Seminars in Cancer Biology xxx (xxxx) xxx

Contents lists available at ScienceDirect

Seminars in Cancer Biology

journal homepage: www.elsevier.com/locate/semcancer

Review

The role of RNA-binding proteins in the processing of mRNAs produced by carcinogenic papillomaviruses

Naoko Kajitani [a,b], Stefan Schwartz [a,b,*]

[a] Department of Medical Biochemistry and Microbiology (IMBIM), Uppsala University, BMC-B9, 751 23, Uppsala, Sweden
[b] Department of Laboratory Medicine, Lund University, BMC-B13, 221 84, Lund, Sweden

ARTICLE INFO

Keywords:
Papillomavirus
Splicing
Polyadenylation
hnRNP
SR protein
Akt
DDR

ABSTRACT

Human papillomaviruses (HPV) are epitheliotropic DNA tumor viruses that are prevalent in the human population. A subset of the HPVs termed high-risk HPVs (HR-HPVs) are causative agents of anogenital cancers and head-and-neck cancers. Cancer is the result of persistent high-risk HPV infections that have not been cleared by the immune system of the host. These infections are characterized by dysregulated HPV gene expression, in particular constitutive high expression of the HPV E6 and E7 oncogenes and absence of the highly immunogenic viral L1 and L2 capsid proteins. HPVs make extensive use of alternative mRNA splicing to express its genes and are therefore highly dependent on cellular RNA-binding proteins for proper gene expression. Levels of RNA-binding proteins are altered in HPV-containing premalignant cervical lesions and in cervical cancer. Here we review our current knowledge of RNA-binding proteins that control HPV gene expression. We focus on RNA-binding proteins that control expression of the E6 and E7 oncogenes since they initiate and drive development of cancer and on the immunogenic L1 and L2 proteins as there silencing may contribute to immune evasion during carcinogenesis. Furthermore, cellular RNA-binding proteins are essential for HPV gene expression and as such may be targets for therapy to HPV infections and HPV-driven cancers.

1. Introduction

Virus infections account for approximately 15 % of all cancers [1]. There are seven human tumor viruses including human papillomavirus (HPV), merkel cell polyomavirus (MCPyV), Epstein-Barr-virus (EBV), human herpesvirus type 8 (HHV-8), hepatitis B virus (HBV), hepatitis C virus (HCV) and human T-lymphotropic virus type 1 (HTLV-1). The basic mechanism by which these viruses contribute to cancer development is diverse and the majority of the cancer-associated viruses are unrelated and belong to different families: HPV (*Papillomaviridae* family), MCPyV (*Polyomaviridae* family), EBV and HHV-8 (*Herpesviridae* family) and HBV (*Hepadnaviridae* family) are DNA viruses while HCV (*Flaviviridae* family) is an RNA virus whereas HTLV-1 is a retrovirus belonging to the *Retroviridae* family. In addition, the origin of the various cancers caused by these viruses differ and largely reflect the tropism of each virus, e.g. HPV is the causative agent of anogenital cancers, primarily cervical cancer or head-and-neck cancers. Each virus has evolved and adapted to a unique environment and cell type in the host. These viruses have also developed their own viral replication cycle that are intimately and individually linked to each virus pathogenic properties.

The vast majority of the human tumor viruses, with the exception of HCV and HTLV-1, are DNA viruses that replicate in the nucleus of the infected cells and have access to the cellular transcription machinery as well to the cellular splicing machinery. HTLV-1 is of interest in that it is an RNA virus by definition and as such encapsidates an RNA genome in the virion but converts the single stranded RNA genome into a double stranded DNA form once it enters human cells. Thus, getting access to the cellular transcriptional control factors as well as to the cellular splicing machinery, both of which HTLV-1 utilizes well for its own purpose. Many of the tumor viruses encode transcriptional activator

Abbreviations: DDR, DNA damage response; eUTR, human papillomavirus early untranslated region; HNSCC, head-and-neck squamous cell carcinomas; hnRNP, heterogeneous nuclear ribonuclear protein; HPV, human papillomavirus; HR-HPV, high-risk human papillomavirus; LCR, long control region; LR-HPV, low-risk human papillomavirus; lUTR, human papillomavirus late untranslated region; ORF, open reading frame; pAE, human papillomavirus early polyadenylation signal; pAL, human papillomavirus late polyadenylation signal; RBP, RNA-binding protein; SA, splice acceptor/3′-splice site; SD, splice donor/5′-splice site; SR-protein, serine- arginine-rich protein; UTR, untranslated region.

* Corresponding author at: Department of Medical Biochemistry and Microbiology (IMBIM), BMC-B9, Uppsala University, Box 582, 751 23, Uppsala, Sweden.
E-mail address: stefan.schwartz@imbim.uu.se (S. Schwartz).

https://doi.org/10.1016/j.semcancer.2022.02.014
Received 7 December 2021; Received in revised form 10 February 2022; Accepted 11 February 2022
Available online 16 February 2022

proteins that interact with viral promoter regions directly or indirectly to direct the cellular transcription machinery to the viral genomes, i.e. EBV EBNA [2], HHV-8 LANA [3], HTLV-1 Tax [4] and HPV E2 [5,6]. Following transcription, the viral mRNAs are processed with help of the cellular splicing and polyadenylation machineries. As a consequence of the many and often overlapping genes on the retroviral and DNA viral genomes, RNA processing plays a major role in the gene expression programs of the tumor viruses, rendering them dependent on cellular RNA-binding proteins that control these processes. HCV, as an RNA virus that replicates in the cytoplasm, is totally dependent on cellular RNA-binding proteins as well. HCV has even evolved to get access to cellular proteins with a predominantly nuclear localization by perturbing subcellular protein localization in such a way that nuclear cellular proteins re-localize to the cytoplasmic compartment in HCV-infected cells [7]. Some of the tumor viruses express viral RNA-binding proteins that regulate posttranscriptional processes during maturation and expression of viral mRNAs such as the HTLV-1 Rex protein [8] and the HHV-8 ORF57 protein [9], whereas the HPV E2 protein controls HPV polyadenylation through protein-protein interactions [10]. In conclusion, the human tumor viruses are dependent on cellular RNA-binding proteins to control viral gene expression. For many of these viruses, the ability to persist and hide from the immune system of the host is intimately connected to their ability to cause cancer. Furthermore, to evade the immune system in an immunocompetent host and to establish viral persistence is dependent on the ability of these viruses to control viral gene expression. In this review article, we focus on cellular RNA-binding proteins that control HPV gene expression and discuss the significance of currently published results in relation to HPV-associated oncogenesis.

2. HPV is a causative agent of cervical cancers and head-and-neck cancers

The HPV family includes more than 200 HPV types that preferentially infect either mucosal or cutaneous epithelium. Mucosal HPV types are further divided into high-risk HPV (HR-HPV) and low-risk HPV (LR-HPV) types depending on their association with cancer, i.e. anogenital cancers and head-and-neck squamous cell carcinomas (HNSCC). Papillomaviruses are classified into genera and HPVs can be found in five different papillomavirus genera [11,12]. Most of our knowledge on the contribution of HPVs to disease, including cancer, is derived from research on the mucosal HPVs in the α-HPV genus. The α-HPV genus is further divided into various “species”, each containing multiple HPV types, of which α-HPV genus 7 and 9 are of interest as they harbor HPV18 and HPV16, respectively, that together account for >80 % of all cases of cervical cancer world-wide [13]. More than 95 % of cervical cancer cases are positive for HR-HPV DNA, but HPV16 (~65 %) and HPV18 (~15 %) dominate. HPVs account for more than 500,000 annual cases of cervical cancer and 300,000 annual deaths world-wide [14,15]. In countries that have implemented vaccination to the major high-risk HPV types HPV16 and HPV18 the risk of invasive cervical cancer is reduced [16]. In addition, some of the low-risk HPVs are known to cause benign cervical lesions and genital warts and may also cause laryngeal papillomatosis, a rare benign proliferative disease that may be life-threatening and require life-long surgical treatment to prevent air way obstruction [17]. Less is known about HPV in HNSCCs, primarily oropharyngeal cancer, but the incidence of HPV-positive oropharyngeal cancer is currently bypassing the incidence of HPV-positive cervical cancers in some countries [18–20]. The vast majority of the HPV-positive HNSCCs are caused by HPV16 and other HPV types are rare to find in HNSCC [21]. Thus, HPV16 is the most common mucosal HPV type. In this review, we will mainly discuss HPV16 and we will refer to nucleotide numbers of the HPV16 reference genome regarding splice sites and polyadenylation signals [11]. Other HPV types, primarily HPV18 and HPV31, will be discussed as well and in relation to HPV16.

3. Molecular mechanism of HPV oncogenesis

3.1. HPV16 genome

HPV is a small DNA tumor virus with a virion diameter of 55 nm that carries a double stranded circular DNA genome of approximately 8 kilobases (kb) genome (Fig. 1). Molecular cloning and sequencing of the papillomaviruses have revealed a typical genomic organization of all HPV family members: eight or nine open reading frames (ORFs) found on the same DNA strand [11]. The HPV genome can be divided into three regions: 1) The long control region (LCR) which is a non-coding region located between the stop codon of the L1 ORF and the start codon of the E6 ORF. The LCR contains most of the regulatory elements for viral DNA replication and transcription. 2) A coding region for the early genes E6, E7, E1, E2, E4 and E5 and 3) a late coding region for the major and minor L1 and L2 capsid proteins, respectively.

3.2. HPV16 life cycle

The HPV infection initiates by the attachment of the HPV virion to cellular primary receptor(s) such as heparan sulfate proteoglycans (HSPG) that are enriched on the mucosal cell surface and in the extracellular matrix (ECM) at the epithelial basal membrane (Fig. 2) [22]. These interactions induce conformational changes that allow clathrin-mediated endosomal trafficking of the virus capsids into the nucleus of the keratinocytes in the basal layer of the stratified epithelium. Next, episomal HPV DNA in the host nucleus is subjected to transcription from the early promoter and expression of the HPV early genes commences [5,23] (Fig. 2). The E1 and E2 proteins enter the nucleus and induce replication of the HPV DNA with the help of the cellular DNA polymerases [24,25]. HPV DNA is maintained as extrachromosomal episomal DNA. Initially HPV DNA is amplified to relatively low numbers, approximately 50–100 copies in each undifferentiated basal cell, a number that is maintained in sub-differentiated cells. HPV16 lacks a DNA polymerase and depends on cellular DNA polymerases for viral genome amplification. However, cellular DNA polymerases in epithelial keratinocytes are inactivated once mitotic epithelial basal cells divide and detach from the basal layer to enter the epithelial differentiation program. HPV therefore produces the promitotic E7 protein [26] that continues to drive mitosis of the HPV-infected cell and the antiapoptotic E6 protein [27] that grants survival of the infected cell by preventing apoptosis and by contributing to immune evasion. In this respect, HPV E5 contributes to cell proliferation and immune evasion by interacting with the epidermal growth factor receptor and by down-regulating MHC-I [28]. As the HPV-infected cells reach higher levels of differentiation in the stratified epithelium, HPV DNA is subjected to vegetative amplification, i.e. a rapid production of a high number of HPV genomes, resulting in thousands of copies of HPV DNA per cell [29,30] (Fig. 2). Once the infected cells reach the upper layers with terminally differentiated cells in the stratified epithelium, the late promoter is also activated and high levels of E2 protein are produced followed by the E1^E4 protein [31]. The HPV late promoter produces mRNAs encoding E1, E2 and E1^E4 proteins, and subsequently, HPV late L1 and L2 mRNAs. The latter are translated into HPV L1 and L2 capsid proteins that self-assemble into capsids with genomic HPV DNA to form virions that are shed from the top of the epithelium [32–34] (Fig. 2). In the next section, we will discuss the ability of HPV to drive cell proliferation to achieve high HPV DNA replication and how this property of HPV may contribute to development of malignancy of the HPV-infected cells.

3.3. Natural course of HPV16 infection

HPV16 is the most common sexually transmitted HPV type and the most common HR-HPV type found in anogenital and head and neck cancer [13,22]. During the natural course of infection, HPV16 infections

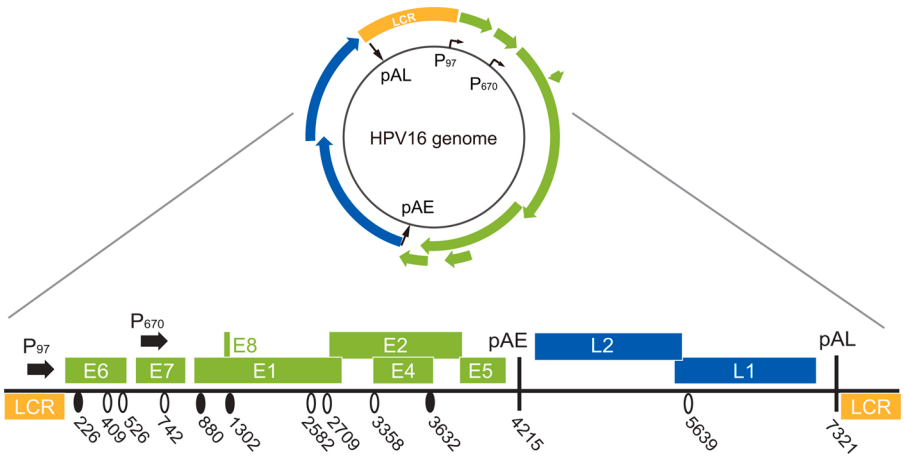
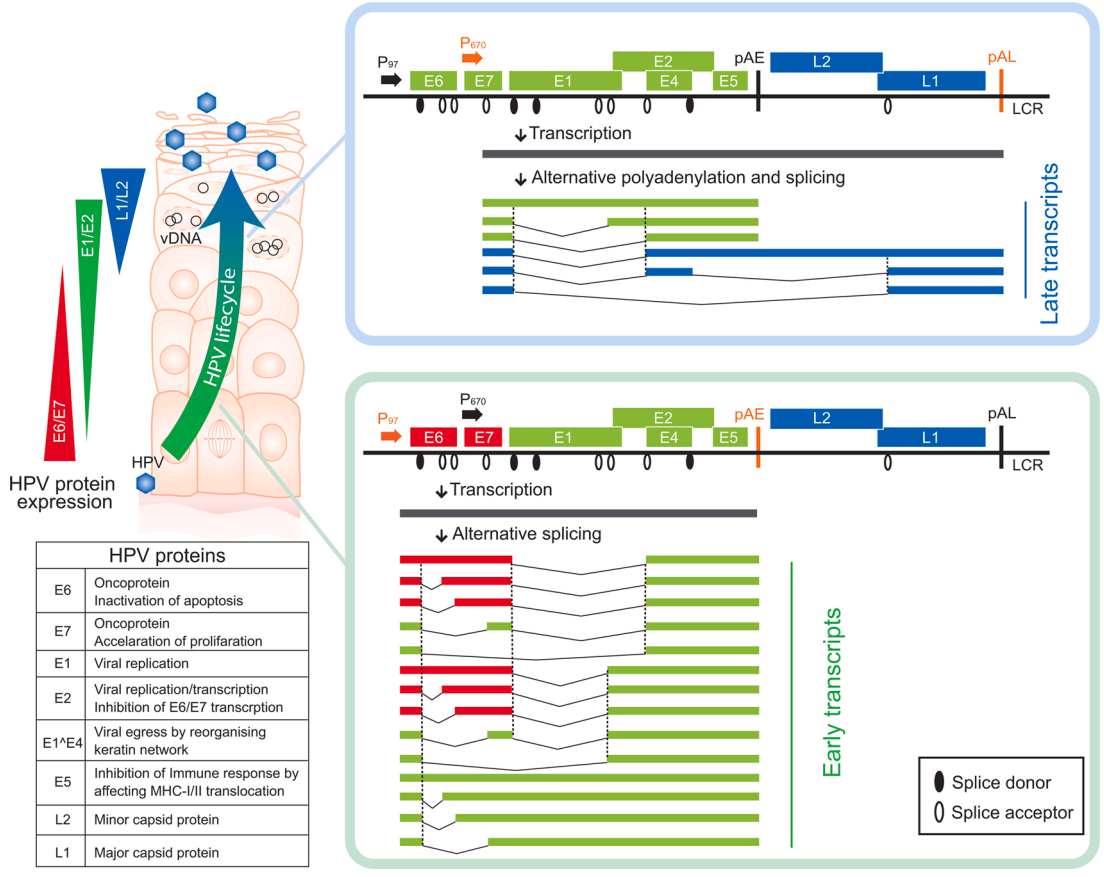


Fig. 1. Human papillomavirus type 16 (HPV16) genome.
Upper part: The circular HPV16 DNA genome. P_{97}: HPV16 early promoter. P_{670}: HPV16 late promoter. pAE: HPV16 early polyadenylation site. pAL: HPV16 late polyadenylation site. Early genes indicated green, late genes in blue and the long control region (LCR) in orange. **Lower part:** HPV16 genome linearized for clarity (numbers refer to the HPV16 reference strain GeneBank: K02718.1). Early genes indicated green and late genes in blue. P_{97}: HPV16 early promoter. P_{670}: HPV16 late promoter. Black oval: 5′-splice site/splice donor. White oval: 3′-splice site/splice acceptor. pAE: HPV16 early polyadenylation site. pAL: HPV16 late polyadenylation site. LCR: long control region (orange).



