High-Risk Human Papillomavirus Oncogenic E6/E7 mRNAs Splicing Regulation

Yunji Zheng[1†], Xue Li[1†], Yisheng Jiao[2] and Chengjun Wu[2*]

[1] School of Pharmacy, Binzhou Medical University, Yantai, China, [2] School of Biomedical Engineering, Dalian University of Technology, Dalian, China

High-risk human papillomavirus infection may develop into a persistent infection that is highly related to the progression of various cancers, including cervical cancer and head and neck squamous cell carcinomas. The most common high-risk subtypes are HPV16 and HPV18. The oncogenic viral proteins expressed by high-risk HPVs E6/E7 are tightly involved in cell proliferation, differentiation, and cancerous transformation since E6/E7 mRNAs are derived from the same pre-mRNA. Hence, the alternative splicing in the E6/E7-coding region affects the balance of the E6/E7 expression level. Interrupting the balance of E6 and E7 levels results in cell apoptosis. Therefore, it is crucial to understand the regulation of E6/E7 splice site selection and the interaction of splicing enhancers and silencers with cellular splicing factors. In this review, we concluded the relationship of different E6/E7 transcripts with cancer progression, the known splicing sites, and the identified cis-regulatory elements within high-risk HPV E6/E7-coding region. Finally, we also reviewed the role of various splicing factors in the regulation of high-risk HPV oncogenic E6/E7 mRNA splicing.

Keywords: high-risk HPVs, cervical cancer, HNSCC, E6/E7, splicing, splicing factors

HUMAN PAPILLOMAVIRUS INFECTION AND RELATED CANCERS

Human papillomaviruses (HPV) are non-enveloped, small, double-stranded DNA viruses. The genome size of HPVs consists of 8,000 nucleotides (zur Hausen, 2002; zur Hausen, 2009). As one of the most common sexually transmitted diseases among almost all sexually active populations, a HPV-caused viral infection usually starts at the cutaneous or mucosal epithelium at the basal layer through sexual activity or a small wound. In most cases, a HPV infection is usually asymptomatic and will disappear within 2 years (Richardson et al., 2003). However, it will develop into a persistent infection in some rare cases, further developing into malignant lesions and progressing to cancer if left untreated (Parkin and Bray, 2006; Zheng and Baker, 2006). Depending on the association of HPV subtypes with cancer progression, they are divided into high-risk HPVs and low-risk HPVs (Zheng and Baker, 2006; Stanley, 2008).

About 15% of human cancers are caused by a viral infection; a high-risk HPV infection nearly accounts for half of all cases. According to the global cancer report in 2021, about 600,000 cancer

cases worldwide are reported to be associated with high-risk HPV infections, causing 340,000 deaths. Cervical cancer is one of the most prevalent cancer types tightly related to HPVs; more than 99% of cases are diagnosed with HPV infections. Moreover, recent studies have shown that anal cancer and head and neck cancer increase rapidly; about 95% of anal cancer and 80% of head and neck cancer are associated with a high-risk HPV infection (Arbyn et al., 2012). Up to now, more than 20 high-risk HPVs have been identified, among which HPV16 and HPV18 are the most popular high-risk strains (Lorincz et al., 1992; de Villiers et al., 2004; Bzhalava et al., 2013; de Villiers, 2013). Nearly half of cervical cancer cases are HPV16-positive, while HPV18 accounts for 20% of the cases (de Sanjose et al., 2010; Li et al., 2011).

High-risk HPV viral oncoproteins E6 and E7 are essential for viral replication and the termination of cell differentiation (Doorbar, 2005; Johansson and Schwartz, 2013). Furthermore, E6 and E7 play a vital role in cancer progression since they interact with cellular tumor suppressor proteins p53 and pRb. Transcription factor p53 contains a DNA binding region to activate the expression of the genes involved in DNA damage repair and cell apoptosis (Howie et al., 2009). HPV16 E6 oncoprotein interacts with p53 with the help of E6-associated protein, causing p53 degradation through the ubiquitination pathway. E6 interrupts p53 DNA binding activity, resulting in transcription inhibition by directly binding to p53. P300/CBP is a histone acetyltransferase essential for p53 production. E6 indirectly downregulates the p53 level by interacting with p300/CBP and inhibiting the p300/CBP expression level (Li and Coffino, 1996; Patel et al., 1999; Zimmermann et al., 2000). The loss of p53 may result in immortalization, followed by cancer occurrence (Scheffner et al., 1993; Martinez-Zapien et al., 2016). The tumor suppressor factor pRb plays a significant role in the regulation of cell cycle. Within the G1 phase, unphosphorylated pRb interacts with transcription factor E2F to suppress the transcription. Once pRb starts to phosphorylate gradually, the cell cycle steps into the S phase (Kato et al., 1993). HPV16 E7 interacts with pRB, causing the inactivation or degradation of pRB, thus leading to the release of E2F and thereby forcing the cell cycle to progress from G1 to the S phase (Munger et al., 2001).

