase. Nucleic Acids Research **40:**8336-8347.
154. **Kanoh, J., and F. Ishikawa.** 2001. spRap1 and spRif1, recruited to telomeres by Taz1, are essential for telomere function in fission yeast. Current Biology **11:**1624-1630.
155. **Karlseder, J., K. Hoke, O. K. Mirzoeva, C. Bakkenist, M. B. Kastan, J. H. J. Petrini, and T. d. Lange.** 2004. The Telomeric Protein TRF2 Binds the ATM Kinase and Can Inhibit the ATM-Dependent DNA Damage Response. PLoS Biol **2:**e240.
156. **Kerppola, T. K., and T. Curran.** 1997. The transcription activation domains of Fos and Jun induce DNA bending through electrostatic interactions. Embo J **16:**2907-16.
157. **Kim, H., O.-H. Lee, H. Xin, L.-Y. Chen, J. Qin, H. K. Chae, S.-Y. Lin, A. Safari, D. Liu, and Z. Songyang.** 2009. TRF2 functions as a protein hub and regulates telomere maintenance by recognizing specific peptide motifs. Nat Struct Mol Biol **16:**372-379.
158. **Kim, K., P. C. Angeletti, E. C. Hassebroek, and P. F. Lambert.** 2005. Identification of cis-acting elements that mediate the replication and maintenance of human papillomavirus type 16 genomes in Saccharomyces cerevisiae. J Virol **79:**5933-42.
159. **Kim, K., and P. F. Lambert.** 2002. E1 Protein of Bovine Papillomavirus 1 Is Not Required for the Maintenance of Viral Plasmid DNA Replication. Virology **293:**10-14.
160. **Kim, S. H.** 2004. TIN2 mediates functions of TRF2 at human telomeres. J. Biol. Chem. **279:**43799-43804.

161. **Kirnbauer, R., F. Booy, N. Cheng, D. R. Lowy, and J. T. Schiller.** 1992. Papillomavirus L1 major capsid protein self-assembles into virus-like particles that are highly immunogenic. Proceedings of the National Academy of Sciences **89:**12180-12184.
162. **Kiyono, T., Foster, S. A., Koop, J. I., McDougall, J. K., Galloway, D. A. & Klingelhutz, A. J.** 1998. Both Rb/p16Ink4a inactivation and telomerase activity are required to immortalize human epithelial cells. Nature **396:**84-88.
163. **Klingelhutz, A. J., S. A. Foster, and J. K. McDougall.** 1996. Telomerase activation by the E6 gene product of human papillomavirus type 16. Nature **380:**79-82.
164. **Kobayashi, T., D. J. Heck, M. Nomura, and T. Horiuchi.** 1998. Expansion and contraction of ribosomal DNA repeats in Saccharomyces cerevisiae: requirement of replication fork blocking (Fob1) protein and the role of RNA polymerase I. Genes & Development **12:**3821-3830.
165. **Kudoh, A., S. Iwahori, Y. Sato, S. Nakayama, H. Isomura, T. Murata, and T. Tsurumi.** 2009. Homologous Recombinational Repair Factors Are Recruited and Loaded onto the Viral DNA Genome in Epstein-Barr Virus Replication Compartments. Journal of Virology **83:**6641-6651.
166. **Kuo, S. R., J. S. Liu, T. R. Broker, and L. T. Chow.** 1994. Cell-free replication of the human papillomavirus DNA with homologous viral E1 and E2 proteins and human cell extracts. J. Biol. Chem. **269:**24058-24065.
167. **Lee, J.-H., and T. T. Paull.** 2005. ATM Activation by DNA Double-Strand Breaks Through the Mre11-Rad50-Nbs1 Complex. Science **308:**551-554.
168. **Lee, J.-H., and T. T. Paull.** 2004. Direct Activation of the ATM Protein Kinase by the Mre11/Rad50/Nbs1 Complex. Science **304:**93-96.
169. **Lee, M.-A., M. E. Diamond, and J. L. Yates.** 1999. Genetic Evidence that EBNA-1 Is Needed for Efficient, Stable Latent Infection by Epstein-Barr Virus. Journal of Virology **73:**2974-2982.
170. **Lehman, C. W., and M. R. Botchan.** 1998. Segregation of viral plasmids depends on tethering to chromosomes and is regulated by phosphorylation. Proc Natl Acad Sci U S A **95:**4338-43.
171. **Lei, M., P. Baumann, and T. R. Cech.** 2002. Cooperative Binding of Single-Stranded Telomeric DNA by the Pot1 Protein of Schizosaccharomyces pombeâ€ Biochemistry **41:**14560-14568.
172. **Lentz, M., S. Stevens, J. Raynes, and N. Elkhoury.** 2006. A phosphorylation map of the bovine papillomavirus E1 helicase. Virology Journal **3:**13.
173. **Li, B., S. Oestreich, and T. de Lange.** 2000. Identification of Human Rap1: Implications for Telomere Evolution. Cell **101:**471-483.
174. **Li, J.-L., R. J. Harrison, A. P. Reszka, R. M. Brosh, V. A. Bohr, S. Neidle, and I. D. Hickson.** 2001. Inhibition of the Bloom's and Werner's Syndrome Helicases by G-Quadruplex Interacting Ligandsâ€ Biochemistry **40:**15194-15202.
175. **Li, R., J. Knight, G. Bream, A. Stenlund, and M. Botchan.** 1989. Specific recognition nucleotides and their DNA context determine the affinity of E2 protein for 17 binding sites in the BPV-1 genome. Genes Dev **3:**510-26.
176. **Li, R., J. Knight, G. Bream, A. Stenlund, and M. Botchan.** 1989. Specific recognition nucleotides and their DNA context determine the affinity of E2 protein for 17 binding sites in the BPV-1 genome. Genes & Development **3:**510-526.