HPV proteins	
E6	Oncoprotein Inactivation of apoptosis
E7	Oncoprotein Accelaration of prolifaration
E1	Viral replication
E2	Viral replication/transcription Inhibition of E6/E7 transcrption
E1^E4	Viral egress by reorganising keratin network
E5	Inhibition of Immune response by affecting MHC-I/II translocation
L2	Minor capsid protein
L1	Major capsid protein

Fig. 2. Cell differentiation-dependent gene expression of HPV16.
Left upper part: HPVs have adapted to stratified squamous epithelium. The HPV life cycle is strictly linked to cell differentiation. HPV genomic DNA replication occurs with the help of HPV E1 and E2 proteins and the cellular DNA polymerase in basal and suprabasal layers of the epithelium. Expression of the late L1 and L2 capsid genes in terminally differentiated cells immediately followed by de novo synthesis of HPV virions. **Left lower part:** The main function of each HPV protein is listed below the epithelium. **Right upper part:** Schematic representation of the HPV16 genome. Early and late genes are indicated. P_{97}: HPV16 early promoter. P_{670}: HPV16 late promoter. Black oval: 5′-splice site/splice donor. White oval: 3′-splice site/splice acceptor. pAE: HPV16 early polyadenylation site. pAL: HPV16 late polyadenylation site. LCR: long control region. Early genes indicated green and late genes in blue. Schematic representation of the HPV16 late pre-mRNA initiated at the late promoter P_{670} and polyadenylated either at pAE or pAL. Alternatively spliced late transcripts are indicated. **Right lower part:** Schematic representation of the HPV16 early pre-mRNA initiated at the early promoter P_{97} and polyadenylated at pAE. Alternatively spliced early transcripts are indicated. Black oval: 5′-splice site/splice donor. White oval: 3′-splice site/splice acceptor.

are asymptomatic and persist for 12–24 months in the infected host, after which they are cleared by host immune system [33]. In rare cases, the immune system fails to clear the HPV16 infections. In these cases, HPV16 gene expression may be dysregulated with the HPV16-infected cells apparently getting stuck in the early stage of viral life cycle defined by mitosis and HPV16 DNA replication. These cells have reduced ability to transit to the late, productive stage of the viral life cycle since this final stage of the viral life cycle is totally dependent on

terminal cell differentiation. Of particular interest is that HPV16 infected cells that persist, do not express the L1 and L2 genes [35], since the mitotic cells that carry the HPV16-infection are non-permissive for L1 and L2 expression. It therefore appears that absence of L1 and L2 expression is a prerequisite for persistence of HPV16 infections in the immunocompetent host. Any HPV16-infected cell that would prematurely express L1 and L2 genes is likely to be cleared by the immune system of the host, precluding persistence. Simultaneously, dysregulated continuous and increased expression of HPV16 E6 and E7 genes drives cell mitosis and promotes cell survival and contributes to genetic instability and accumulation of mutations in cells that can no longer activate cell cycle arrest or apoptosis. As a result of continuous cell proliferation, the HPV16 driven cells will increase in number and may cause macroscopic premalignant cervical intraepithelial lesions [36]. Over time, these cells will occupy a larger part of the epidermal space and may progress to more severe, high-grade premalignant lesions that constitute a risk factor for development of cervical cancer. If high-grade cervical lesions driven by high-risk HPV types, e.g. HPV16, are left untreated, they may progress to malignancy. Therefore, the HPV16 E6 and E7 proteins have been termed viral oncogenes [37]. In contrast, the HPV16 E2 protein indirectly plays a role in oncogenesis since it has an inhibitory effect on the expression of E6 and E7 proteins [38]. During a productive HPV infection E2 accumulates to reach a critical concentration that eventually shuts down the HPV16 early promoter P_{97} promoter that is essential for E6 and E7 expression. In its capacity as transcriptional repressor, E2 will counteract mitosis driven by E6 and E7, and may induce apoptosis [39,40]. Indeed, E2 is also considered proapoptotic due to its ability to counteract the anti-apoptotic E6 protein by inducing caspase-8 oligomerization [41]. It is reasonous therefore, that the E2 gene is frequently, although not invariably inactivated in HR-HPV infected malignant cells. This is the consequence of preferential integration of the HPV16 episomal DNA genome into cellular chromosomes in the E2 gene [42,43]. The E8ˆE2 protein produced from an mRNA spliced between HPV16 SD1302 and SA2709 is an even greater inhibitor of the HPV16 early promoter and thereby inhibiting expression of all genes under control of the early promoter, apparently causing a general reduction of HPV16 DNA replication [44]. Combined, dysregulated expression of HPV16 E6 and E7 oncoproteins, absence of the pro-apoptotic HPV16 E2 protein and the highly immunogenic HPV16 late L1 and L2 proteins are hallmarks of the HPV16-driven cancer cells. These properties contribute to immune escape and malignant progression of HPV16-infected cells during HPV16 persistence.

3.4. HPV E6 and E7: oncoproteins

HPV E6 and E7 proteins together deregulate crucial cellular pathways including the cell cycle, apoptosis, DNA repair response, senescence and differentiation [26,27,37,45]. The effects of E6 and E7 on the cells is brought about by the binding of the E6 and E7 proteins to multiple cellular proteins with key roles in these pathways [37]. The cellular target proteins may differ between E6 and E7 proteins derived from different HPV genera. The E7 protein from HPV of all known genera binds to members of the pocket protein family including key cellular tumor repressor protein pRb and liberates members of E2F family of transcription factors that govern transcription of a myriad of genes involved in the S-phase cell cycle transition [26,37,45]. As a consequence, the cellular DNA polymerase that catalyzes HPV genomic DNA replication is turned on in HPV-infected cells. Interestingly, E7 protein derived from HR-HPV types binds with higher affinity to members of the pocket protein family than E7 from LR-HPV types does. Furthermore, only E7 proteins derived from high-risk HPVs E7 degrade pRb, indicating that this activity is significant for the ability of HR-HPVs to cause cancer.

HPV DNA replication and unscheduled cell proliferation induced by HPV could potentially cause cell-cycle arrest and even apoptosis of the infected cells, but this is prevented by HPV E6 protein that targets the cellular p53 tumor suppressor protein. E6 proteins derived from high-risk HPVs have the ability to induce proteasomal degradation of p53 [27]. HR-HPV-derived E6 proteins also have a PDZ-binding motif and can bind to and degrade PDZ-proteins [27,37,45]. The PDZ proteins belong to a large family of proteins that is characterized by the presence of the PDZ-homology domain. The interactions between HR-HPV E6 proteins with the PDZ-proteins disrupt the polarity of the epithelial cells thereby disturbing the balance of symmetric and asymmetric epithelial cell division to regulate viral DNA replication [46]. HR-HPV E6 proteins interact with the PDZ-proteins including a number of cell signaling regulators such as 14-3-3, and core polarity proteins such as DLG1, Scrib, MAGI 1/2/3 or PAR3 [47]. Cell polarity plays a crucial role in the organization of signaling pathway to control epithelial cell proliferation, metabolism, apoptosis, differentiation and motility [46]. The interaction of PDZ proteins with HR-HPV E6 is believed to create an environment that is favorable for viral DNA replication. Indeed, the PDZ-binding motif of E6 is required for the replication of viral DNA [48] and the episomal maintenance [49]. Furthermore, HR-HPV E6 protein activates transcription of the human telomerase reverse transcriptase (hTERT) gene by interfering with the transcriptional repressor NFX-91 and by activating the Myc transcriptional activator [50,51]. hTERT encodes the catalytic subunit of the telomerase complex and therefore contributes to telomere length maintenance and infinite proliferation of the HR-HPV-infected cells.

HPV E6 and E7 have additional properties mediated via interactions with a vast number of cellular proteins. One of the more significant properties of E6 and E7 is their contribution to evasion of the host immune system by either transcriptionally inactivating the expression of immune response genes, including both innate and adaptive immune modulators [52], or interfering with factors in the cGAS/STING- and JAK/STAT-pathways [53]. It is worth noting that the more elusive HPV E5 protein also contributes to immune evasion by down-regulating MHC-I and to cell proliferation by enhancing dimerization of the epidermal growth factor receptor molecules [28]. Taken together, E6 and E7 promote proliferation of the HPV-infected cell to create an intracellular environment that facilitates HPV DNA replication, while simultaneously preventing apoptosis and contributing to evasion of cellular immune responses [45]. All as part of a successful completion of the productive HPV life cycle. The vast majority of the HPV infections will be cleared by the host within 12–24 months [33]. However, in the rare event of persistence beyond this time-period, the same properties of E6 and E7 may contribute to genetic instability and transformation of the HR-HPV infection to malignancy. Of particular interest for this review is that HR-HPV E6 and E7 mRNAs are produced from different alternatively spliced mRNAs generated in a mutually exclusive manner from the same pre-mRNA produced from the HPV early promoter (Fig. 2). Thus, regulation of HPV mRNA splicing is of paramount importance during the HPV life cycle as well as during malignancy, since both E6 and E7 are essential for HPV replication as well as for the induction and maintenance of HPV-induced malignancy [26,27,37].

3.5. HPV E2: a DNA replication factor that also controls E6 and E7 gene expression

HPV E1 and E2 proteins are committed to replication of the HPV DNA genome [5,23,30]. HPV E2 protein binds to HPV DNA and to HPV E1, which enhances the association of HPV E1 protein and HPV DNA. The E1 and E2 proteins together promote the formation of HPV DNA replication complexes, recruiting the cellular DNA polymerase to the HPV DNA and promoting efficient HPV DNA replication. The HPV E1 DNA helicase protein is the only HPV protein that possesses enzymatic activity and acts by unwinding the double-stranded DNA of the HPV genome to facilitate HPV genome replication [23]. The E2 protein also plays a part in the segregation of equal numbers of HPV genomes into the daughter cells by tethering HPV genomes to cellular chromosomes [54]. This property of E2 contributes to HPV persistence and the

increase in the number of HPV-infected cells at the site of HPV infection.

The E2 protein also serves as a transcriptional regulator, activating HPV transcription at low levels and repressing HPV transcription at high levels, by binding to E2-binding sites in the HPV LCR [6]. As the E2 protein accumulates to higher and higher levels through cellular differentiation (Fig. 2), E2 represses the HPV early promoter by competing with essential basal transcription factors for HPV DNA, thereby shutting down the HPV early promoter and E6 and E7 gene expression [38]. In the HPV life cycle, E2 would reduce the pro-mitotic effects mediated by E6 and E7 to secure a return to terminal cell differentiation and subsequent activation of late L1 and L2 expression and virus production. In sharp contrast, in persistently HPV-infected cells at the verge of transformation, E2 could be an obstacle to transformation and malignancy, and the E2 gene is often destroyed by integration of the HPV16 genome in the E2 gene into cellular chromosomes [55]. The E8ˆE2 protein strongly suppresses HPV16 early transcription and as such secures an efficient shut down of HPV16 early gene expression [44]. Integration is not an absolute requirement for cell transformation and malignancy caused by HPV [56], indicating that E2 levels can be controlled at the level of gene expression, which in essence would be at the level of RNA processing since HPV E2, E6 and E7 would presumably all be expressed from the HPV early promoter in HPV-driven, transformed cells. In addition to regulation of gene expression at the level of RNA splicing and polyadenylation, the methylation status of the E2 binding sites in the HPV early promoter may also affect expression of the HPV early genes [57].

3.6. HPV L1 and L2: strongly immunogenic antigens with expression restricted to terminally differentiated keratinocytes

HPV L1 and L2 proteins are viral structural capsid proteins that are produced exclusively in terminally differentiated cells (Fig. 2) [32,33,58]. Terminal cell differentiation is required for induction of L1 and L2 gene expression. The delay in L1 and L2 protein production to the uppermost cells in the squamous epithelium likely serves two purposes: 1) L1 and L2 proteins will be produced when HPV infected cells contain enough of HPV DNA genomes to secure efficient virus production and 2) restricted expression of the highly immunogenic L1 and L2 proteins to the uppermost cells in the epithelium minimizes exposure of HPV to the host immune system. It is reasonable to speculate that this restricted expression of L1 and L2 contributes to the ability of HPV to persist in the presence of the immune surveillance. Indeed, L1 proteins are highly immunogenic and L1 produced in eukaryotic cells self-assemble into icosahedral virus-like particles that are the source of the highly effective prophylactic vaccine to HPV [59]. Taken together, the HPV gene expression program is a highly orchestrated process that efficiently controls the spatiotemporal expression of each HPV gene in the differentiating environment of the multilayered squamous epithelium (Fig. 2). If HPV gene expression is dysregulated and HPV fails to control expression of each HPV gene, HPV infected cells would either be rapidly cleared by immune surveillance or they could persist and progress to a state of uncontrolled cell growth and genetic instability, undetected by the immune system and with a high risk of progression to cancer. Thus, a better understanding of the control of HPV gene expression will enhance our understanding of the HPV life cycle and pathogenesis and may uncover novel targets for therapy to HPV-infections and to HPV-driven cancer. Due to the low number of promoters on the HPV genome, RNA processing plays a significant role in the HPV gene expression program and cellular RNA-binding proteins are key players [60].