177. **Li, R., Knight, J, Bream, G, Stenlund, A, and M Botchan.** 1989. Specific recognition nucleotides and their DNA context determine the affinity of E2 protein for 17 binding sites in the BPV-1 genome. Genes & Development **3:**510-526.

178. **Liebe, B., M. Alsheimer, C. HÃ¶Ã¶g, R. Benavente, and H. Scherthan.** 2004. Telomere Attachment, Meiotic Chromosome Condensation, Pairing, and Bouquet Stage Duration Are Modified in Spermatocytes Lacking Axial Elements. Molecular Biology of the Cell **15:**827-837.
179. **Lindner, S. E., and B. Sugden.** 2007. The plasmid replicon of Epstein-Barr virus: mechanistic insights into efficient, licensed, extrachromosomal replication in human cells. Plasmid **58:**1-12.
180. **Liu, D.** 2004. PTOP interacts with POT1 and regulates its localization to telomeres. Nat. Cell Biol. **6:**673-680.
181. **Liu, D., M. S. O'Connor, J. Qin, and Z. Songyang.** 2004. Telosome, a Mammalian Telomere-associated Complex Formed by Multiple Telomeric Proteins. Journal of Biological Chemistry **279:**51338-51342.
182. **Liu, L., S. Franco, B. Spyropoulos, P. B. Moens, M. A. Blasco, and D. L. Keefe.** 2004. Irregular telomeres impair meiotic synapsis and recombination in mice. Proceedings of the National Academy of Sciences of the United States of America **101:**6496-6501.
183. **Loayza, D., H. Parsons, J. Donigian, K. Hoke, and T. de Lange.** 2004. DNA Binding Features of Human POT1. Journal of Biological Chemistry **279:**13241-13248.
184. **Longworth, M. S., and L. A. Laimins.** 2004. Pathogenesis of human papillomaviruses in differentiating epithelia. Microbiol Mol Biol Rev **68:**362-72.
185. **Loo, Y.-M., and T. Melendy.** 2004. Recruitment of Replication Protein A by the Papillomavirus E1 Protein and Modulation by Single-Stranded DNA. Journal of Virology **78:**1605-1615.
186. **Lundblad, V.** 1998. Telomerase catalysis: A phylogenetically conserved reverse transcriptase. Proceedings of the National Academy of Sciences **95:**8415-8416.
187. **Lundblad, V.** 2002. Telomere maintenance without telomerase. Oncogene **21:**522-531.
188. **Lupton, S., and A. J. Levine.** 1985. Mapping genetic elements of Epstein-Barr virus that facilitate extrachromosomal persistence of Epstein-Barr virus-derived plasmids in human cells. Mol Cell Biol **5:**2533-2542.
189. **Lustig, A. J.** 2001. Cdc13 subcomplexes regulate multiple telomere functions. Nat Struct Mol Biol **8:**297-299.
190. **Lustig, A. J.** 2003. Clues to catastrophic telomere loss in mammals from yeast telomere rapid deletion. Nat Rev Genet **4:**916-923.
191. **M Remm, R. B., and J R Jenkins.** 1992. The E2 binding sites determine the efficiency of replication for the origin of human papillomavirus type 18. Nucleic Acids Res **20:**6015-6021.
192. **MÃ¼nger, K., W. C. Phelps, V. Bubb, P. M. Howley, and R. Schlegel.** 1989. The E6 and E7 genes of the human papillomavirus type 16 together are necessary and sufficient for transformation of primary human keratinocytes. Journal of Virology **63:**4417-4421.
193. **Maddar, H., N. Ratzkovsky, and A. Krauskopf.** 2001. Role for Telomere Cap Structure in Meiosis. Molecular Biology of the Cell **12:**3191-3203.
194. **Makovets, S., I. Herskowitz, and E. H. Blackburn.** 2004. Anatomy and Dynamics of DNA Replication Fork Movement in Yeast Telomeric Regions. Molecular and Cellular Biology **24:**4019-4031.
195. **Mao, Z., A. Seluanov, Y. Jiang, and V. Gorbunova.** 2007. TRF2 is required for repair of nontelomeric DNA double-strand breaks by homologous recombination. Proceedings of the National Academy of Sciences **104:**13068-13073.
196. **Marahrens, Y., and Bruce Stillman.** 1992. A Yeast Chromosomal Origin of DNA Replication Defined by Multiple Functional Elements. Science **255:**820.

197. **Masterson, P. J., M. A. Stanley, A. P. Lewis, and M. A. Romanos.** 1998. A C-Terminal Helicase Domain of the Human Papillomavirus E1 Protein Binds E2 and the DNA Polymerase Î±-Primase p68 Subunit. Journal of Virology **72:**7407-7419.
198. **Matsuoka, S., B. A. Ballif, A. Smogorzewska, E. R. McDonald, K. E. Hurov, J. Luo, C. E. Bakalarski, Z. Zhao, N. Solimini, Y. Lerenthal, Y. Shiloh, S. P. Gygi, and S. J. Elledge.** 2007. ATM and ATR Substrate Analysis Reveals Extensive Protein Networks Responsive to DNA Damage. Science **316:**1160-1166.
199. **Mattern, K. A., S. J. J. Swiggers, A. L. Nigg, B. LÃ¶wenberg, A. B. Houtsmuller, and J. M. J. M. Zijlmans.** 2004. Dynamics of Protein Binding to Telomeres in Living Cells: Implications for Telomere Structure and Function. Molecular and Cellular Biology **24:**5587-5594.
200. **McBride, A. A., M. G. McPhillips, and J. G. Oliveira.** 2004. Brd4: tethering, segregation and beyond. Trends Microbiol **12:**527-9.
201. **McBride, A. A., J. G. Oliveira, and M. G. McPhillips.** 2006. Partitioning Viral Genomes in Mitosis: Same Idea, Different Targets. Cell Cycle **5:**1499-1502.
202. **McBride, A. A., H. Romanczuk, and P. M. Howley.** 1991. The papillomavirus E2 regulatory proteins. J Biol Chem **266:**18411-4.