4. Expression of HPV genes is meticulously controlled at the level of RNA processing

Structures of alternatively spliced transcripts of different types of papillomaviruses have been determined and compiled with big efforts by different laboratories [11]. The combined results have revealed that the fundamental structure of the HPV transcriptome is similar among different HPVs, namely that the HPV genes share a restricted number of major promoters, often one early and one late promoter, followed by one early and one late polyadenylation site (Fig. 1). HPV mRNAs initiated at these promoters and terminated at the HPV polyadenylation signals are alternatively spliced to a great extent to generate a variety of transcripts at each stage of the HPV life cycle (Fig. 2). The spatiotemporal HPV gene expression program consists primarily of three fundamental series of events: an HPV promoter switch, an HPV polyadenylation switch and HPV RNA processing regulation, including alternative mRNA splicing (Fig. 2). Promoter switch and polyadenylation switch are essential for the transition of HPV life cycle from early to late stage, whereas alternative mRNA splicing regulation contributes to production of a variety of HPV transcripts at each stage of the HPV gene expression program (Fig. 2) [60–63]. The RNA splicing machinery is extensively utilized for alternative splicing of HPV mRNAs with examples of alternative 3′-splice sites, alternative 5′-splice sites, exon skipping and exon retention (Fig. 2). For example, transcription of the HPV16 early genes during the early stage of the HPV16 life cycle, including the E6 and E7 oncogenes, is initiated at the early promoter (P_{97}) located in the Long Control Region (LCR) (Fig. 2) and is terminated at the early polyadenylation signal (pAE) located between the E5 and the L2 genes (Fig. 2). As a result, all HPV16 early mRNAs, including E6 and E7 mRNAs, are generated from the same precursor mRNA (pre-mRNA), followed by mRNA alternative splicing. The arrangement on the HPV16 genome of the E6 and E7 genes is of particular interest. While the HPV16 E7 protein is produced from an mRNA that is spliced in the upstream E6 coding region, E6 is translated from and mRNA with a retained intron and this intron is coding for E6 (Fig. 3A and B). As both mRNAs are derived from the early promoter, the E6 and E7 mRNAs are produced in a mutually exclusive fashion from the same pre-mRNA. Thus, splicing efficiency will determine the relative levels of the E6 and E7 mRNAs as well as the relative levels of the E6 and E7 proteins. Several research groups have independently presented results that support the conclusion that E7 protein is produced primarily from the spliced mRNA [64–67]. E6 and E7 are functionally entangled during the HPV16 life cycle as well as during cell transformation [37]. Both E6 and E7 are expressed in cancer cells derived from HPV infections depend on both proteins to maintain the carcinogenic state of these cells. Thus, meticulous control of this splicing event is paramount to generate sufficient quantities of each mRNA and to maintain the optimal ratio of E6 to E7 protein in HPV-infected cells and in cancer cells.

Transcription during the HPV16 late life cycle initially requires a switch of promoter, i.e. downregulation of the early promoter (P_{97}) and activation of the late promoter (P_{670}) [68,69] (Fig. 2), similarly to various HPVs [58,70,71]. The major advantage with the late promoter is the bypassing of E6 and E7 genes and the ability to create mRNAs on which the E1- or the E2- open reading frames are favorably positioned appearing as first major ORFs on their respective mRNA, thereby allowing enhanced E1 and E2 production, which is likely to be necessary for vegetative HPV DNA replication (Fig. 2). Similarly to E6 and E7, production of the E1- and E2-encoding mRNA is a mutually exclusive event in which E1 and E2 are generated from the same pre-mRNA initiated from the late promoter (Fig. 3A and C) [11]. As the E2-encoding mRNA is generated by splicing, a splicing event that removes the E1 coding region, the E1 mRNAs are generated by retention of the E1-encoding intron to produce unspliced E1 mRNAs from the same pre-mRNAs as the E2 mRNAs (Fig. 3A and C) [11]. Since E1 and E2 proteins function together to replicate HPV16 DNA, a balanced production of the mutually exclusively produced unspliced E1 mRNAs and spliced E2 mRNAs is of utmost importance.

As a consequence of cell differentiation there is not only a shift in HPV promoter usage, there is also a shift in HPV polyadenylation usage, from polyadenylation primarily at the HPV early polyadenylation signal (pAE) to polyadenylation at the downstream located late polyadenylation signal (pAL) (Figs. 2 and 4) [32,33,58]. Of note, these

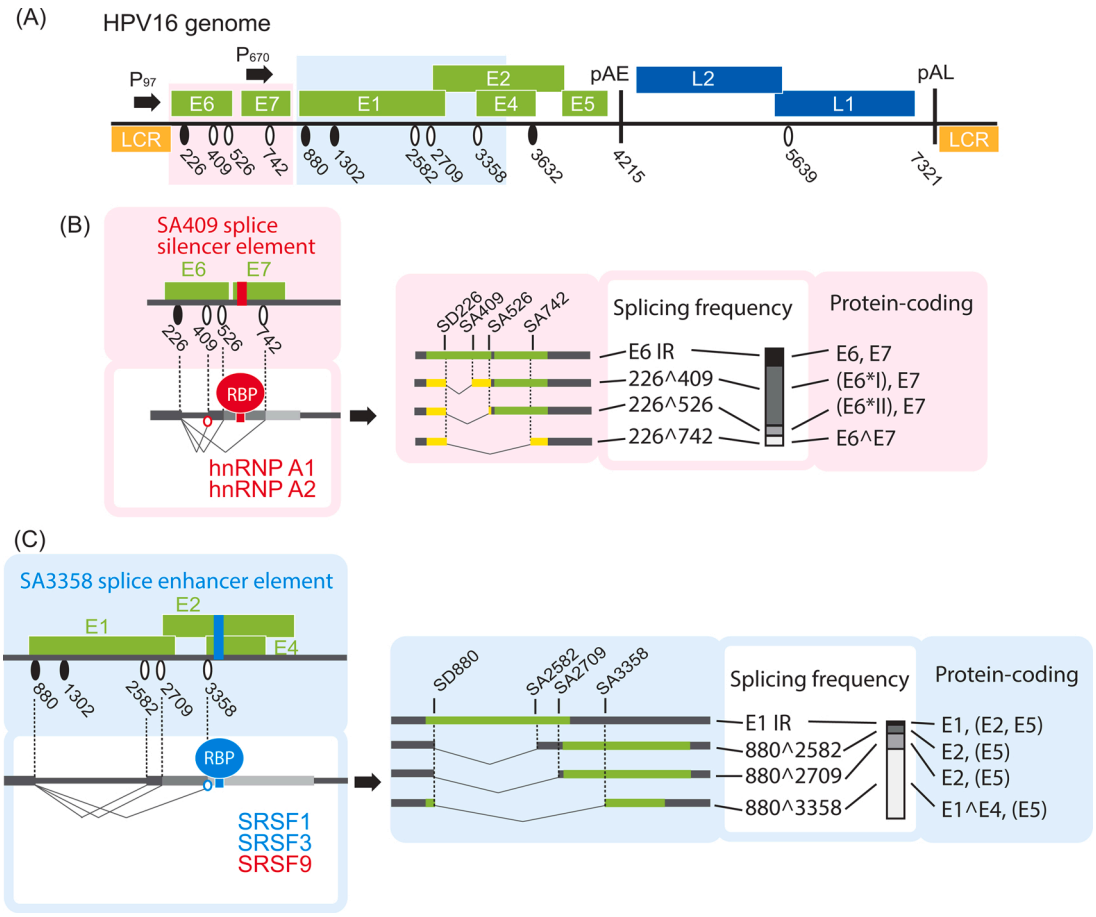


Fig. 3. Cis-acting HPV16 RNA elements and trans-acting cellular factors that control HPV16 early mRNA splicing.
(A) HPV16 genome linearized for clarity (numbers refer to the HPV16 reference strain GeneBank: K02718.1). Early genes indicated in green, late genes in blue and the long control region (LCR) in orange. Early and late genes are indicated. P_{97}: HPV16 early promoter. P_{670}: HPV16 late promoter. Black oval: 5′-splice site/splice donor. White oval: 3′-splice site/splice acceptor. pAE: HPV16 early polyadenylation site. pAL: HPV16 late polyadenylation site. **(B) Left:** Blow up of the region encoding HPV16 E6 and E7 oncogenes. E6 and E7 coding sequences are in green. Splice sites used for alternative splicing in the E6/E7 coding region are indicated (226, 409, 526 and 742). Thin black lines represent various splicing alternatives utilized to create a subset of alternatively spliced mRNAs from this region. Red box represents a splicing-regulatory RNA-element that interacts with RNA-binding proteins (RBPs) hnRNP A1 and hnRNP A2 that control the indicated HPV16 splice sites. **Right**: Schematic representation of the various HPV16 alternatively spliced mRNAs generated from the E6/E7 coding region. Intact E6 and E7 open reading frames are indicated in green various the shorter, E6-derived protein coding sequences E6*I, E6*II and E6^E7 are indicated in yellow. Splice sites in this region are indicated (SD226, SA409, SA526 and SA742). Splice sites used to generate the various alternatively spliced mRNAs are indicated to the right: 226^409, 226^526 and 226^742. mRNAs with retained introns encode full-length E6 (E6 IR). The abundance of each alternatively spliced mRNAs in HPV16-infected cells is schematically represented by the various white/grey/black areas in the vertical bar to the right. Proteins potentially produced from each mRNA is indicated to the right. **(C) Left:** Blow up of the region encoding HPV16 E1, E2 and E1^E4 proteins. E1, E2 and E4 coding sequences are in green. Splice sites used for alternative splicing in this region are indicated (880, 1302, 2582, 2709 and 3358). Thin black lines represent various splicing alternatives utilized to create a subset of alternatively spliced mRNAs from this region. Blue box represents a splicing-regulatory RNA-element that interacts with RNA-binding proteins (RBPs) SRSF1, SRSF3 and SRSF9 that control the indicated HPV16 splice sites. **Right**: Schematic representation of the various HPV16 alternatively spliced mRNAs generated from the E1, E2 and E4 coding regions. E1, E2 and E4 open reading frames are indicated in green. Splice sites in this region are indicated (SD880, SA2582, SA2709 and SA3358). Splice sites used to generate the various alternatively spliced mRNAs are indicated to the right: 880^2582, 226^2709 and 226^3358. mRNAs with retained introns encode full-length E1 (E1 IR). The abundance of each alternatively spliced mRNAs in HPV16-infected cells is schematically represented by the various white/grey/black areas in the vertical bar to the right. Proteins potentially produced from each mRNA is indicated to the right.

switches are dependent on the cellular differentiation status and differentiation-specific trans-acting factors, although the detailed mechanisms remain unclear [60]. This switch in polyadenylation leaves open the possibility of generating HPV late L1 and L2 encoding mRNAs. By polyadenylation at the HPV16 late polyadenylation signal pAL and by usage of splice sites that are exclusively utilized for L1 mRNAs, i.e. SD3632 and SA5639 (Fig. 5). Similarly to E6- and E7-mRNAs and E1- and E2-mRNAs that are produced in a mutually exclusively fashion from alternatively spliced mRNAs, L1 and L2 mRNAs are also generated from the same pre-mRNA in a mutually exclusive manner (Fig. 5) [11]. While L1 mRNAs are generated by splicing, using HPV late-mRNA specific splice sites SD3632 and SA5639, the L2 ORF is encoded by an mRNA generated by intron retention (Fig. 5). The intron in this case encodes the L2 ORF.

In conclusion, regulation of HPV alternative splicing is vital to HPV for the following reasons: 1) HPV pre-mRNAs are polycistronic and encode multiple overlapping viral open reading frames, 2) HPV mRNAs are translated by 5′-cap dependent mechanism for translation. HPV mRNAs are capped and polyadenylation by cellular machineries. The cellular translation machinery is recruited to the 5′-end of the mRNA in a CAP-dependent manner after which the ribosome commences a scanning movement across the mRNA in a 5′-to-3′-direction in search for translational start codons in optimal context [65,66,72]. Translation initiation at the E6 ATG is special in the sense that the ATG is very close to the 5′-end of the mRNA and the CAP structure. Efficient recognition of the HPV18 E6 ATG therefore appears to be dependent on a unique RNA

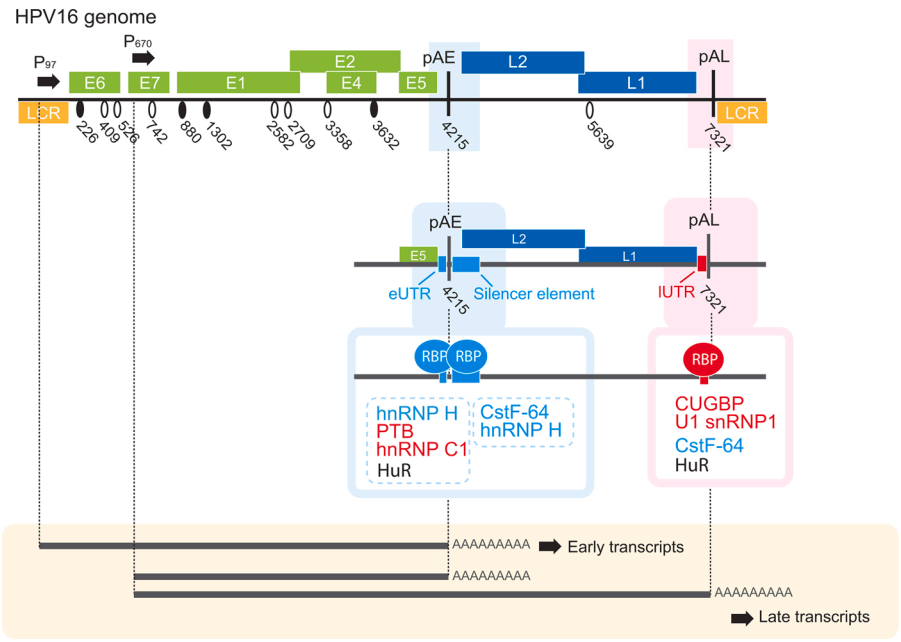


Fig. 4. Trans-acting cellular factors that interact with cis-acting HPV16 RNA elements at the early and late HPV16 polyadenylation signals.
HPV16 genome linearized for clarity (numbers refer to the HPV16 reference strain GeneBank: K02718.1). Early genes indicated in green, late genes in blue and the long control region (LCR) in orange. Early and late genes are indicated. P_{97}: HPV16 early promoter. P_{670}: HPV16 late promoter. Black oval: 5′-splice site/splice donor. White oval: 3′-splice site/splice acceptor. pAE: HPV16 early polyadenylation site. pAL: HPV16 late polyadenylation site. Below the schematic representation of the HPV16 genome, the region spanning the early polyadenylation signal (pAE), the late region and the late polyadenylation signal (pAL). RNA elements upstream and downstream of pAE are indicated in blue and the RNA element upstream of pAL is indicated in red. Various RNA-binding proteins (RBPs) that interact with the indicated HPV sequences are listed below each HPV RNA element. Bottom part of the figure: HPV16 early pre-mRNA initiated at the early promoter $P_{97}$and polyadenylated at the early polyadenylation signal pAE, HPV16 late pre-mRNA initiated at the late promoter P_{670} and polyadenylated at the early polyadenylation signal pAE and HPV16 late pre-mRNA initiated at the late promoter P_{670} and polyadenylated at pAL.

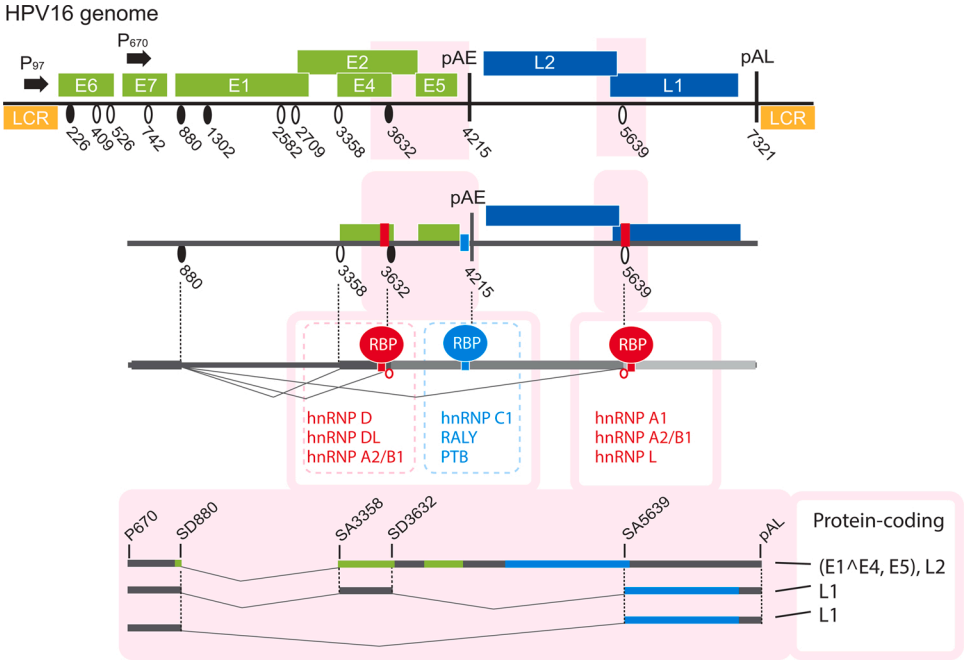


Fig. 5. Cis-acting HPV16 RNA elements and trans-acting cellular factors that control HPV16 late mRNA splicing.
Top: HPV16 genome linearized for clarity (numbers refer to the HPV16 reference strain GeneBank: K02718.1). Early genes indicated in green, late genes in blue and the long control region (LCR) in orange. Early and late genes are indicated. P_{97}: HPV16 early promoter. P_{670}: HPV16 late promoter. Black oval: 5′-splice site/splice donor. White oval: 3′-splice site/splice acceptor. pAE: HPV16 early polyadenylation site. pAL: HPV16 late polyadenylation site. **Middle:** Below the HPV16 genome, a schematic representation of the HPV16 late pre-mRNA initiated at the late promoter P_{670} and polyadenylated at the late polyadenylation signal pAL. Selected splice sites used by the major late mRNAs are indicated (SD880. SA3358, SD3632 and SA5639). Thin black lines represent various splicing alternatives utilized to create a subset of alternatively spliced late mRNAs. mRNAs with retained introns will encode either E1 or L2. Red boxes represent splicing regulatory RNA elements that interact with RNA-binding proteins (RBPs) that control adjacent splice sites. Various RNA-binding proteins that interact with the HPV RNA sequences are listed below each RNA element. **Bottom:** The coding potential of a subset of alternatively spliced HPV16 late mRNAs is shown. E1^E4 and E5 open reading frames are indicated in green and L1 and L2 open reading frames in blue. Schematic representation of the various HPV16 alternatively spliced late mRNAs generated encoding L1 and L2. E4 and the E5 open reading frames present on the L2 encoding mRNAs are indicated in green and the L1 and L2 coding sequences are in blue. Splice sites utilized to generate these late mRNAs are indicated (SD880, SA3358, SD3632 and SA5639). Proteins potentially produced from each mRNA is indicated to the right.

sequence in the 5′-end of these mRNAs due to the close proximity of the E6 ATG to the 5′-end of the mRNA [73].Efficiency of translational initiation of other translational start codons on the HPV mRNAs is likely to be determined by the context of the ATG in the same manner as on cellular mRNAs [74]. As ribosomes generally leave the mRNAs as they translate a stop codon, mRNAs translated in human cells generally produce only one protein each. For HPV with multiple overlapping open reading frames and only two major promoters, alternative splicing is essential for production of at least one unique, alternatively spliced mRNA from which each papillomavirus protein can be translated with appropriate efficiency. Thus, alternative mRNA splicing is paramount for production of HPV mRNAs with the ability to express the entire HPV proteome. Regulation of HPV alternative splicing combined with transcriptional control therefore determine when each HPV protein is

expressed during the HPV life cycle. Furthermore, HPV-infected cells that progress to malignancy need to maintain an intracellular environment that allows for efficient production of both E6 and E7. Since the relative levels of E6 and E7 proteins is determined largely by the alternative splicing in the E6 and E7 coding regions, it is reasonable to assume that cellular RNA-binding proteins and alternative splicing play important roles during induction of cancer by HPV.

5. RNA-binding proteins contribute to HPV gene expression

5.1. Exons and introns on HPV mRNAs

The regulation of HPV mRNA alternative splicing is complex. For instance, HPV16 protein coding regions occupy 92 % of the HPV16 genome. Apparently, there is a significant difference of exon/intron organization between HPV and the cellular genome. For example, the average exon and intron lengths of HPV are 1004 and 2285 base pairs, respectively, while exon and intron lengths of mammalian cells are 160 and 6938 base pairs, respectively [75]. Consequently, cis-regulatory RNA elements that control HPV RNA processing are more condensed than on cellular mRNAs, predicting that trans-acting proteins may interact more frequently with HPV mRNAs than with cellular mRNAs. Indeed, a number of hot spots for RNA binding proteins on HPV16 mRNAs have been identified [76]. The high exon-occupancy and overlap with protein coding regions in the majority of the genes on the HPV16 genome renders the definition of exons and introns somewhat problematic. There are simply no sequences on the HPV16 genome that are located between the early promoter and the late polyadenylation site that are invariably spliced out and are never present on at least one of the many alternatively spliced, mature HPV16 mRNAs (Figs. 1 and 2). Thus, all HPV16 sequences between the early promoter and the late polyadenylation site could be defined as exonic, but in the view of individual mRNAs, selected sequences will have to be spliced out as introns to generate at least one HPV16 mRNA that can be efficiently translated to each of the viral proteins. Some protein coding regions on the HPV genome could therefore also be defined as introns.

Exon or intron definition is the primary step of mRNA alternative splicing in which RNA-binding splicing factors (repressors or activators), initiate splice site recognition by binding with cognate factors to cis-regulatory RNA target elements adjacent to splice sites. This is of particular importance if splice donors and acceptors conform poorly to the consensus sequences for splice sites, which is true for the majority of the HPV splice sites [60]. These RNA elements are better known as exonic splicing enhancers (ESE), exonic splicing silencers (ESS), intronic splicing enhancers (ISE) and intronic splicing silencers (ISS) [77,78]. RNA-binding proteins that recognize the splicing regulatory elements will define exons and/or introns and recruit the core splicing machinery to execute splicing with relatively high precision at defined splice sites. Splicing regulatory RNA elements on papillomaviruses were initially identified in bovine papillomavirus [63]. On the HPV pre-mRNAs, the vast majority of the splicing regulatory RNA elements will overlap with protein coding sequences, and in at least one case these RNA elements overlap with two overlapping protein coding sequences (E2 and E4) (Figs. 3 and 5). Thus, the HPV protein coding sequences may inflict restrictions to the primary sequence of the splicing regulatory RNA elements on HPV. However, it has been shown in at least two cases that HPV RNA regulatory elements could be inactivated without affecting the protein coding sequence of the overlapping ORF [79–81]. Splicing silencers overlapping the HPV16 L1 coding region were inactivated by mutations that left the L1 protein sequence unaffected [79,80] and polyadenylation-enhancing RNA elements overlapping the HPV16 L2 coding region were inactivated by mutations that did not alter the L2 protein sequence [81]. However, in addition to recognizing the RNA primary sequence, the secondary structure of the RNA is also a critical factor for RNA recognition by RNA-binding proteins, adding a dimension to the complexity of the cis-acting RNA elements. The ability of RNA-binding proteins to identify key regulatory RNA elements may be enhanced by the association of RNA-binding proteins with RNA polymerase II or by recruitment to chromatin or by specific DNA-binding proteins [82,83,84,85,86,87,88,89). The latter may be applicable to HPV, for which recent results suggest that CCCTC-binding factor (CTCF) binds to HPV DNA and alters HPV mRNA splicing [90]. Finally, recognition of the cis-acting RNA elements may be affected by reversible chemical modifications of individual nucleotides in an RNA element such as the methylated adenosine m6A [91–93]. Combined, these properties of HPV mRNAs may contribute to the successful utilization of alternative splicing and recruitment of regulatory RNA-binding proteins to control HPV gene expression in a spaciotemporal manner, at the level of RNA processing.

5.2. Alternative mRNA splicing within the HPV E6 and E7 coding region

The E6 and E7 genes are transcribed only from the unique HPV early promoter. In HPV16, the HPV16 early promoter P_{97} has a transcriptional start site at nucleotide position 97 followed by the E6 and E7 coding regions. HPV16 mRNAs produced from P_{97} contain a first splice site at position 226, a splice donor site named SD226, followed by three splice acceptor sites named SA409, SA526 and SA742. Additional minor or cryptic splice sites have been described [94,95]. Thus, four splicing outcomes are possible within this region: E6 intron retention (E6 IR) that generates the mRNA encoding full-length E6, 226ˆ409-, 226ˆ526- and 226ˆ742-splicing (Fig. 3A and B). Not all HPV types have these four splicing events within the E6/E7 region. As a matter of fact, most HPVs do not, but all HPVs produce mRNAs with retained E6-encoding intron and at least one spliced mRNA corresponding to HPV16 226ˆ409-mRNA [96]. The latter mRNA is also referred to as the “E6*I”- transcript since it produces a splicing-maimed E6 protein termed E6*I [11]. However, the E6*I”- transcript may also be the major E7 mRNA since E6*I appears to be less of an obstacle than the full-length E6 ORF to translation initiation at the E7 ORF [65]. Interestingly, E6*I-transcripts are unique to high-risk HPV types, since low risk HPVs express E6 and E7 from two independent promoters [97], and as such are diagnostic for high-risk HPV infections [98]. Furthermore, it is worth to note that there is a difference in splice site utilization frequency among the E6/E7 splice sites. The 226ˆ409-spliced mRNA is more abundant than any of the other alternatively spliced mRNAs including the E6 mRNA with retained intron in HPV16-infected cervical cancer cells including premalignant lesions and tumors [97], as well as in HPV16-positive cervical cancer cell line CaSki [97] and HPV16-positive tonsillar cancer cell line HN26 [99].

It has been difficult to ascertain which of the alternatively spliced E6/E7 transcripts that is the major E7 producing mRNA [64–67,100]. With regard to coding potential, the E6 intron-retained transcript has the potential to be translated to both E6 and E7, whereas the E6*I (226ˆ409) and the E6*II (226ˆ526) transcripts have the potential to produce either E6*I or E6*II and E7 (Fig. 3B). In contrast, the E6ˆE7 (226ˆ742) transcript is incapable of producing E6 and E7, but encodes an E6-E7 fusion protein named E6ˆE7. However, less is known about the E6*I, E6*II or E6ˆE7 proteins than about E6 and E7 [101–104]. Although they may be of functional significance, it is of interest that the vast majority of the HPV risk HPV types produced only mRNAs corresponding to the HPV16 E6 mRNA with retained intron and the HPV16 E6*I mRNA spliced between SD226 and SA409 [96]. The primary functional significance of the E6*I transcript may also be to secure efficient translation and production of the E7 protein since the short upstream E6*I ORF should be less of an obstacle for translation initiation at the E7 ATG than the full-length E6 ORF according to the scanning model for translation initiation [74]. Although there is ample evidence that the E7 protein is translated primarily from mRNAs that are spliced in the E6 coding region, i.e. the E6*I mRNAs [64–67], it has also been suggested that HPV16 mRNAs with retained E6-encoding intron can produce significant quantities of an E7-GFP fusion protein [100]. The significance of E7

protein production from the E6-encoding mRNAs remains to be determined. Finally, E6 intron retention and E6*I (226ˆ409)-, E6*II (226ˆ526)- and E6ˆE7 (226ˆ742)-splicing events are all mutually exclusive, thereby potentially affecting the balance between the E6 and E7 mRNA and protein levels.

The regulation of E6/E7 alternative mRNA splicing has been investigated. Recently two publications reported the identification of cognate cis-acting, negative RNA elements in the E7-coding region of HPV16 [105] and HPV18 [67]. These RNA elements acted as exonic splicing silencers for the E6*I mRNA and they bound directly to cellular RNA-binding protein hnRNP A1. As a consequence of the inhibitory effect on E6*I-mRNA splicing by hnRNP A1, hnRNP A1 enhanced intron retention and promoted production of E6-encoding mRNAs with retained E6-encoding intron (Fig. 3B). These results were similar for HPV16 and HPV18. Interestingly, for HPV16, the closely related hnRNP A2 protein also interacted with the same cis-regulatory element to inhibit splicing between SD226 and SA409, but in contrast to hnRNP A1, hnRNP A2 redirected splicing to the alternative 3′-splice site SA742 rather than promoting intron retention, thereby enhancing production of the E6ˆE7-encoding mRNA (HPV16 226ˆ742) (Fig. 3B) [105]. It has previously been shown that hnRNP A1 and hnRNP A2 affect HPV16 E6/E7 mRNA splicing in an epidermal growth factor dependent manner, indicating that the control of HPV16 E6/E7 mRNA splicing is dynamic and may change during the HPV16 life cycle and/or during carcinogenesis [106]. Indeed, the hnRNP A1 protein levels gradually increase from low grade HPV-positive cervical lesions to high grade lesions and cancer [107]. However, the HPV16 E6*I mRNAs spliced between SD226 and SA409 dominate over the intron-retained E6 mRNAs suggesting that if the levels of hnRNP A1 increase, hnRNP A1 is either inactive in the control of HPV16 mRNA splicing, or that the levels of a yet unidentified cellular factor with enhancing effect of splicing is increasing as well. It has also been shown that knock-down of SR protein SC35/SRSF2 reduced the levels of all E6/E7 mRNAs, but it was unclear if splicing was affected [108]. Although hnRNP A1 and A2 apparently contribute to the control of HPV16 E6/E7 splicing through their splicing inhibitory effect, it would be of great interest to identify RNA elements and trans-acting factors that enhance splicing and contributes to production of the HR-HPV-specific spliced E6*I mRNAs, i.e. the E7 mRNAs.

5.3. Alternative mRNA splicing within the HPV E1/E2 gene region

The HPV E1 and E2 genes are transcribed from both HPV early and late promoters (Fig. 2). In the HPV16 E1/E2 gene region, there are three splice donor sites (SD880, SD1302 and SD3632) and three splice acceptor sites (SA2582, SA2709 and SA3358) [11] (Fig. 3C). Splice sites SD880, SD1302, SA2582 and SA2709 are all located within the E1 ORF whereas SA3358 is located in the E2 ORF (Fig. 3C). HPV16 SD3632 is dedicated to late mRNA splicing and is therefore discussed later. Studies carried out so far on HPV16 indicate that of the two splice donors in this region (SD880 and SD1302), splice donor SD880 is normally more active than SD1302. However, the splicing plasticity of the HPV16 transcriptome encompasses also SD880. Even if SD880 is efficiently utilized to produce the spliced E1ˆE4 mRNAs (880ˆ3358) and spliced E2 mRNAs (880ˆ2709), inactivation of the SD880 is necessary for generation of E1 mRNAs as they are generated by retention of the E1-encoding intron sequence located downstream of SD880. As splicing ligation partners of SD880, splice acceptors SA2582, SA2709 and SA3358 in the E1 and E2 coding regions could potentially all be simultaneously silenced to retain the E1-encoding intron to generate E1 mRNAs, but it is more reasonable to assume that the major splice site SD880 is subject to regulation. Posttranscriptional regulation of E1 and E2 expression has been proposed for HPV31 E1 and E2 expression [109]. This splice site is also required for production of E6 and E7 mRNAs, indicating that the majority of the E7 mRNAs are spliced at SD880 as well [110]. Thus, splice site SD880 has a key role in the regulation of HPV16 gene expression. However, little is known about the control of SD880.