203. **McMillan, N. A. J., E. Payne, I. H. Frazer, and M. Evander.** 1999. Expression of the Î±6 Integrin Confers Papillomavirus Binding upon Receptor-Negative B-Cells. Virology **261:**271-279.
204. **McPhillips, M. G., J. G. Oliveira, J. E. Spindler, R. Mitra, and A. A. McBride.** 2006. Brd4 Is Required for E2-Mediated Transcriptional Activation but Not Genome Partitioning of All Papillomaviruses. J. Virol. **80:**9530-9543.
205. **McPhillips, M. G., J. G. Oliveira, J. E. Spindler, R. Mitra, and A. A. McBride.** 2006. Brd4 is required for E2-mediated transcriptional activation but not genome partitioning of all papillomaviruses. J Virol **80:**9530-43.
206. **McPhillips, M. G., J. G. Oliveira, J. E. Spindler, R. Mitra, and A. A. McBride.** 2006. Brd4 Is Required for E2-Mediated Transcriptional Activation but Not Genome Partitioning of All Papillomaviruses. Journal of Virology **80:**9530-9543.
207. **McPhillips, M. G., K. Ozato, and A. A. McBride.** 2005. Interaction of bovine papillomavirus E2 protein with Brd4 stabilizes its association with chromatin. J Virol **79:**8920-32.
208. **McPhillips, M. G., K. Ozato, and A. A. McBride.** 2005. Interaction of Bovine Papillomavirus E2 Protein with Brd4 Stabilizes Its Association with Chromatin. Journal of Virology **79:**8920-8932.
209. **Mecsas, J., and B. Sugden.** 1987. Replication of plasmids derived from bovine papilloma virus type 1 and Epstein-Barr virus in cells in culture. Annu Rev Cell Biol **3:**87-108.
210. **Melendy, T., J. Sedman, and A. Stenlund.** 1995. Cellular factors required for papillomavirus DNA replication. Journal of Virology **69:**7857-67.
211. **Miller, K. M., O. Rog, and J. P. Cooper.** 2006. Semi-conservative DNA replication through telomeres requires Taz1. Nature **440:**824-828.
212. **Mirzoeva, O. K., and J. H. J. Petrini.** 2001. DNA Damage-Dependent Nuclear Dynamics of the Mre11 Complex. Molecular and Cellular Biology **21:**281-288.
213. **Mistry, N., M. Simonsson, and M. Evander.** 2007. Transcriptional activation of the human papillomavirus type 5 and 16 long control region in cells from cutaneous and mucosal origin. Virol J **4:**27.

214. **Mitton-Fry, R. M., E. M. Anderson, T. R. Hughes, V. Lundblad, and D. S. Wuttke.** 2002. Conserved Structure for Single-Stranded Telomeric DNA Recognition. Science **296:**145-147.
215. **Miyake, Y., M. Nakamura, A. Nabetani, S. Shimamura, M. Tamura, S. Yonehara, M. Saito, and F. Ishikawa.** 2009. RPA-like Mammalian Ctc1-Stn1-Ten1 Complex Binds to Single-Stranded DNA and Protects Telomeres Independently of the Pot1 Pathway. Molecular Cell **36:**193-206.
216. **Mohaghegh, P., J. K. Karow, R. M. Brosh Jr, Jr., V. A. Bohr, and I. D. Hickson.** 2001. The Bloom's and Werner's syndrome proteins are DNA structure-specific helicases. Nucleic Acids Res **29:**2843-9.
217. **Mohr, I. J., R. Clark, S. Sun, E. J. Androphy, P. MacPherson, and M. R. Botchan.** 1990. Targeting the E1 replication protein to the papillomavirus origin of replication by complex formation with the E2 transactivator. Science **250:**1694-9.
218. **Moody, C. A., and L. A. Laimins.** 2009. Human Papillomaviruses Activate the ATM DNA Damage Pathway for Viral Genome Amplification upon Differentiation. PLoS Pathog **5:**e1000605.
219. **Muller, F., Y. S. Seo, and J. Hurwitz.** 1994. Replication of bovine papillomavirus type 1 origin-containing DNA in crude extracts and with purified proteins. J. Biol. Chem. **269:**17086-17094.
220. **Munoz, N., F. X. Bosch, S. de Sanjose, R. Herrero, X. Castellsague, K. V. Shah, P. J. Snijders, and C. J. Meijer.** 2003. Epidemiologic classification of human papillomavirus types associated with cervical cancer. N Engl J Med **348:**518-27.
221. **Munoz, N., F. X. Bosch, S. de Sanjose, R. Herrero, X. Castellsague, K. V. Shah, P. J. F. Snijders, and C. J. L. M. Meijer.** 2003. Epidemiologic Classification of Human Papillomavirus Types Associated with Cervical Cancer. New England Journal of Medicine **348:**518-527.
222. **Nakamura, M., A. Nabetani, T. Mizuno, F. Hanaoka, and F. Ishikawa.** 2005. Alterations of DNA and Chromatin Structures at Telomeres and Genetic Instability in Mouse Cells Defective in DNA Polymerase Î±. Molecular and Cellular Biology **25:**11073-11088.
223. **Narechania, A., M. Terai, and R. D. Burk.** 2005. Overlapping reading frames in closely related human papillomaviruses result in modular rates of selection within E2. J Gen Virol **86:**1307-13.
224. **Ng'andwe, C., J. Lowe, P. Richards, L. Hause, C. Wood, and P. Angeletti.** 2007. The distribution of sexually-transmitted Human Papillomaviruses in HIV positive and negative patients in Zambia, Africa. BMC Infectious Diseases **7:**77.
225. **Niller, H. H., G. Glaser, R. KnÃ¼chel, and H. Wolf.** 1995. Nucleoprotein Complexes and DNA 5â€²-Ends at oriP of Epstein-Barr Virus. Journal of Biological Chemistry **270:**12864-12868.