HPV16 splice acceptors SA2582 and SA2709 are located immediately upstream of the E2 start codon within the E1 ORF, thereby being potentially dedicated to E2 mRNA expression (Fig. 3C) [111]. However, SA2582 is inefficiently utilized and it has been shown that mRNAs spliced to SA2582 initiate translation at the E2 ATG less efficiently than mRNAs spliced to SA2709 [105], so the significance of SA2582 remains unknown. It is reasonable to assume that SA2079 is required for production of the essential HPV transcription and replication factor E2. In essence, HPV16 SD880 can choose to splice to either SA2709 to generate E2 mRNAs, or to SA3358 to generate E1ˆE4 mRNAs (Fig. 3C). The two major splice acceptors SA2709 and SA3358 are used in a mutually exclusive fashion and are therefore competing for splicing to the upstream SD880 splice donor. The 880ˆ3358-splicing event generates an in-frame fusion of the E1 ATG and the 5′-end of the E1 ORF encoding the N-terminal five amino acids of E1 located upstream of SD880 to the E4 coding sequence starting at HPV16 nucleotide position 3358. This splicing arrangement and production of an E1ˆE4 fusion protein is a common property of HPVs [11]. This mRNA produces the E1ˆE4 protein, which is one of the most abundantly expressed HPV16 proteins [31]. Several findings suggest that the E1ˆE4 protein has a distinct role in the late stage of the HPV life cycle: E1ˆE4 may promote viral DNA amplification in differentiated cells by inducing G2/M cell cycle arrest and may facilitate viral egress by reorganizing the cellular cytoskeleton network [31]. In cancer cells that carry integrated HPV16 DNA, integration of the HPV16 genome often occurs in such a manner that the E2 gene is destroyed or cannot be expressed [55]. In such cases, splice donor SD880 is often finding a matching splice acceptor in the cellular genome to generate HPV16-cell hybrid mRNAs with the cellular mRNA contributing a splice acceptor and a polyA tail. In such HPV16-driven cancers, HPV16 splice sites SD226, SA409, SA526, SA742 and SD880 are still active and are required for production of the cancer driving HPV proteins E6 and E7. Thus, splicing control of HPV16 E6 and E7 mRNAs extend to these hybrid-mRNAs that are vital for survival of the cancer cell.

So far, regulation of E1/E2 mRNA splicing has been poorly researched. In HPV16 the elusive SD1302- and SA2582-splice sites in the E1 coding region are rarely used and will be challenging to study. HPV16 5′-splice site SD1302 (and its equivalent in HPV31, SD1296) have been mutated and functionally analyzed in strongly indicating that this donor is mainly used to produce the E8ˆE2 RNA [112,113]. Compared to SD1302, SD880 is very frequently used but its regulatory fate is still disclosed. However, it has been shown in RNA-mediated protein pull-down assays that an RNA oligo encoding HPV16 SD880 changed affinity for the U1snRNP component U1-70K if cell extracts were derived from human cells treated with Akt kinase inhibitor, suggesting that interactions of splicing components with SD880 is under the control of the Akt kinase [76]. However, a detailed molecular mechanism is absent and further investigation is needed to determine which and how RNA-binding proteins regulate SD880. Little, if anything is known about the control of the major E2 mRNA splice site SA2709 in HPV16, or in any other HPV type. In contrast, a positive cis-regulatory element immediately downstream of HPV16 SA3358 has been identified [114–116]. Cellular RNA binding proteins including the SR-proteins SRSF1 (ASF/SF2), SRSF3 (SRp20) and SRSF9 (SRp30c) were shown to act as trans-acting factors and control splicing to SA3358 [114,116,117]. A splicing enhancer interacting with SR-proteins is present in a corresponding position in bovine papillomavirus [118]. SRSF proteins belong to the serine-arginine (SR) protein family that are massively phosphorylated and regulated by phosphorylation [119]. Multiple kinases appear to phosphorylate SR proteins and SRPK (SR protein kinase) is one among them [120]. Indeed, the significance of SRPK1-SRSF1 axis in the HPV lifecycle was suggested by the finding that stable overexpression of HPV16 E2 caused an increase in SRPK1 protein levels by an unknown mechanism [121]. E2 has also been shown to activate expression of SRSF1, a target of SRPK1 [122]. The active phosphorylated form of SRSF1 was produced in a keratinocyte

differentiation-specific manner [121] and was counteracted by the E1^E4 protein [123]. In summary, a better understanding of cis- and trans-acting factors that control the major HPV16 splice sites involved in production of E1, E2 and E1^E4 mRNAs would enhance our understanding of HPV replication and its role in carcinogenesis.

5.4. Polyadenylation at HPV pAE or pAL and the role of adjacent RNA elements

There are two major polyadenylation sites for HPV transcripts: the early polyadenylation signal (pAE) located downstream of the E5 gene and the late polyadenylation signal (pAL) located downstream of the L1 gene (Figs. 1 and 2). HPV pAE is highly active in undifferentiated epithelial cells while pAL is strongly inhibited in these cells (Fig. 2). The active pAE restricts transcription-readthrough at pAE and allows transcription only within HPV early gene region and not into the late gene region. Upon differentiation, the readthrough at pAE is significantly increased, thereby reducing polyadenylation at pAE and allowing polyadenylation at pAL (Fig. 2). In order to allow for readthrough into the late L1 and L2 coding region, the HPV early polyadenylation pAE must be subject to control [10,60,61]. As differentiation of the HPV infected cell proceeds and the cell is set up for HPV L1 and L2 capsid protein production and virus production, polyadenylation at the pAE must be reduced, whereas pAL must be active or activated (Fig. 2). As a consequence, HPV pAE must be regulatable, whereas pAL must be totally silent at the early stage and have high activity in terminally differentiated cells. Thus, HPVs should encode RNA control elements in addition to the core RNA elements present at regular polyadenylation signals, which include the upstream consensus AAUAAA element (polyadenylation signal) and a downstream GU- or U-rich element. The consensus AAUAAA element is recognized by cleavage and polyadenylation specificity factor (CPSF), which is required for both cleavage and poly(A) addition [124]. The GU- or U-rich element is recognized by cleavage stimulatory factor (CstF) are required for RNA cleavage and for stability of the CPSF-RNA complex [124].

In HPV16 and HPV31 a single consensus AAUAAA polyadenylation element (pAE) is present downstream of the E5 gene. In HPV31, pAE has been shown to be under control of several weak binding sites for CstF-64 (GU- or U-rich element) upstream and downstream of pAE [125,126]. The downstream CstF-64 binding sites are located in the L2 coding region [125,126]. In HPV16, sequences required for polyadenylation at pAE also extend into the L2 coding region and interact with CstF-64 [81,127]. A number of GGG-motifs downstream of HPV16 pAE interact with hnRNP H and enhance polyadenylation at pAE [81,127]. For both HPV16 and HPV31, cleavage of the pre-mRNA prior to polyadenylation appears to be heterogeneous [126,128]. HPV16 pAE is also dependent on U-rich sequences in the HPV early 3′-UTR (eUTR) that interact with multiple factors such as Fip-1, hnRNP C, hnRNP H, HuR and PTB [128,129] (Fig. 4). While Fip1 binds U-rich sequences upstream of polyadenylation signals to promote polyadenylation [130], transfection experiments showed that binding of hnRNP C and RALYL to U-rich sequences in the HPV16 early UTR inhibited pAE and favored utilization of the upstream splice donor SD3632 (Fig. 4) [131]. Thus, hnRNP C and RALYL interacted with U-rich RNA elements in the HPV16 early UTR, inhibited early polyadenylation and activated the upstream late SD3632 splice site to enhance late L1 mRNA splicing and polyadenylation at pAL [131] (Fig. 4). Overexpression of PTB had a similar effect on HPV16 late gene expression [132] (Fig. 4), supporting the idea that pAE and SD3632 are used in a mutually exclusive manner. Thus, activation of either of the two RNA processing sites, will reduce processing at the other.

In contrast to the single polyadenylation signal at [133] HPV16 pAE, there are multiple polyadenylation signals downstream of the HPV16 L1 stop codon [133]. The untranslated region of the late mRNAs also controls RNA processing by negatively affecting mRNAs in a cis-acting manner, at least in mitotic cells and cancer cells [60,134]. The inhibitory function of the late UTR is relatively well conserved among various HPVs (Fig. 4) [135]. Rather than influencing polyadenylation, the late UTR elements appear to affect late mRNA stability, transport or translation, either by interacting with cellular proteins such as CstF64, CUGBP and HuR or with U1snRNP that binds to 5′-splice site-like sequences located in the late UTR (Fig. 4) [136–140]. The late UTR in the cutaneous HPV1 is a classical AU-rich RNA instability element with high homology to AU-rich RNA instability elements in c-fos oncogene mRNAs and cytokine mRNAs [141,142]. In addition to reducing late mRNA stability, this element also inhibited mRNA translation [143]. It interacts specifically with hnRNP C, HuR and polyA-binding protein [141,143–146]. In summary, it appears that the late UTR of HPVs functions to guarantee silence of late L1 and L2 gene expression in undifferentiated cells, although they differ in mode of action.

5.5. Alternative splicing at HPV late splice sites dedicated to late L1 and L2 mRNA production

HPV late L1 and L2 gene expression is suppressed in HPV-induced cancer cells, presumably since cancer cells must evade immune surveillance to persist. One may therefore speculate that RNA-binding proteins that suppress HPV late gene expression operate in cancer cells and that inhibition of L1 and L2 gene expression is a prerequisite for persistence, progression and maintenance of cancer. HPV16 splice sites SD3632 and SA5639 are dedicated to the generation of HPV16 spliced late L1 mRNAs (Fig. 5). These splice sites are conditionally inactive during the early stage of the HPV16 life cycle but are activated upon cellular differentiation. Since HPV16 splice site SD3632 is located between the major HPV16 splice acceptor SA3358 and the early polyadenylation signal pAE (Fig. 5), SD3632 and pAE are used in a mutually exclusive manner. Therefore, it is of paramount importance that SD3632 is silent in mitotic cells to allow early mRNAs to be polyadenylation at pAE (Fig. 2). Splice donor sequences upstream of polyadenylation signals may have an inhibitory effect on mRNA levels [138,140], further supporting the idea that SD3632 must be suppressed at the early stage of the HPV life cycle. Furthermore, it is worth noting that once the HPV16 late splice sites SD3632 and SA5639 are activated, activation cannot be complete since the intronic sequence between SD3632 and SA5639 encodes L2. Infectious HPV capsids contain both L1 and L2 protein, and L1 and L2 mRNAs are produced in a mutually exclusive manner. It is imperative to obtain a balance between retention of the L2 encoding intron between SD3632 and SA5639 and splicing between these two splice sites to generate both L2 and L1 mRNAs. Thus, utilization of HPV16 splice sites SD3632 and SA5639 is strictly controlled [79,147].

hnRNPD, hnRNP DL, hnRNP AB and hnRNP A2/B1 were identified as suppressors of HPV16 late 5′-splice site SD3632 (Fig. 5). These proteins were shown to interact with cis-acting negative RNA elements located upstream of SD3632 [115,147]. As a matter of fact, sequences upstream of SD3632 were shown to be a hot spot for cellular RNA-binding proteins [60], including hnRNP A1 and hnRNP L, supporting the idea that the control of SD3632 is complex with multiple trans-acting factors involved (Fig. 5). In addition to these inhibitory RNA elements, U-rich sequences in the early UTR also control SD3632 (Fig. 5) [131]. The rational for this control-axis is that the early polyadenylation signal and SD3632 are used in a mutually exclusive manner. Thus, cellular RNA-binding protein hnRNP C1 and its cognate factor RALYL, can activate SD3632 and inhibit pAE in an early UTR-dependent manner [131]. In addition, the HPV16 late L1 mRNA-specific splice acceptor SA5639 is also suppressed by splicing silencers (Fig. 5) [79]. These splicing inhibitory RNA elements are located downstream of SA5639 [79]. The SA5639 is surrounded by hot spots for RNA-binding proteins including hnRNP A1 and hnRNP L (Fig. 5) [76]. hnRNP A1 interacted with wild type RNA sequences in the HPV16 L1 coding region that suppress SA5639, but not with mutant sequences in which splicing silencers had been inactivated, indicating that hnRNP A1 is specifically suppressing HPV16 late splice site SA5639 [79,80]. Interestingly, interactions of hnRNP L with the cis-acting, inhibitory RNA elements that suppress SD3632 and SA5639

are regulated by hnRNP L-phosphorylation by the Akt kinase [76]. These interactions correlated well with the activation of HPV16 late L1 mRNA splicing in response to Akt-kinase inhibition in cervical cancer cells [76]. In addition to the Akt-kinase-dependent effects on HPV16 alternative mRNA polyadenylation and splicing [76], activation of the DNA-damage response (DDR) pathway also had major effects on HPV16 mRNA polyadenylation and splicing [129,148]. The interactions of hnRNP C with the cis-elements at SD3632 and the interactions of Fip1 and CPSF30 with the cis-elements in the early UTR were influenced by DNA damage response [129,148]. The effect of the DDR on these factors strongly inhibited the HPV16 early polyadenylation signal and caused a read-through into the late HPV16 L1 and L2 coding regions followed by efficient polyadenylation at the HPV16 late polyadenylation signal pAL [129,148]. In combination with significant effects on HPV16 L1 mRNAs splicing, activation of the DDR was canalized into a strong activation of HPV16 late gene expression solely as a consequence of altered HPV16 mRNA processing mediated by various RNA-binding proteins [129,148]. In addition to the effects of processing of the late HPV16 mRNAs, a significant increase in splicing of HPV16 E2 and E4-mRNAs was observed [129,148], indicating that also these mRNAs were conditionally controlled by factors under the DDR pathway. These observations are of particular significance since HPVs are known to activate the DDR and utilize the DDR for HPV DNA replication [32,149]. Apparently, HPV can utilize the DDR pathway to alter processing of the HPV mRNAs with profound effects on HPV16 gene expression including a robust activation of HPV16 late gene expression. Taken together, multiple cellular RNA-binding proteins control HPV16 late gene expression. As a matter of fact, the majority of hnRNP- and SR-proteins had the ability to alter HPV16 late mRNA levels or splicing [150]. It will be of interest to investigate how the cellular RNA-binding proteins themselves are controlled by posttranslational modifications.

5.6. Stability and translation of HPV mRNAs and the role of RNA binding proteins

Since HPV early E6 and E7 mRNAs counteract cell differentiation, as crucial step in the activation of HPV late gene expression and virus production, an efficient shut down of E6 and E7 mRNAs is essential. Thus, it is reasonable to assume that early mRNAs and E6 and E7 proteins have short half-lives to quickly vanish once the early promoter is shut down, thereby paving the way for differentiation and virus production. As a matter of fact, blocking transcription in HPV16-driven tonsillar cancer cells with the DNA-alkylating cancer drug melphalan, resulted in reduced levels of HPV16 early mRNAs and the return of p53 protein within hours [99]. Cis-acting RNA elements that confer short half-life to the early HPV mRNAs remains to be determined, although the HPV16 early UTR is one candidate [151,152]. The HPV16 early UTR interacts with multiple RNA binding proteins (Fig. 4) [128,129,131], but the role of these proteins in HPV early mRNA stability has not been investigated. HPV16 late L1 and L2 coding region confer short half-life to these mRNAs [153], but the RNA binding proteins mediating this effect have not been identified. Finally, HPV late mRNAs are poorly translated, either as a result of intragenic RNA elements [154] or the presence of the late UTR [143]. HPV16 late L2 mRNAs interact specifically with hnRNP E/PCBP and hnRNP K that reduce translation efficiency of the L2 mRNAs [154] and the late UTR of HPV1 efficiently reduced translation of late HPV1 mRNAs [143], an effect that correlates with the binding to HuR and polyA-binding protein. Taken together, the results support the idea that RNA-binding proteins are involved in multiple steps in the HPV gene expression program.