226. **Nugent, C. I., T. R. Hughes, N. F. Lue, and V. Lundblad.** 1996. Cdc13p: A Single-Strand Telomeric DNA-Binding Protein with a Dual Role in Yeast Telomere Maintenance. Science **274:**249-252.
227. **O'Connor, M. S., A. Safari, H. Xin, D. Liu, and Z. Songyang.** 2006. A critical role for TPP1 and TIN2 interaction in high-order telomeric complex assembly. Proc. Natl. Acad. Sci. USA **103:**11874-11879.
228. **O'Donovan, P. J., and D. M. Livingston.** 2010. BRCA1 and BRCA2: breast/ovarian cancer susceptibility gene products and participants in DNA double-strand break repair. Carcinogenesis **31:**961-967.

229. **Ofir, R., A. C. Wong, H. E. McDermid, K. L. Skorecki, and S. Selig.** 1999. Position effect of human telomeric repeats on replication timing. Proc Natl Acad Sci U S A **96:**11434-9.
230. **Oliveira, J. G., L. A. Colf, and A. A. McBride.** 2006. Variations in the association of papillomavirus E2 proteins with mitotic chromosomes. Proc Natl Acad Sci U S A **103:**1047-1052.
231. **Oliveira, J. G., L. A. Colf, and A. A. McBride.** 2006. Variations in the association of papillomavirus E2 proteins with mitotic chromosomes. Proceedings of the National Academy of Sciences of the United States of America **103:**1047-1052.
232. **Olson M, H. L., Cantor C, Botstein D.** 1989. A common language for physical mapping of the human genome. Science **245:**1434–1435.
233. **Ono, Y., K. Tomita, A. Matsuura, T. Nakagawa, H. Masukata, M. Uritani, T. Ushimaru, and M. Ueno.** 2003. A novel allele of fission yeast rad11 that causes defects in DNA repair and telomere length regulation. Nucleic Acids Research **31:**7141-7149.
234. **Opresko, P. L., P. A. Mason, E. R. Podell, M. Lei, I. D. Hickson, T. R. Cech, and V. A. Bohr.** 2005. POT1 Stimulates RecQ Helicases WRN and BLM to Unwind Telomeric DNA Substrates. Journal of Biological Chemistry **280:**32069-32080.
235. **Opresko, P. L., M. Otterlei, J. GraakjÃ¦r, P. Bruheim, L. Dawut, S. KÃ¸lvraa, A. May, M. M. Seidman, and V. A. Bohr.** 2004. The Werner Syndrome Helicase and Exonuclease Cooperate to Resolve Telomeric D Loops in a Manner Regulated by TRF1 and TRF2. Molecular Cell **14:**763-774.
236. **Opresko, P. L., C. von Kobbe, J.-P. Laine, J. Harrigan, I. D. Hickson, and V. A. Bohr.** 2002. Telomere-binding Protein TRF2 Binds to and Stimulates the Werner and Bloom Syndrome Helicases. Journal of Biological Chemistry **277:**41110-41119.
237. **Palm, W., and T. de Lange.** 2008. How Shelterin Protects Mammalian Telomeres. Annual Review of Genetics **42:**301-334.
238. **Papadopoulos, J. S., and R. Agarwala.** 2007. COBALT: constraint-based alignment tool for multiple protein sequences. Bioinformatics **23:**1073-1079.
239. **Parish, J. L., A. M. Bean, R. B. Park, and E. J. Androphy.** 2006. ChlR1 Is Required for Loading Papillomavirus E2 onto Mitotic Chromosomes and Viral Genome Maintenance. Molecular Cell **24:**867-876.
240. **Park, P., W. Copeland, L. Yang, T. Wang, M. R. Botchan, and I. J. Mohr.** 1994. The cellular DNA polymerase alpha-primase is required for papillomavirus DNA replication and associates with the viral E1 helicase. Proc Natl Acad Sci U S A **91:**8700-4.
241. **Park, R. B., and E. J. Androphy.** 2002. Genetic analysis of high-risk e6 in episomal maintenance of human papillomavirus genomes in primary human keratinocytes. J Virol **76:**11359-64.
242. **Patel, S. S., and K. M. Picha.** 2000. Structure and function of hexameric helicases. Annu Rev Biochem **69:**651-97.
243. **Pereira, R., Hitzeroth, II, and E. P. Rybicki.** 2009. Insights into the role and function of L2, the minor capsid protein of papillomaviruses. Arch Virol **154:**187-97.
244. **Piirsoo, M., E. Ustav, T. Mandel, A. Stenlund, and M. Ustav.** 1996. Cis and trans requirements for stable episomal maintenance of the BPV-1 replicator. Embo J **15:**1-11.
245. **Pittayakhajonwut, D., and P. C. Angeletti.** 2008. Analysis of cis-elements that facilitate extrachromosomal persistence of human papillomavirus genomes. Virology **374:**304-314.
246. **Poddar, A., S. C. Reed, M. G. McPhillips, J. E. Spindler, and A. A. McBride.** 2009. The Human Papillomavirus Type 8 E2 Tethering Protein Targets the Ribosomal DNA Loci of Host Mitotic Chromosomes. Journal of Virology **83:**640-650.

247. **Podgornaya, O. I., E. A. Bugaeva, A. P. Voronin, E. Gilson, and A. R. Mitchell.** 2000. Nuclear envelope associated protein that binds telomeric DNAs. Molecular Reproduction and Development **57:**16-25.
248. **Pruitt KD, T., T, Maglott DR** 2007. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins Nucleic Acids Res **Jan 1:**35(Database issue):D61-5.
249. **R Li, J Knight, G Bream, a. A Stenlund, and a. M. Botchan.** 1989. Specific recognition nucleotides and their DNA context determine the affinity of E2 protein for 17 binding sites in the BPV-1 genome. . Genes & Development**:**510-526.
250. **Rabson, M. S., C. Yee, Y. C. Yang, and P. M. Howley.** 1986. Bovine papillomavirus type 1 3' early region transformation and plasmid maintenance functions. J Virol **60:**626-34.
251. **Reinson, T., M. Toots, M. Kadaja, R. Pipitch, M. Allik, E. Ustav, and M. Ustav.** 2013. Engagement of the ATR-Dependent DNA Damage Response at the Human Papillomavirus 18 Replication Centers during the Initial Amplification. Journal of Virology **87:**951-964.
252. **Rogoff, H. A., M. T. Pickering, F. M. Frame, M. E. Debatis, Y. Sanchez, S. Jones, and T. F. Kowalik.** 2004. Apoptosis Associated with Deregulated E2F Activity Is Dependent on E2F1 and Atm/Nbs1/Chk2. Molecular and Cellular Biology **24:**2968-2977.
253. **Sailaja, G., R. M. Watts, and H. U. Bernard.** 1999. Many different papillomaviruses have low transcriptional activity in spite of strong epithelial specific enhancers. J Gen Virol **80 (Pt 7):**1715-24.
254. **Sakai, H., T. Yasugi, J. D. Benson, J. J. Dowhanick, and P. M. Howley.** 1996. Targeted mutagenesis of the human papillomavirus type 16 E2 transactivation domain reveals separable transcriptional activation and DNA replication functions. J Virol **70:**1602-11.
255. **Sakakibara, N., R. Mitra, and A. A. McBride.** 2011. The Papillomavirus E1 Helicase Activates a Cellular DNA Damage Response in Viral Replication Foci. Journal of Virology **85:**8981-8995.
256. **Salas, T. R., I. Petruseva, O. Lavrik, A. Bourdoncle, J.-L. Mergny, A. Favre, and C. SaintomÃ©.** 2006. Human replication protein A unfolds telomeric G-quadruplexes. Nucleic Acids Research **34:**4857-4865.
257. **Sanchez, I. E., M. Dellarole, K. Gaston, and G. de Prat Gay.** 2008. Comprehensive comparison of the interaction of the E2 master regulator with its cognate target DNA sites in 73 human papillomavirus types by sequence statistics. Nucl. Acids Res. **36:**756-769.
258. **Sanders, C. M., and A. Stenlund.** 1998. Recruitment and loading of the E1 initiator protein: an ATP-dependent process catalysed by a transcription factor. EMBO J **17:**7044-7055.
259. **Sarafi, T. R., and A. A. McBride.** 1995. Domains of the BPV-1 E1 Replication Protein Required for Origin-Specific DNA Binding and Interaction with the E2 Transactivator. Virology **211:**385-396.
260. **Scherthan, H., A. Sfeir, and T. de Lange.** Rap1-independent telomere attachment and bouquet formation in mammalian meiosis. Chromosoma **120:**151-157.
261. **Schlegel, R., W. C. Phelps, Y. L. Zhang, and B. M.** 1988. Quantitative keratinocyte assay detects two biological activities of human papillomavirus DNA and identifies viral types associated with cervical carcinoma. Embo J **7:**3181-3187.
262. **Schuck, S., and A. Stenlund.** 2005. Assembly of a Double Hexameric Helicase. Molecular Cell **20:**377-389.

263. **Schweiger, M. R., J. You, and P. M. Howley.** 2006. Bromodomain protein 4 mediates the papillomavirus E2 transcriptional activation function. J Virol **80:**4276-85.
264. **Sedman, J., and A. Stenlund.** 1998. The Papillomavirus E1 Protein Forms a DNA-Dependent Hexameric Complex with ATPase and DNA Helicase Activities. Journal of Virology **72:**6893-6897.
265. **Sedman, T., J. Sedman, and A. Stenlund.** 1997. Binding of the E1 and E2 proteins to the origin of replication of bovine papillomavirus. Journal of Virology **71:**2887-96.
266. **Sfeir, A., and T. de Lange.** 2012. Removal of Shelterin Reveals the Telomere End-Protection Problem. Science **336:**593-597.
267. **Sfeir, A., and T. de Lange.** Removal of Shelterin Reveals the Telomere End-Protection Problem. Science **336:**593-597.
268. **Sfeir, A., S. T. Kosiyatrakul, D. Hockemeyer, S. L. MacRae, J. Karlseder, C. L. Schildkraut, and T. de Lange.** 2009. Mammalian Telomeres Resemble Fragile Sites and Require TRF1 for Efficient Replication. Cell **138:**90-103.
269. **Shire, K., D. F. J. Ceccarelli, T. M. Avolio-Hunter, and L. Frappier.** 1999. EBP2, a Human Protein That Interacts with Sequences of the Epstein-Barr Virus Nuclear Antigen 1 Important for Plasmid Maintenance. Journal of Virology **73:**2587-2595.
270. **Si, H., S. C. Verma, and E. S. Robertson.** 2006. Proteomic analysis of the Kaposi's sarcoma-associated herpesvirus terminal repeat element binding proteins. J Virol **80:**9017-30.
271. **Silla, T., I. Haal, J. Geimanen, K. Janikson, A. Abroi, E. Ustav, and M. Ustav.** 2005. Episomal maintenance of plasmids with hybrid origins in mouse cells. J Virol **79:**15277-88.
272. **Skalsky, R. L., J. Hu, and R. Renne.** 2007. Analysis of viral cis elements conferring Kaposi's sarcoma-associated herpesvirus episome partitioning and maintenance. J Virol **81:**9825-37.
273. **Skiadopoulos, M. H., and A. A. McBride.** 1998. Bovine papillomavirus type 1 genomes and the E2 transactivator protein are closely associated with mitotic chromatin. J Virol **72:**2079-88.
274. **Skiadopoulos, M. H., and A. A. McBride.** 1998. Bovine Papillomavirus Type 1 Genomes and the E2 Transactivator Protein Are Closely Associated with Mitotic Chromatin. Journal of Virology **72:**2079-2088.