6. Future perspective

There is a number of outstanding questions in the area of HPV mRNA processing and control of HPV gene expression that would be of interest to address in the future. For example, splicing is required to produce the E2 and E7 mRNAs but RNA elements required for production of these abundant spliced mRNAs have not yet been identified (Fig. 6i). More enigmatic is perhaps the generation of E1, E6 and L2 mRNAs, as they are only partially spliced and require an intron retention mechanism to generate the mature mRNAs with intact E1, E6 or L2 open reading frames. Identification of the cellular factors, e.g RNA-binding proteins, that are required for these splicing or intron retention processes would be of great interest to understand how these events are controlled (Fig. 6ii). The activity of the RNA-binding proteins themselves is likely subject to regulation by posttranslational modifications (Fig. 6iii). Also the HPV RNA elements may be modified, e.g. by m6A methylation, which could affect the ability of the RNA elements to interact with RNA-binding proteins. The location of m6A sites on HPV mRNAs would therefore by highly informative as would information on the role of various m6A-"writers", -"erasers" and -"readers" on the processing and control of HPV mRNA processing. As an extension, it would be equally interesting to determine signal transduction pathways that converge on the RNA-binding proteins that control HPV mRNA processing. In addition to these factors, the HPV chromatin may recruit various RNA-binding proteins to the HPV genome to facilitate a swift identification of regulatory RNA elements on the newly synthesized HPV mRNAs. Since the HPV life cycle is adapted to a differentiating environment, the effect of cell differentiation of the RNA elements and their cognate factors is of particular interest (Fig. 6iv and v). Of greater interest perhaps, is the activity of the same factors during carcinogenesis since HPV-induced carcinogenesis is restricted to cells that maintain an environment that allows for optimal splicing of the E6 and E7 mRNAs, as continuous production of both E6 and E7 proteins is required for maintenance of the carcinogenic state of the HPV-infected cell. Splicing that occurs within the E6 region to generate the E7 producing mRNAs, e.g. the so called E6*I/E7 mRNA with splicing between SD226 and SA4509 in HPV16, disserves special attention in that it is specific for the high-risk, carcinogenic HPV types. An understanding of the evolution of these RNA elements may enhance our understanding of the carcinogenic properties of HPV (Fig. 6vi). Similar RNA elements may be present on cellular mRNAs encoding proteins with a role in carcinogenesis, why furthering our knowledge of high-risk HPV RNA processing may shed a light on carcinogenesis in a wider perspective. In conjunction with cellular proteins controlling HPV RNA processing, it is known that HPV protein E2 can activate expression of cellular genes encoding splicing factors and control polyadenylation of the HPV early polyadenylation signal. It remains to be investigated if other HPV proteins alter expression of cellular RNA-binding proteins, alter their subcellular localization or affect HPV RNA processing (Fig. 6vii). Nevertheless, factors controlling HPV splice sites are potential targets for therapy to HPV infections or HPV-caused cancer. Since posttranslational modifications that control the activity of the RNA-binding proteins are carried out by enzymes, these enzymes constitute potential targets for therapy by various small molecule inhibitors. Such molecules could potentially perturb HPV mRNA processing and shift the balance in HPV gene expression to uncover the HPV infection for the immune surveillance, for example by inducting premature expression of the late genes encoding the highly immunogenic L1 and L2 proteins. Alternatively, small chemical molecules could perturb HPV mRNA processing to alter expression levels of HPV E6 and E7 oncoproteins, thereby potentially inducing apoptosis of HPV-infected cells and/or of HPV-driven cancer cells.

As a consequence of the recent remarkable developments of RNA research technology, more comprehensive and sensitive bioinformatical analysis tools are available. Advanced technologies including nanopore sequencing, single cell sequencing or CLIP-seq may pave the way for further research on RNA-binding proteins that control HPV mRNA alternative splicing and polyadenylation. These tools could be combined with experimental systems that support productive HPV replication such as 3D organotypic raft culture, methyl cellulose suspension culture systems or humanized mouse infection models. As a result, changes in HPV RNA processing or changes interactions of RNA-binding proteins

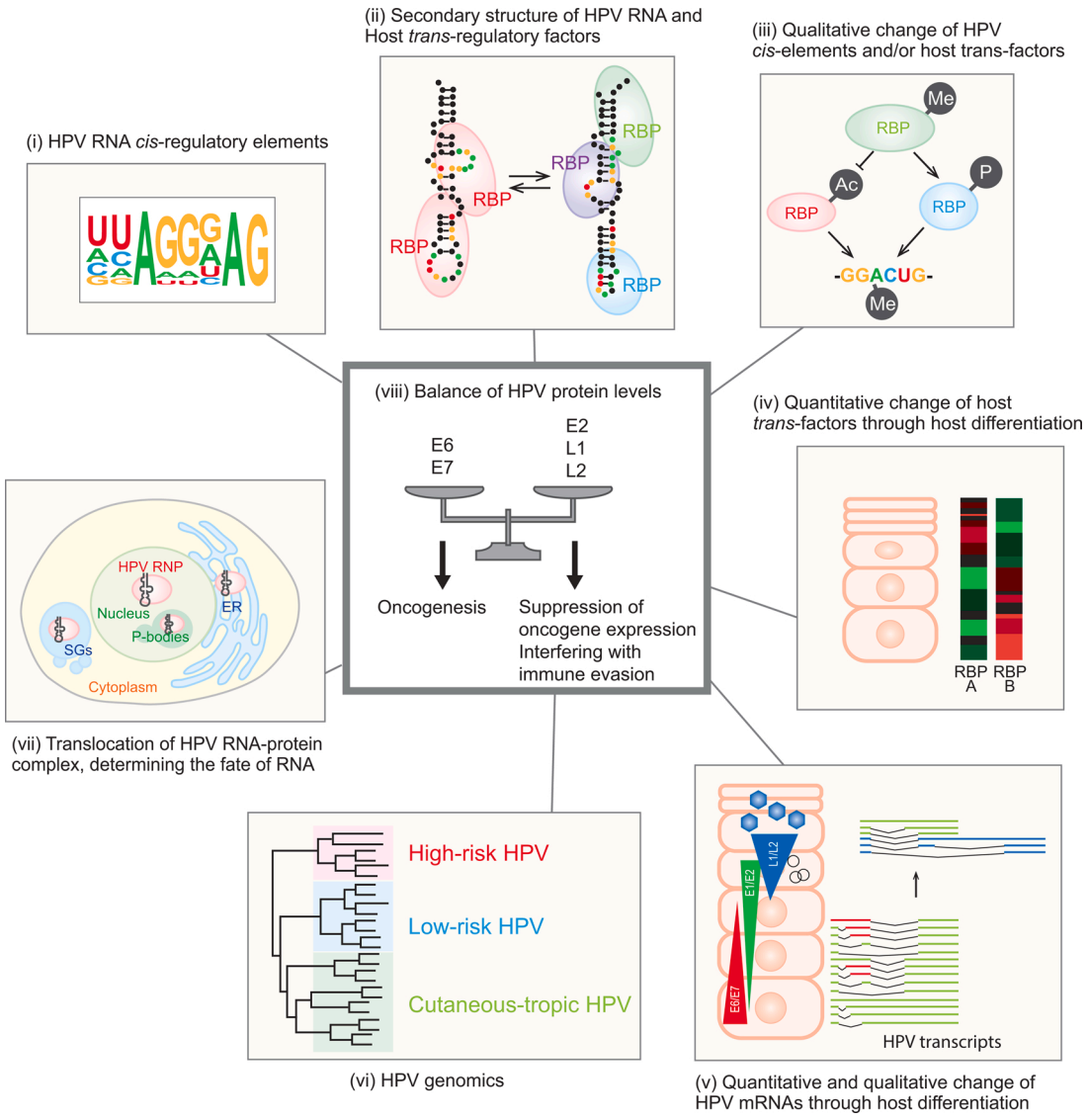


Fig. 6. Future perspective: areas of research on HPV mRNAs and their interactions with cellular RNA-binding proteins – significance for cancer biology. HPV-associated oncogenesis requires a balanced expression of the HPV proteins that is determined by a combination of cis-acting RNA elements on the HPV mRNAs and cellular trans acting factor, e.g. RNA-binding proteins that interact with HPV mRNAs. Analysis of primary **(i)** and secondary RNA structure **(ii)** of the cis-acting RNA elements and the significance for interactions with trans-actin RNA-binding proteins(RBPs) **(ii, iii)**. The interactions between mRNAs and RBPs may be influenced both by qualitative changes e.g. posttranslational modifications **(iii)**, or by quantitative changes of RNA-binding proteins **(iv)**, in response to cell differentiation or cell transformation. Altered interactions between HPV RNAs and cellular RNA-binding proteins may affect the alternative splicing regulation on the RNA, resulting in the production of HPV mRNA variants **(v)**. Evolution of the cis-acting HPV RNA elements in high-risk and low-risk HPVs and in cutaneous and mucosal HPV types may contribute to the different profile of HPV transcriptome **(vi)**. The interaction between HPV mRNAs and cellular RNA-binding proteins may further affect the fate of various HPV mRNAs, including effects on RNA localization, stability and translation efficiency **(vii)**. Perturbed interactions between HPV RNAs and cellular RBPs may affect the entire HPV gene expression program by altering the balance between oncogenic HPV proteins (E6, E7) levels and pro-apoptotic E2 proteins, thereby affecting the pathogenic properties of HPV and the outcome of the HPV infection **(vii)**. Expression of the HPV late genes encoding the highly immunogenic L1 and L2 capsid proteins is also controlled by interactions between HPV mRNAs and RNA-binding proteins. Alterations of these interactions may cause premature late gene expression that uncovers the HPV-infected cells for the immune surveillance or may silence L1 and L2 expression thereby contributing to immune evasion and HPV persistence in the presence of a functional immune system. Me: methylation, P: phosphorylation, Ac: acetylation, HPV RNP: HPV Ribonucleoprotein complex (HPV RNA-protein complex), ER: endoplasmic reticulum, SG: stress granule.

with HPV mRNAs could be investigated in a temporospatial manner in response to cell differentiation or carcinogenesis (Fig. 6iv and vi).

Finally, these research activities could have a great impact on public health by the development of specific antiviral strategies to treat pre-malignant HPV-positive cervical lesions, or oropharyngeal HPV infections and HPV-driven cancers and/or other HPV-caused proliferative diseases. Research on RNA-binding proteins that control HPV RNA processing will also enhance our understanding of cancer-associated alterations in the cellular RNA processing machinery, potentially identifying novel targets for therapy to cancer.

Declaration of Competing Interest

The authors declare that there are no conflicts of interest.

Acknowledgements

This work was supported by the Swedish Research Council-Medicine [VR2019-01210 to S.S.] and the Swedish Cancer Society [CAN2018/702 to S.S.].

References

[1] M.E. McLaughlin-Drubin, K. Munger, Viruses associated with human cancer, Biochim. Biophys. Acta 1782 (2008) 127–150.
[2] E.R. Leight, B. Sugden, EBNA-1: a protein pivotal to latent infection by Epstein-Barr virus, Rev. Med. Virol. 10 (2000) 83–100.
[3] M.E. Ballestas, K.M. Kaye, The latency-associated nuclear antigen, a multifunctional protein central to Kaposi's sarcoma-associated herpesvirus latency, Future Microbiol. 6 (2011) 1399–1413.
[4] C.Z. Giam, K.T. Jeang, HTLV-1 tax and adult T-cell leukemia, Front. Biosci. 12 (2007) 1496–1507.
[5] A.A. McBride, The papillomavirus E2 proteins, Virology 445 (2013) 57–79.
[6] F. Thierry, Transcriptional regulation of the papillomavirus oncogenes by cellular and viral transcription factors in cervical carcinoma, Virology 384 (2009) 375–379.
[7] R.E. Lloyd, Nuclear proteins hijacked by mammalian cytoplasmic plus strand RNA viruses, Virology 479-480 (2015) 457–474.
[8] D.M. D'Agostino, I. Cavallari, M.G. Romanelli, V. Ciminale, Post-transcriptional regulation of HTLV gene expression: rex to the rescue, Front. Microbiol. 10 (2019) 1958.
[9] V. Majerciak, Z.M. Zheng, KSHV ORF57, a protein of many faces, Viruses 7 (2015) 604–633.
[10] C. Johansson, M. Somberg, X. Li, E. Backström Winquist, J. Fay, F. Ryan, D. Pim, L. Banks, S. Schwartz, HPV-16 E2 contributes to induction of HPV-16 late gene expression by inhibiting early polyadenylation, EMBO J. 13 (2012) 3212–3227.
[11] K. Van Doorslaer, Q. Tan, S. Xirasagar, S. Bandaru, V. Gopalan, Y. Mohamoud, Y. Huyen, A.A. McBride, The Papillomavirus Episteme: a central resource for papillomavirus sequence data and analysis, Nucleic Acids Res. 41 (2013) D571–578.
[12] E.M. de Villiers, Cross-roads in the classification of papillomaviruses, Virology 445 (2013) 2–10.
[13] J.M. Walboomers, M.V. Jacobs, M.M. Manos, F.X. Bosch, J.A. Kummer, K.V. Shah, P.J. Snijders, J. Peto, C.J. Meijer, N. Munoz, Human papillomavirus is a necessary cause of invasive cervical cancer worldwide, J. Pathol. 189 (1999) 12–19.
[14] F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre, A. Jemal, Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA Cancer J. Clin. 68 (2018) 394–424.
[15] J. Ferlay, M. Colombet, I. Soerjomataram, C. Mathers, D.M. Parkin, M. Pineros, A. Znaor, F. Bray, Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods, Int. J. Cancer 144 (2019) 1941–1953.
[16] J. Lei, A. Ploner, K.M. Elfstrom, J. Wang, A. Roth, F. Fang, K. Sundstrom, J. Dillner, P. Sparen, HPV vaccination and the risk of invasive cervical cancer, N. Engl. J. Med. 383 (2020) 1340–1348.
[17] G.A. Rivera, F. Morell, StatPearls, Treasure Island (FL), 2021.
[18] F. Faraji, E.M. Rettig, H.L. Tsai, M. El Asmar, N. Fung, D.W. Eisele, C. Fakhry, The prevalence of human papillomavirus in oropharyngeal cancer is increasing regardless of sex or race, and the influence of sex and race on survival is modified by human papillomavirus tumor status, Cancer 125 (2019) 761–769.
[19] J.E. Tota, A.F. Best, Z.S. Zumsteg, M.L. Gillison, P.S. Rosenberg, A.K. Chaturvedi, Evolution of the oropharynx cancer epidemic in the United States: moderation of increasing incidence in younger individuals and shift in the burden to older individuals, J. Clin. Oncol. 37 (2019) 1538–1546.
[20] M. Lechner, J. Liu, L. Masterson, T.R. Fenton, HPV-associated oropharyngeal cancer: epidemiology, molecular biology and clinical management, Nat. Rev. Clin. Oncol. (2022).
[21] A.R. Kreimer, G.M. Clifford, P. Boyle, S. Franceschi, Human papillomavirus types in head and neck squamous cell carcinomas worldwide: a systematic review, Cancer Epidemiol. Biomarkers Prev. 14 (2005) 467–475.
[22] S. DiGiuseppe, M. Bienkowska-Haba, L.G. Guion, M. Sapp, Cruising the cellular highways: how human papillomavirus travels from the surface to the nucleus, Virus Res. 231 (2017) 1–9.
[23] M. Bergvall, T. Melendy, J. Archambault, The E1 proteins, Virology 445 (2013) 35–56.
[24] M. Chojnacki, T. Melendy, The human papillomavirus DNA helicase E1 binds, stimulates, and confers processivity to cellular DNA polymerase epsilon, Nucleic Acids Res. 46 (2018) 229–241.
[25] T. Melendy, J. Sedman, A. Stenlund, Cellular factors required for papillomavirus DNA replication, J. Virol. 69 (1995) 7857–7867.
[26] A. Roman, K. Munger, The papillomavirus E7 proteins, Virology 445 (2013) 138–168.
[27] S.B. Vande Pol, A.J. Klingelhutz, Papillomavirus E6 oncoproteins, Virology 445 (2013) 115–137.
[28] D. DiMaio, L.M. Petti, The E5 proteins, Virology 445 (2013) 99–114.
[29] A. Warburton, A.N. Della Fera, A.A. McBride, Dangerous liaisons: long-term replication with an extrachromosomal HPV genome, Viruses 13 (2021).
[30] M. Kadaja, T. Silla, E. Ustav, M. Ustav, Papillomavirus DNA replication - from initiation to genomic instability, Virology 384 (2009) 360–368.
[31] J. Doorbar, The E4 protein; structure, function and patterns of expression, Virology 445 (2013) 80–98.
[32] S. Hong, L.A. Laimins, Regulation of the life cycle of HPVs by differentiation and the DNA damage response, Future Microbiol. 8 (2013) 1547–1557.
[33] L.T. Chow, T.R. Broker, B.M. Steinberg, The natural history of human papillomavirus infections of the mucosal epithelia, APMIS 118 (2010) 422–449.
[34] N. Kajitani, A. Satsuka, A. Kawate, H. Sakai, Productive lifecycle of human papillomaviruses that depends upon squamous epithelial differentiation, Front. Microbiol. 3 (2012) 152.
[35] K. Middleton, W. Peh, S. Southern, H. Griffin, K. Sotlar, T. Nakahara, A. El-Sherif, L. Morris, R. Seth, M. Hibma, et al., Organization of human papillomavirus productive cycle during neoplastic progression provides a basis for selection of diagnostic markers, J. Virol. 77 (2003) 10186–10201.
[36] M. Schiffman, J. Doorbar, N. Wentzensen, S. de Sanjose, C. Fakhry, B.J. Monk, M. A. Stanley, S. Franceschi, Carcinogenic human papillomavirus infection, Nat. Rev. Dis. Primers 2 (2016) 16086.
[37] J.A. Scarth, M.R. Patterson, E.L. Morgan, A. Macdonald, The human papillomavirus oncoproteins: a review of the host pathways targeted on the road to transformation, J. Gen. Virol. 102 (2021).
[38] A. Nishimura, T. Ono, A. Ishimoto, J.J. Dowhanick, M.A. Frizzell, P.M. Howley, H. Sakai, Mechanisms of human papillomavirus E2-mediated repression of viral oncogene expression and cervical cancer cell growth inhibition, J. Virol. 74 (2000) 3752–3760.
[39] J.L. Parish, A. Kowalczyk, H.T. Chen, G.E. Roeder, R. Sessions, M. Buckle, K. Gaston, E2 proteins from high- and low-risk human papillomavirus types differ in their ability to bind p53 and induce apoptotic cell death, J. Virol. 80 (2006) 4580–4590.
[40] S. Blachon, S. Bellanger, C. Demeret, F. Thierry, Nucleo-cytoplasmic shuttling of high risk human Papillomavirus E2 proteins induces apoptosis, J. Biol. Chem. 280 (2005) 36088–36098.
[41] F. Thierry, C. Demeret, Direct activation of caspase 8 by the proapoptotic E2 protein of HPV18 independent of adaptor proteins, Cell Death Differ. 15 (2008) 1356–1363.
[42] A.M. Sanchez-Perez, S. Soriano, A.R. Clarke, K. Gaston, Disruption of the human papillomavirus type 16 E2 gene protects cervical carcinoma cells from E2F-induced apoptosis, J. Gen. Virol. 78 (Pt 11) (1997) 3009–3018.
[43] S.I. Collins, C. Constandinou-Williams, K. Wen, L.S. Young, S. Roberts, P. G. Murray, C.B. Woodman, Disruption of the E2 gene is a common and early event in the natural history of cervical human papillomavirus infection: a longitudinal cohort study, Cancer Res. 69 (2009) 3828–3832.
[44] M. Dreer, S. van de Poel, F. Stubenrauch, Control of viral replication and transcription by the papillomavirus E8^E2 protein, Virus Res. 231 (2017) 96–102.
[45] O. Basukala, L. Banks, The not-so-good, the bad and the ugly: HPV E5, E6 and E7 oncoproteins in the orchestration of carcinogenesis, Viruses 13 (2021).
[46] C. Royer, X. Lu, Epithelial cell polarity: a major gatekeeper against cancer? Cell Death Differ. 18 (2011) 1470–1477.
[47] K. Ganti, J. Broniarczyk, W. Manoubi, P. Massimi, S. Mittal, D. Pim, A. Szalmas, J. Thatte, M. Thomas, V. Tomaic, et al., The human papillomavirus E6 PDZ binding motif: from life cycle to malignancy, Viruses 7 (2015) 3530–3551.
[48] C.P. Delury, E.K. Marsh, C.D. James, S.S. Boon, L. Banks, G.L. Knight, S. Roberts, The role of protein kinase A regulation of the E6 PDZ-binding domain during the differentiation-dependent life cycle of human papillomavirus type 18, J. Virol. 87 (2013) 9463–9472.
[49] L. Nicolaides, C. Davy, K. Raj, C. Kranjec, L. Banks, J. Doorbar, Stabilization of HPV16 E6 protein by PDZ proteins, and potential implications for genome maintenance, Virology 414 (2011) 137–145.
[50] L. Gewin, H. Myers, T. Kiyono, D.A. Galloway, Identification of a novel telomerase repressor that interacts with the human papillomavirus type-16 E6/E6-AP complex, Genes Dev. 18 (2004) 2269–2282.
[51] T. Veldman, X. Liu, H. Yuan, R. Schlegel, Human papillomavirus E6 and Myc proteins associate in vivo and bind to and cooperatively activate the telomerase reverse transcriptase promoter, Proc. Natl. Acad. Sci. U. S. A. 100 (2003) 8211–8216.
[52] J.A. Westrich, C.J. Warren, D. Pyeon, Evasion of host immune defenses by human papillomavirus, Virus Res. 231 (2017) 21–33.
[53] I. Lo Cigno, F. Calati, S. Albertini, M. Gariglio, Subversion of host innate immunity by human papillomavirus oncoproteins, Pathogens 9 (2020).
[54] T.L. Coursey, A.A. McBride, Hitchhiking of viral genomes on cellular chromosomes, Annu. Rev. Virol. 6 (2019) 275–296.
[55] I.J. Groves, N. Coleman, Human papillomavirus genome integration in squamous carcinogenesis: what have next-generation sequencing studies taught us? J. Pathol. 245 (2018) 9–18.
[56] Y. Xue, D. Lim, L. Zhi, P. He, J.P. Abastado, F. Thierry, Loss of HPV16 E2 protein expression without disruption of the E2 ORF correlates with carcinogenic progression, Open Virol. J. 6 (2012) 163–172.
[57] M. von Knebel Doeberitz, E.S. Prigge, Role of DNA methylation in HPV associated lesions, Papillomavirus Res. 7 (2019) 180–183.
[58] J. Doorbar, W. Quint, L. Banks, I.G. Bravo, M. Stoler, T.R. Broker, M.A. Stanley, The biology and life-cycle of human papillomaviruses, Vaccine 30 (Suppl. 5) (2012) F55–70.
[59] C.M. de Oliveira, J. Fregnani, L.L. Villa, HPV vaccine: updates and highlights, Acta Cytol. 63 (2019) 159–168.
[60] N. Kajitani, S. Schwartz, Role of viral ribonucleoproteins in human papillomavirus type 16 gene expression, Viruses 12 (2020).
[61] C. Johansson, S. Schwartz, Regulation of human papillomavirus gene expression by splicing and polyadenylation, Nature Rev. Microbiol. 11 (2013) 239–251.
[62] S.V. Graham, A.A. Faizo, Control of human papillomavirus gene expression by alternative splicing, Virus Res. 231 (2017) 83–95.
[63] R. Jia, Z.M. Zheng, Regulation of bovine papillomavirus type 1 gene expression by RNA processing, Front. Biosci. 14 (2009) 1270–1282.
[64] Y. Zheng, J. Jonsson, C. Hao, S. Shoja Chaghervand, X. Cui, N. Kajitani, L. Gong, C. Wu, S. Schwartz, Heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1) and hnRNP A2 inhibit splicing to human papillomavirus 16 splice site SA409 through a UAG-containing sequence in the E7 coding region, J. Virol. 94 (2020).

[65] S. Tang, M. Tao, J.P. McCoy Jr., Z.M. Zheng, The E7 oncoprotein is translated from spliced E6*I transcripts in high-risk human papillomavirus type 16- or type 18-positive cervical cancer cell lines via translation reinitiation, J. Virol. 80 (2006) 4249–4263.
[66] S.N. Stacey, D. Jordan, A.J. Williamson, M. Brown, J.H. Coote, J.R. Arrand, Leaky scanning is the predominant mechanism for translation of human papillomavirus type 16 E7 oncoprotein from E6/E7 bicistronic mRNA, J. Virol. 74 (2000) 7284–7297.
[67] M. Ajiro, S. Tang, J. Doorbar, Z.M. Zheng, Serine/arginine-rich splicing factor 3 and heterogeneous nuclear ribonucleoprotein A1 regulate alternative RNA splicing and gene expression of human papillomavirus 18 through two functionally distinguishable cis elements, J. Virol. 90 (2016) 9138–9152.
[68] K. Grassmann, B. Rapp, H. Maschek, K.U. Petry, T. Iftner, Identification of a differentiation-inducible promoter in the E7 open reading frame of human papillomavirus type 16 (HPV-16) in raft cultures of a new cell line containing high copy numbers of episomal HPV-16 DNA, J. Virol. 70 (1996) 2339–2349.
[69] D. Smotkin, F.O. Wettstein, Transcription of human papillomavirus type 16 early genes in a cervical cancer and a cancer-derived cell line and identification of the E7 protein, Proc. Natl. Acad. Sci. U. S. A. 83 (1986) 4680–4684.
[70] M.A. Ozbun, C. Meyers, Temporal usage of multiple promoters during the life cycle of human papillomavirus type 31b, J. Virol. 72 (1998) 2715–2722.
[71] M. Hummel, J.B. Hudson, L.A. Laimins, Differentiation-induced and constitutive transcription of human papillomavirus type 31b in cell lines containing viral episomes, J. Virol 66 (1992) 6070–6080.
[72] Z.M. Zheng, M. Tao, K. Yamanegi, S. Bodaghi, W. Xiao, Splicing of a cap-proximal human Papillomavirus 16 E6E7 intron promotes E7 expression, but can be restrained by distance of the intron from its RNA 5' cap, J. Mol. Biol. 337 (2004) 1091–1108.
[73] A. Garcia, G. Maldonado, J.L. Gonzalez, Y. Svitkin, D. Cantu, A. Garcia-Carranca, N. Sonenberg, G. Hernandez, High-risk human papillomavirus-18 uses an mRNA sequence to synthesize oncoprotein E6 in tumors, Proc. Natl. Acad. Sci. U. S. A. 118 (2021).
[74] M. Kozak, Regulation of translation inititaion in eucaryotic systems, Annu. Rev. Cell Biol. 8 (1992) 197–225.
[75] A. Cerasuolo, L. Buonaguro, F.M. Buonaguro, M.L. Tornesello, The role of RNA splicing factors in cancer: regulation of viral and human gene expression in human papillomavirus-related cervical Cancer, Front. Cell Dev. Biol. 8 (2020) 474.
[76] N. Kajitani, J. Glahder, C. Wu, H. Yu, K. Nilsson, S. Schwartz, hnRNP L controls HPV16 RNA polyadenylation and splicing in an Akt kinase-dependent manner, Nucleic Acids Res. 45 (2017) 9654–9678.
[77] B.J. Blencowe, Exonic splicing enhancers: mechanism of action, diversity and role in human genetic diseases, Trends Biochem. Sci. 25 (2000) 106–110.
[78] L. Cartegni, S.L. Chew, A.R. Krainer, Listening to silence and understanding nonsense: exonic mutations that affect splicing, Nat. Rev. Genet. 3 (2002) 285–298.
[79] X. Zhao, M. Rush, S. Schwartz, Identification of an hnRNP A1 dependent splicing silencer in the HPV-16 L1 coding region that prevents premature expression of the late L1 gene, J. Virol. 78 (2004) 10888–10905.
[80] B. Collier, D. Oberg, X. Zhao, S. Schwartz, Specific inactivation of inhibitory sequences in the 5' end of the human papillomavirus type 16 L1 open reading frame results in production of high levels of L1 protein in human epithelial cells, J. Virol. 76 (2002) 2739–2752.
[81] D. Oberg, J. Fay, H. Lambkin, S. Schwartz, A downstream polyadenylation element in human papillomavirus type 16 L2 encodes multiple GGG motifs and interacts with hnRNP H, J. Virol. 79 (2005) 9254–9269.
[82] S. Naftelberg, I.E. Schor, G. Ast, A.R. Kornblihtt, Regulation of alternative splicing through coupling with transcription and chromatin structure, Annu. Rev. Biochem. 84 (2015) 165–198.
[83] E. Albert, L. Laimins, Regulation of the human papillomavirus life cycle by DNA damage repair pathways and epigenetic factors, Viruses (2020) 12.
[84] X. Wang, H. Liu, H. Ge, M. Ajiro, N.R. Sharma, C. Meyers, P. Morozov, T. Tuschl, A. Klar, D. Court, et al., Viral DNA replication orientation and hnRNPs regulate transcription of the human papillomavirus 18 late promoter, mBio 8 (2017).
[85] M. Burley, S. Roberts, J.L. Parish, Epigenetic regulation of human papillomavirus transcription in the productive virus life cycle, Semin. Immunopathol. 42 (2020) 159–171.
[86] P. Cramer, A. Srebrow, S. Kadener, S. Werbajh, M. de la Mata, G. Melen, G. Nogues, A.R. Kornblihtt, Coordination between transcription and pre-mRNA processing, FEBS Lett. 498 (2001) 179–182.
[87] C. Johansson, T. Jamal Fattah, H. Yu, J. Nygren, A.K. Mossberg, S. Schwartz, Acetylation of intragenic histones on HPV16 correlates with enhanced HPV16 gene expression, Virology 482 (2015) 244–259.
[88] E. Johannsen, P.F. Lambert, Epigenetics of human papillomaviruses, Virology 445 (2013) 205–212.
[89] J. Durzynska, K. Lesniewicz, E. Poreba, Human papillomaviruses in epigenetic regulations, Mutat. Res. Rev. Mutat. Res. 772 (2017) 36–50.
[90] J. Ferguson, K. Campos-Leon, I. Pentland, J.D. Stockton, T. Gunther, A.D. Beggs, A. Grundhoff, S. Roberts, B. Noyvert, J.L. Parish, The chromatin insulator CTCF regulates HPV18 transcript splicing and differentiation-dependent late gene expression, PLoS Pathog. 17 (2021), e1010032.
[91] S. Adhikari, W. Xiao, Y.L. Zhao, Y.G. Yang, M(6)A: signaling for mRNA splicing, RNA Biol. 13 (2016) 756–759.
[92] C.R. Alarcon, H. Goodarzi, H. Lee, X. Liu, S. Tavazoie, S.F. Tavazoie, HNRNPA2B1 is a mediator of m(6)A-dependent nuclear RNA processing events, Cell 162 (2015) 1299–1308.
[93] J. Zhao, E.E. Lee, J. Kim, R. Yang, B. Chamseddin, C. Ni, E. Gusho, Y. Xie, C. M. Chiang, M. Buszczak, et al., Transforming activity of an oncoprotein-encoding circular RNA from human papillomavirus, Nat. Commun. 10 (2019) 2300.
[94] M. Ajiro, R. Jia, L. Zhang, X. Liu, Z.M. Zheng, Intron definition and a branch site adenosine at nt 385 control RNA splicing of HPV16 E6*I and E7 expression, PLoS One 7 (2012), e46412.
[95] S. Islam, H. Dasgupta, A. Roychowdhury, R. Bhattacharya, N. Mukherjee, A. Roy, G.K. Mandal, N. Alam, J. Biswas, S. Mandal, et al., Study of association and molecular analysis of human papillomavirus in breast cancer of Indian patients: clinical and prognostic implication, PLoS One 12 (2017), e0172760.
[96] D.M. Schutze, P.J. Snijders, L. Bosch, D. Kramer, C.J. Meijer, R.D. Steenbergen, Differential in vitro immortalization capacity of eleven (probable) [corrected] high-risk human papillomavirus types, J. Virol. 88 (2014) 1714–1724.
[97] D. Smotkin, H. Prokoph, F.O. Wettstein, Oncogenic and nononcogenic human genital papillomaviruses generate the E7 mRNA by different mechanisms, J. Virol. 63 (1989) 1441–1447.
[98] M. Schmitt, V. Dalstein, T. Waterboer, C. Clavel, L. Gissmann, M. Pawlita, Diagnosing cervical cancer and high-grade precursors by HPV16 transcription patterns, Cancer Res. 70 (2010) 249–256.
[99] C. Wu, K. Nilsson, Y. Zheng, C. Ekenstierna, N. Sugiyama, O. Forslund, N. Kajitani, H. Yu, J. Wennerberg, L. Ekblad, et al., Short half-life of HPV16 E6 and E7 mRNAs sensitizes HPV16-positive tonsillar cancer cell line HN26 to DNA-damaging drugs, Int. J. Cancer 144 (2019) 297–310.
[100] O. del Moral-Hernandez, E. Lopez-Urrutia, R. Bonilla-Moreno, M. Martinez-Salazar, E. Arechaga-Ocampo, J. Berumen, N. Villegas-Sepulveda, The HPV-16 E7 oncoprotein is expressed mainly from the unspliced E6/E7 transcript in cervical carcinoma C33-A cells, Arch. Virol. 155 (2010) 1959–1970.
[101] M. Ajiro, Z.M. Zheng, E6^E7, a novel splice isoform protein of human papillomavirus 16, stabilizes viral E6 and E7 oncoproteins via HSP90 and GRP78, mBio 6 (2015) e02068–02014.
[102] D. Pim, V. Tomaic, L. Banks, The human papillomavirus (HPV) E6* proteins from high-risk, mucosal HPVs can direct degradation of cellular proteins in the absence of full-length E6 protein, J. Virol. 83 (2009) 9863–9874.
[103] E. Guccione, D. Pim, L. Banks, HPV-18 E6*I modulates HPV-18 full-length E6 functions in a cell cycle dependent manner, Int. J. Cancer 110 (2004) 928–933.
[104] M. Filippova, M.M. Johnson, M. Bautista, V. Filippov, N. Fodor, S. S. Tungteakkhun, K. Williams, P.J. Duerksen-Hughes, The large and small isoforms of human papillomavirus type 16 E6 bind to and differentially affect procaspase 8 stability and activity, J. Virol. 81 (2007) 4116–4129.
[105] Y. Zheng, X. Cui, K. Nilsson, H. Yu, L. Gong, C. Wu, S. Schwartz, Efficient production of HPV16 E2 protein from HPV16 late mRNAs spliced from SD880 to SA2709, Virus Res. 285 (2020), 198004.
[106] S. Rosenberger, J. De-Castro Arce, L. Langbein, R.D. Steenbergen, F. Rosl, Alternative splicing of human papillomavirus type-16 E6/E6* early mRNA is coupled to EGF signaling via Erk1/2 activation, Proc. Natl. Acad. Sci. U. S. A. 107 (2010) 7006–7011.
[107] J. Fay, P. Kelehan, H. Lambkin, S. Schwartz, Increased expression of cellular RNA-binding proteins in HPV-induced neoplasia and cervical cancer, J. Med. Virol. 81 (2009) 897–907.
[108] M. McFarlane, A.I. MacDonald, A. Stevenson, S.V. Graham, Human papillomavirus 16 oncoprotein expression is controlled by the cellular splicing factor SRSF2 (SC35), J. Virol. 89 (2015) 5276–5287.
[109] W.G. Hubert, L.A. Laimins, Human papillomavirus type 31 replication modes during the early phases of the viral life cycle depend on transcriptional and posttranscriptional regulation of E1 and E2 expression, J. Virol. 76 (2002) 2263–2273.
[110] N.S. Belaguli, M.M. Pater, A. Pater, Nucleotide 880 splice donor site required for efficient transformation and RNA accumulation by human papillomavirus type 16 E7 gene, J. Virol. 66 (1992) 2724–2730.
[111] L. Sherman, N. Alloul, Human papillomavirus type 16 expresses a variety of alternatively spliced mRNAs putatively encoding the E2 protein, Virology 191 (1992) 953–959.
[112] M.J. Lace, J.R. Anson, G.S. Thomas, L.P. Turek, T.H. Haugen, The E8–E2 gene product of human papillomavirus type 16 represses early transcription and replication but is dispensable for viral plasmid persistence in keratinocytes, J. Virol. 82 (2008) 10841–10853.
[113] F. Stubenrauch, M. Hummel, T. Iftner, L.A. Laimins, The E8E2C protein, a negative regulator of viral transcription and replication, is required for extrachromosomal maintenance of human papillomavirus type 31 in keratinocytes, J. Virol. 74 (2000) 1178–1186.
[114] M. Somberg, S. Schwartz, Multiple ASF/SF2 sites in the human papillomavirus type 16 (HPV-16) E4-coding region promote splicing to the most commonly used 3'-splice site on the HPV-16 genome, J. Virol. 84 (2010) 8219–8230.
[115] M. Rush, X. Zhao, S. Schwartz, A splicing enhancer in the E4 coding region of human papillomavirus type 16 is required for early mRNA splicing and polyadenylation as well as inhibition of premature late gene expression, J. Virol. 79 (2005) 12002–12015.
[116] X. Li, C. Johansson, C. Cardoso-Palacios, A. Mossberg, S. Dhanjal, M. Bergvall, S. Schwartz, Eight nucleotide substitutions inhibit splicing to HPV-16 3'-splice site SA3358 and reduce the efficiency by which HPV-16 increases the life span of primary human keratinocytes, PLoS One 8 (2013), e72776.
[117] M. Somberg, X. Li, C. Johansson, B. Orru, R. Chang, M. Rush, J. Fay, F. Ryan, S. Schwartz, Serine/arginine-rich protein 30c activates human papillomavirus type 16 L1 mRNA expression via a bimodal mechanism, J. Gen. Virol. 92 (2011) 2411–2421.