275. **Slijepcevic, P.** 2006. The role of DNA damage response proteins at telomeres--an "integrative" model. DNA Repair (Amst) **5:**1299-306.
276. **Smith, J., H. Zou, and R. Rothstein.** 2000. Characterization of genetic interactions with RFA1: the role of RPA in DNA replication and telomere maintenance. Biochimie **82:**71-78.
277. **Smith, J. L., S. K. Campos, and M. A. Ozbun.** 2007. Human Papillomavirus Type 31 Uses a Caveolin 1- and Dynamin 2-Mediated Entry Pathway for Infection of Human Keratinocytes. Journal of Virology **81:**9922-9931.
278. **Smogorzewska, A., and T. de Lange.** 2004. Regulation of telomerase by telomeric proteins. Annu. Rev. Biochem. **73:**177-208.
279. **Songyang, Z., and D. Liu.** 2006. Inside the mammalian telomere interactome: regulation and regulatory activities of telomeres. Crit Rev Eukaryot Gene Expr **16:**103-18.
280. **Sowd, G., H. Wang, D. Pretto, W. J. Chazin, and P. L. Opresko.** 2009. Replication Protein A Stimulates the Werner Syndrome Protein Branch Migration Activity. Journal of Biological Chemistry **284:**34682-34691.

281. **Stagno D'Alcontres, M., A. Mendez-Bermudez, J. L. Foxon, N. J. Royle, and P. Salomoni.** 2007. Lack of TRF2 in ALT cells causes PML-dependent p53 activation and loss of telomeric DNA. The Journal of Cell Biology **179:**855-867.
282. **Stanley, M. A., H. Browne, A. M, and M. AC.** 1989. Properties of a non-tumorigenic human cervical keratinocyte cell line. Int J Cancer **43:**672-676.
283. **Stansel, R. M., T. de Lange, and J. D. Griffith.** 2001. T-loop assembly in vitro involves binding of TRF2 near the 3[prime] telomeric overhang. Embo J **20:**5532-5540.
284. **Stiff, T., M. Oâ€™Driscoll, N. Rief, K. Iwabuchi, M. LÃ¶brich, and P. A. Jeggo.** 2004. ATM and DNA-PK Function Redundantly to Phosphorylate H2AX after Exposure to Ionizing Radiation. Cancer Res **64:**2390-2396.
285. **Strauss, J. K., and L. J. Maher, 3rd.** 1994. DNA bending by asymmetric phosphate neutralization. Science **266:**1829-34.
286. **Surovtseva, Y. V., D. Churikov, K. A. Boltz, X. Song, J. C. Lamb, R. Warrington, K. Leehy, M. Heacock, C. M. Price, and D. E. Shippen.** 2009. Conserved Telomere Maintenance Component 1 Interacts with STN1 and Maintains Chromosome Ends in Higher Eukaryotes. Molecular Cell **36:**207-218.
287. **Swindle, C. S., N. Zou, B. A. Van Tine, G. M. Shaw, J. A. Engler, and L. T. Chow.** 1999. Human Papillomavirus DNA Replication Compartments in a Transient DNA Replication System. Journal of Virology **73:**1001-1009.
288. **Takai, K. K., S. Hooper, S. Blackwood, R. Gandhi, and T. de Lange.** In Vivo Stoichiometry of Shelterin Components. Journal of Biological Chemistry **285:**1457-1467.
289. **Takata, M., M. S. Sasaki, E. Sonoda, C. Morrison, M. Hashimoto, H. Utsumi, Y. Yamaguchi-Iwai, A. Shinohara, and S. Takeda.** 1998. Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. EMBO J **17:**5497-5508.
290. **Tan, M., C. L. Jahn, and C. M. Price.** 2003. Origin usage during euplotes ribosomal DNA amplification. Eukaryot Cell **2:**115-22.
291. **Tanaka, H., M. S. Mendonca, P. S. Bradshaw, D. J. Hoelz, L. H. Malkas, M. S. Meyn, and D. Gilley.** 2005. DNA damage-induced phosphorylation of the human telomere-associated protein TRF2. Proceedings of the National Academy of Sciences of the United States of America **102:**15539-15544.
292. **Tatsumi, Y., K. Ezura, K. Yoshida, T. Yugawa, M. Narisawa-Saito, T. Kiyono, S. Ohta, C. Obuse, and M. Fujita.** 2008. Involvement of human ORC and TRF2 in pre-replication complex assembly at telomeres. Genes to Cells **13:**1045-1059.
293. **Thain, A., O. Jenkins, A. Clarke, and K. Gaston.** 1996. CpG methylation directly inhibits binding of the human papillomavirus type 16 E2 protein to specific DNA sequences. J. Virol. **70:**7233-7235.
294. **Theobald, D. L., R. B. Cervantes, V. Lundblad, and D. S. Wuttke.** 2003. Homology Among Telomeric End-Protection Proteins. Structure **11:**1049-1050.
295. **Thomas, M. C., and C. M. Chiang.** 2005. E6 oncoprotein represses p53-dependent gene activation via inhibition of protein acetylation independently of inducing p53 degradation. Mol Cell **17:**251-64.
296. **Titolo, S., A. Pelletier, A.-M. Pulichino, K. Brault, E. Wardrop, P. W. White, M. G. Cordingley, and J. Archambault.** 2000. Identification of Domains of the Human Papillomavirus Type 11 E1 Helicase Involved in Oligomerization and Binding to the Viral Origin. Journal of Virology **74:**7349-7361.

297. **Tommerup, H., A. Dousmanis, and T. de Lange.** 1994. Unusual chromatin in human telomeres. Molecular and Cellular Biology **14:**5777-5785.
298. **TP, C., H. TH, T. JP, T. F, S. P. 3rd, D. M, L. Gissmann, R. A, and T. LP.** 1987. Transcriptional regulation of the human papillomavirus-16 E6-E7 promoter by a keratinocyte-dependent enhancer, and by viral E2 trans-activator and repressor gene products: implications for cervical carcinogenesis. EMBO J **6:**3745-53.