[118] R. Jia, X. Liu, M. Tao, M. Kruhlak, M. Guo, C. Meyers, C.C. Baker, Z.M. Zheng, Control of the papillomavirus early-to-late switch by differentially expressed SRp20, J. Virol. 83 (2009) 167–180.
[119] P.J. Shepard, K.J. Hertel, The SR protein family, Genome Biol. 10 (2009) 242.
[120] Z. Zhou, X.D. Fu, Regulation of splicing by SR proteins and SR protein-specific kinases, Chromosoma 122 (2013) 191–207.
[121] S. Mole, A.A.A. Faizo, H. Hernandez-Lopez, M. Griffiths, A. Stevenson, S. Roberts, S.V. Graham, Human papillomavirus type 16 infection activates the host serine arginine protein kinase 1 (SRPK1) - splicing factor axis, J. Gen. Virol. 101 (2020) 523–532.
[122] S. Mole, S.G. Milligan, S.V. Graham, Human papillomavirus type 16 E2 protein transcriptionally activates the promoter of a key cellular splicing factor, SF2/ASF, J. Virol. 83 (2009) 357–367.
[123] E.L. Prescott, C.L. Brimacombe, M. Hartley, I. Bell, S. Graham, S. Roberts, Human papillomavirus type 1 E1^E4 protein is a potent inhibitor of the serine-arginine (SR) protein kinase SRPK1 and inhibits phosphorylation of host SR proteins and of the viral transcription and replication regulator E2, J. Virol. 88 (2014) 12599–12611.
[124] B. Tian, J.L. Manley, Alternative polyadenylation of mRNA precursors, Nat. Rev. Mol. Cell Biol. 18 (2017) 18–30.
[125] S.S. Terhune, W.G. Hubert, J.T. Thomas, L.A. Laimins, Early polyadenylation signals of human papillomavirus type 31 negatively regulate capsid gene expression, J. Virol. 75 (2001) 8147–8157.
[126] S.S. Terhune, C. Milcarek, L.A. Laimins, Regulation of human papillomavirus 31 polyadenylation during the differentiation-dependent life cycle, J. Virol. 73 (1999) 7185–7192.
[127] D. Oberg, B. Collier, X. Zhao, S. Schwartz, Mutational inactivation of two distinct negative RNA elements in the human papillomavirus type 16 L2 coding region induces production of high levels of L2 in human cells, J. Virol. 77 (2003) 11674–11684.
[128] X. Zhao, D. Öberg, M. Rush, J. Fay, H. Lambkin, S. Schwartz, A 57 nucleotide upstream early polyadenylation element in human papillomavirus type 16 interacts with hFip1, CstF-64, hnRNP C1/C2 and PTB, J. Virol. 79 (2005) 4270–4288.
[129] K. Nilsson, C. Wu, N. Kajitani, H. Yu, E. Tsimtsirakis, L. Gong, E.B. Winquist, J. Glahder, L. Ekblad, J. Wennerberg, et al., The DNA damage response activates HPV16 late gene expression at the level of RNA processing, Nucleic Acids Res. 46 (2018) 5029–5049.
[130] I. Kaufmann, G. Martin, A. Friedlein, H. Langen, W. Keller, Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase, EMBO J. 23 (2004) 616–626.
[131] S. Dhanjal, N. Kajitani, J. Glahder, A.K. Mossberg, C. Johansson, S. Schwartz, Heterogeneous nuclear ribonucleoprotein C proteins interact with the human papillomavirus type 16 (HPV16) early 3'-untranslated region and alleviate suppression of HPV16 late L1 mRNA splicing, J. Biol. Chem. 290 (2015) 13354–13371.
[132] M. Somberg, X. Zhao, M. Fröhlich, M. Evander, S. Schwartz, PTB induces HPV-16 late gene expression by interfering with splicing inhibitory elements at the major late 5'-splice site SD3632, J. Virol. 82 (2008) 3665–3678.
[133] I.M. Kennedy, J.K. Haddow, J.B. Clements, Analysis of human papillomavirus type 16 late mRNA 3' processing signals in vitro and in vivo, J. Virol. 64 (1990) 1825–1829.
[134] S.V. Graham, Papillomavirus 3' UTR regulatory elements, Front. Biosci. 13 (2008) 5646–5663.
[135] X. Zhao, M. Rush, A. Carlsson, S. Schwartz, The presence of inhibitory RNA elements in the late 3'-untranslated region is a conserved property of human papillomaviruses, Virus Res. 125 (2007) 135–144.
[136] R. Goraczniak, S.I. Gunderson, The regulatory element in the 3' untranslated region of human papillomavirus 16 inhibits expression by binding CUG binding protein 1, J. Biol. Chem. 283 (2008) 2286–2296.
[137] S.A. Cumming, M.G. McPhillips, T. Veerapraditsin, S.G. Milligan, S.V. Graham, Activity of the human papillomavirus type 16 late negative regulatory element is partly due to four weak consensus 5' splice sites that bind a U1 snRNP-like complex, J. Virol. 77 (2003) 5167–5177.
[138] S.I. Gunderson, M. Polycarpou-Schwarz, I.W. Mattaj, U1 snRNP inhibits pre-mRNA polyadenylation through a direct interaction between U1 70K and poly(A) polymerase, Mol. Cell 1 (1998) 255–264.
[139] S.A. Cumming, T. Chuen-Im, J. Zhang, S.V. Graham, The RNA stability regulator HuR regulates L1 protein expression in vivo in differentiating cervical epithelial cells, Virology 383 (2009) 142–149.
[140] P.A. Furth, C.C. Baker, An element in the bovine papillomavirus late 3' untranslated region reduces polyadenylated cytoplasmic RNA levels, J. Virol. 65 (1991) 5806–5812.
[141] M. Sokolowski, S. Schwartz, Heterogeneous nuclear ribonucleoprotein C binds exclusively to the functionally important UUUUU-motifs in the human papillomavirus type-1 AU-rich inhibitory element, Virus Res. 73 (2001) 163–175.
[142] W. Tan, S. Schwartz, The Rev protein of human immunodeficiency virus type 1 counteracts the effect of an AU-rich negative element in the human papillomavirus type 1 late 3' untranslated region, J. Virol. 69 (1995) 2932–2945.
[143] L. Wiklund, M. Sokolowski, A. Carlsson, M. Rush, S. Schwartz, Inhibtion of translation by UAUUUAU and UAUUUUUAU motifs of the AU-rich RNA instability in the HPV-1 late 3' untranslated region, J. Biol. Chem. 277 (2002) 40462–40471.
[144] C. Zhao, W. Tan, M. Sokolowski, S. Schwartz, Identification of nuclear and cytoplasmic factors that interact specifically with an AU-rich, cis-acting inhibitory sequence in the 3' untranslated region of human papillomavirus type 1 late mRNAs, J. Virol. 70 (1996) 3659–3667.
[145] M. Sokolowski, H. Furneaux, S. Schwartz, The inhibitory activity of the AU-rich RNA element in the human papillomavirus type 1 late 3' untranslated region correlates with its affinity for the elav-like HuR protein, J. Virol. 73 (1999) 1080–1091.
[146] M. Sokolowski, C. Zhao, W. Tan, S. Schwartz, AU-rich mRNA instability elements on human papillomavirus type 1 late mRNAs and c-fos mRNAs interact with the same cellular factors, Oncogene 15 (1997) 2303–2319.
[147] X. Li, C. Johansson, J. Glahder, A.K. Mossberg, S. Schwartz, Suppression of HPV-16 late L1 5'-splice site SD3632 by binding of hnRNP D proteins and hnRNP A2/B1 to upstream AUAGUA RNA motifs, Nucleic Acids Res. 22 (2013) 10488–10508.
[148] K. Nilsson, C. Wu, S. Schwartz, Role of the DNA damage response in human papillomavirus RNA splicing and polyadenylation, Int. J. Mol. Sci. 19 (2018).
[149] D.C. Anacker, C.A. Moody, Modulation of the DNA damage response during the life cycle of human papillomaviruses, Virus Res. 231 (2017) 41–49.
[150] C. Hao, L. Gong, X. Cui, J. Jönsson, Y. Zheng, C. Wu, N. Kajitani, S. Schwartz, Identification of heterogenous nuclear ribonucleoproteins (hnRNPs) and serine and arginine rich (SR) proteins that induce human papillomavirus type 16 late gene expression and alter L1 mRNA splicing, Arch. Virol. 167 (2022) 563–570.
[151] S. Jeon, B.L. Allen-Hoffmann, P.F. Lambert, Integration of human papillomavirus type 16 into the human genome correlates with a selective growth advantage of cells, J. Virol. 69 (1995) 2989–2997.
[152] S. Jeon, P.F. Lambert, Integration of human papillomavirus type 16 DNA into the human genome leads to increased stability of E6 and E7 mRNAs: implications for cervical carcinogenesis, Proc. Natl. Acad. Sci. U. S. A. 92 (1995) 1654–1658.
[153] M. Sokolowski, W. Tan, M. Jellne, S. Schwartz, mRNA instability elements in the human papillomavirus type 16 L2 coding region, J. Virol. 72 (1998) 1504–1515.
[154] B. Collier, L. Goobar-Larsson, M. Sokolowski, S. Schwartz, Translational inhibition in vitro of human papillomavirus type 16 L2 mRNA mediated through interaction with heterogenous ribonucleoprotein K and poly(rC)-binding proteins 1 and 2, J. Biol. Chem. 273 (1998) 22648–22656.