299. **Trujillo, K. M., S.-S. F. Yuan, E. Y.-H. P. Lee, and P. Sung.** 1998. Nuclease Activities in a Complex of Human Recombination and DNA Repair Factors Rad50, Mre11, and p95. Journal of Biological Chemistry **273:**21447-21450.
300. **Tzur, Y. B., K. L. Wilson, and Y. Gruenbaum.** 2006. SUN-domain proteins: 'Velcro' that links the nucleoskeleton to the cytoskeleton. Nat Rev Mol Cell Biol **7:**782-788.
301. **Ustav, M., E. Ustav, P. Szymanski, and A. Stenlund.** 1991. Identification of the origin of replication of bovine papillomavirus and characterization of the viral origin recognition factor E1. Embo J **10:**4321-9.
302. **van Steensel, B., A. Smogorzewska, and T. de Lange.** 1998. TRF2 Protects Human Telomeres from End-to-End Fusions. Cell **92:**401-413.
303. **Veldman, T., I. Horikawa, J. C. Barrett, and R. Schlegel.** 2001. Transcriptional Activation of the Telomerase hTERT Gene by Human Papillomavirus Type 16 E6 Oncoprotein. Journal of Virology **75:**4467-4472.
304. **Verdun, R. E., and J. Karlseder.** 2006. The DNA Damage Machinery and Homologous Recombination Pathway Act Consecutively to Protect Human Telomeres. Cell **127:**709-720.
305. **Voet, T., B. Liebe, C. Labaere, P. Marynen, and H. Scherthan.** 2003. Telomere-independent homologue pairing and checkpoint escape of accessory ring chromosomes in male mouse meiosis. The Journal of Cell Biology **162:**795-808.
306. **Walbloomers, J. M. M., Jacobs, M.V., Manos, M.M., Bosch, F.X., Kummer, J.A., Shah, K.V., Snijders, P.J.F., Peto, J., Meijer, C.J.L.M., Muñoz, N.** 1999. Human Papillomavirus is a Necessary Cause of Invasive Cervical Cancer Worldwide. The Journal of Pathology **189:**12-19.
307. **Walboomers, J. M. M., M. V. Jacobs, M. M. Manos, F. X. Bosch, J. A. Kummer, K. V. Shah, P. J. F. Snijders, J. Peto, C. J. L. M. Meijer, and N. Muñoz.** 1999. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. The Journal of Pathology **189:**12-19.
308. **Wang, F.** 2007. The POT1-TPP1 telomere complex is a telomerase processivity factor. Nature **445:**506-510.
309. **Wang, R. C., A. Smogorzewska, and T. de Lange.** 2004. Homologous Recombination Generates T-Loop-Sized Deletions at Human Telomeres. Cell **119:**355-368.
310. **Wei, C., and C. M. Price.** 2004. Cell Cycle Localization, Dimerization, and Binding Domain Architecture of the Telomere Protein cPot1. Molecular and Cellular Biology **24:**2091-2102.
311. **Wellinger, R. J., A. J. Wolf, and V. A. Zakian.** 1993. Origin activation and formation of single-strand TG1-3 tails occur sequentially in late S phase on a yeast linear plasmid. Molecular and Cellular Biology **13:**4057-4065.
312. **Wellinger, R. J., A. J. Wolf, and V. A. Zakian.** 1993. Saccharomyces telomeres acquire single-strand TG1â€“3 tails late in S phase. Cell **72:**51-60.
313. **Wentzensen, N., S. Vinokurova, and M. von Knebel Doeberitz.** 2004. Systematic review of genomic integration sites of human papillomavirus genomes in epithelial dysplasia and invasive cancer of the female lower genital tract. Cancer Res **64:**3878-84.

314. **Wilkinson, D. E., and S. K. Weller.** 2004. Recruitment of Cellular Recombination and Repair Proteins to Sites of Herpes Simplex Virus Type 1 DNA Replication Is Dependent on the Composition of Viral Proteins within Prereplicative Sites and Correlates with the Induction of the DNA Damage Response. Journal of Virology **78:**4783-4796.
315. **Wilson, P. F., P. B. Nham, S. S. Urbin, J. M. Hinz, I. M. Jones, and L. H. Thompson.** 2010. Inter-individual variation in DNA double-strand break repair in human fibroblasts before and after exposure to low doses of ionizing radiation. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis **683:**91-97.
316. **Wold, M. S.** 1997. REPLICATION PROTEIN A:A Heterotrimeric, Single-Stranded DNA-Binding Protein Required for Eukaryotic DNA Metabolism. Annual Review of Biochemistry **66:**61-92.
317. **Wright, W. E., V. M. Tesmer, M. L. Liao, and J. W. Shay.** 1999. Normal human telomeres are not late replicating. Exp Cell Res **251:**492-9.
318. **Wu, S. Y., and C. M. Chiang.** 2007. The double bromodomain-containing chromatin adaptor Brd4 and transcriptional regulation. J Biol Chem **282:**13141-13145.
319. **Xin, H.** 2007. TPP1 is a homologue of ciliate TEBP-beta and interacts with POT1 to recruit telomerase. Nature **445:**559-562.
320. **Xue, Y., S. Bellanger, W. Zhang, D. Lim, J. Low, D. Lunny, and F. o. Thierry.** HPV16 E2 Is an Immediate Early Marker of Viral Infection, Preceding E7 Expression in Precursor Structures of Cervical Carcinoma. Cancer Res **70:**5316-5325.
321. **Yang, Q., Y.-L. Zheng, and C. C. Harris.** 2005. POT1 and TRF2 Cooperate To Maintain Telomeric Integrity. Molecular and Cellular Biology **25:**1070-1080.
322. **Yang, Q., Y. L. Zheng, and C. C. Harris.** 2005. POT1 and TRF2 cooperate to maintain telomeric integrity. Mol. Cell. Biol. **25:**1070-1080.
323. **Yates, J., N. Warren, D. Reisman, and B. Sugden.** 1984. A cis-acting element from the Epstein-Barr viral genome that permits stable replication of recombinant plasmids in latently infected cells. Proceedings of the National Academy of Sciences **81:**3806-3810.
324. **Yates, J., N. Warren, D. Reisman, and B. Sugden.** 1984. A cis-acting element from the Epstein-Barr viral genome that permits stable replication of recombinant plasmids in latently infected cells. Proc Natl Acad Sci U S A **81:**3806-10.
325. **Yates, J. L., S. M. Camiolo, and J. M. Bashaw.** 2000. The Minimal Replicator of Epstein-Barr VirusoriP. Journal of Virology **74:**4512-4522.
326. **Ye, J., C. Lenain, S. Bauwens, A. Rizzo, A. d. Saint-LÃ©ger, A. s. Poulet, D. Benarroch, F. d. r. Magdinier, J. Morere, S. Amiard, E. Verhoeyen, S. b. Britton, P. Calsou, B. Salles, A. Bizard, M. Nadal, E. Salvati, L. Sabatier, Y. Wu, A. Biroccio, A. LondoÃ±o-Vallejo, M.-J. p. Giraud-Panis, and E. Gilson.** TRF2 and Apollo Cooperate with Topoisomerase 2Î± to Protect Human Telomeres from Replicative Damage. Cell **142:**230-242.
327. **Ye, J. Z.-S., J. R. Donigian, M. van Overbeek, D. Loayza, Y. Luo, A. N. Krutchinsky, B. T. Chait, and T. de Lange.** 2004. TIN2 Binds TRF1 and TRF2 Simultaneously and Stabilizes the TRF2 Complex on Telomeres. Journal of Biological Chemistry **279:**47264-47271.
328. **Yongli Zhang, Zhiqun Xi, Rashmi S. Hegde, Zippora Shakked, and a. D. M. Crothers.** 2004. Predicting indirect readout effects in protein–DNA interactions PNAS **101:** 8337-8341.
329. **You, J., J. L. Croyle, A. Nishimura, K. Ozato, and P. M. Howley.** 2004. Interaction of the bovine papillomavirus E2 protein with Brd4 tethers the viral DNA to host mitotic chromosomes. Cell **117:**349-60.

330. **You, J., J. L. Croyle, A. Nishimura, K. Ozato, and P. M. Howley.** 2004. Interaction of the Bovine Papillomavirus E2 Protein with Brd4 Tethers the Viral DNA to Host Mitotic Chromosomes. Cell **117:**349-360.
331. **You, J., M.-R. Schweiger, and P. M. Howley.** 2005. Inhibition of E2 Binding to Brd4 Enhances Viral Genome Loss and Phenotypic Reversion of Bovine Papillomavirus-Transformed Cells. Journal of Virology **79:**14956-14961.
332. **Zachow, K., RSOstrow, M. Bender, S. Watts, T. Okagaki, F. Pass, and A. Faras.** 1982. Detection of Human Papillomavirus DNA in Anogenital Neoplasias. Nature **300:**771-773.
333. **Zaug, A. J., E. R. Podell, and T. R. Cech.** 2005. Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension in vitro. Proceedings of the National Academy of Sciences of the United States of America **102:**10864-10869.
334. **Zhang, B., P. Li, E. Wang, Z. Brahmi, K. W. Dunn, J. S. Blum, and A. Roman.** 2003. The E5 protein of human papillomavirus type 16 perturbs MHC class II antigen maturation in human foreskin keratinocytes treated with interferon-Î³. Virology **310:**100-108.
335. **Zhao, K.-N., X.-Y. Sun, I. H. Frazer, and J. Zhou.** 1998. DNA Packaging by L1 and L2 Capsid Proteins of Bovine Papillomavirus Type 1. Virology **243:**482-491.
336. **Zhao, K. N., and I. H. Frazer.** 2002. Saccharomyces cerevisiae is permissive for replication of bovine papillomavirus type 1. J Virol **76:**12265-73.
337. **Zhao, K. N., J. Zhou, and I. H. Frazer.** 2001. Bovine papillomavirus type-1 virions can be amplified in yeast (Saccharomyces cerevisiae) cultures infected with natural BPV-1 virions. Submitted.
338. **Zhao, Y., A. J. Sfeir, Y. Zou, C. M. Buseman, T. T. Chow, J. W. Shay, and W. E. Wright.** 2009. Telomere Extension Occurs at Most Chromosome Ends and Is Uncoupled from Fill-In in Human Cancer Cells. Cell **138:**463-475.
339. **Zheng, Z. M., and C. C. Baker.** 2006. Papillomavirus genome structure, expression, and post-transcriptional regulation. Front Biosci **11:**2286-302.
340. **Zhou, J., Z. Deng, J. Norseen, and P. M. Lieberman.** Regulation of Epstein-Barr virus origin of plasmid replication (OriP) by the S-phase checkpoint kinase Chk2. J Virol **84:**4979-87.
341. **Zhou, J., X. Y. Sun, D. J. Stenzel, and I. H. Frazer.** 1991. Expression of vaccinia recombinant HPV 16 L1 and L2 ORF proteins in epithelial cells is sufficient for assembly of HPV virion-like particles. Virology **185:**251-257.
342. **Zickler, D., and N. Kleckner.** 1998. THE LEPTOTENE-ZYGOTENE TRANSITION OF MEIOSIS. Annual Review of Genetics **32:**619-697.
343. **Zou, Y., Y. Liu, X. Wu, and S. M. Shell.** 2006. Functions of human replication protein A (RPA): From DNA replication to DNA damage and stress responses. Journal of Cellular Physiology **208:**267-273.
344. **zur Hausen, H.** 2009. Papillomaviruses in the causation of human cancers â€" a brief historical account. Virology **384:**260-265.