The mRNAs that encode E6 and E7 are bicistronic and generated from the same pre-mRNA (Mesplede et al., 2012). Various splice sites are identified in high-risk HPV E6- and E7-coding regions. Alternative splicing decides the selection of the splice site to produce different mRNA expression variants. However, no splicing in E6 and E7 in low-risk HPVs has been reported since all the spliced E6/E7 mRNA isoforms excluded the most part of the E6-coding sequence. Hence, E6 is only translated from unspliced mRNA, and E7 is generated from several spliced mRNAs but not the unspliced mRNAs. One can speculate that the production of E7 produced from the unspliced mRNA is too low to reach the threshold that can be detected (Zheng and Baker, 2006; Ajiro and Zheng, 2014). The balance of E6/E7 level is directly controlled by alternative splicing, thereby impacting the fate of infected cells developing to apoptosis or immortalization. This review concerns the regulation of E6/E7 splicing isoforms, summarizing the identified cis-regulatory elements and the splicing factors involved in E6/E7 splicing.

RNA PROCESSING IN GENERAL

The gaps in between the intronic region and the exonic coding regions are 5′ splice site (5′ ss) and 3′ splice site (3′ ss). The 5′ ss contains a highly conserved invariable "GU" dinucleotide, while the 3′ ss contains a conserved "AG" dinucleotide (Black, 2003; Hallegger et al., 2010). Besides the splice sites, a branch point sequence (BPS) contains a conserved adenine located at 1–40 nucleotides upstream of 3′ ss, and a polypyrimidine tract is also needed for splicing (Taggart et al., 2012). The general process of mRNA processing includes 5′ capping, polyadenylation, and mRNA splicing. 5′ capping adds a 7-methylguanosine cap to the 5′ end of the growing transcript by a 5′- to -5′ phosphate linkage (Topisirovic et al., 2011). If 5′ capping provides the ribosome with a "start-checkpoint" for translation and mRNA stabilization, the 3′ polyadenylation is an "end-check point" by adding a poly-A tail to the mRNA (Bienroth et al., 1993; Beaudoing et al., 2000). Pre-mRNA splicing is a process that brings exons together in different combinations by removing the introns to generate mature mRNA-encoding functional proteins. The occurrence of splicing requires cis-acting elements that function as splicing enhancers or silencers to interact with trans-acting factors and further constitute spliceosomes (Black, 2003; Hallegger et al., 2010). Besides RNA splicing, post-transcriptional modification is also essential to mature mRNA production, such as N6-methyl adenosine (m6A), which is the most prevalent conserved internal modification in prokaryotic and eukaryotic RNAs (Liu and Pan, 2016). However, we will only focus on the regulation of HPV oncogene E6/E7 pre-mRNA splicing in this review.

HPV GENE EXPRESSION IS REPLICATION CYCLE DEPENDENT

Various splice sites have been identified in the HPV genome; the selection of different splice sites generates a variety of HPV mRNA transcripts to ensure that HPV expresses enough viral protein to drive their replication cycle. Therefore, the regulation of HPV gene expression is highly linked to the HPV replication cycle (Doorbar, 2005; Johansson and Schwartz, 2013). The HPV replication cycle can be divided into early and late stages. The study of the HPV replication cycle is best characterized for HPV16. In HPV16, the early promoter p97 is active during the early phase, while the late promoter p670 is active at the late phase. The HPV16 infection usually starts by gaining entry to the host cell's nucleus in an epithelial cell located at the basal layer of the mucosal epithelium. The cellular transcription machinery helps the HPV16 genome to start replicating from early promoter p97 (Egawa, 2003; Borgogna et al., 2012; Howley and Pfister, 2015; Schafer et al., 2015). HPV16

expresses only early genes during the early stage, including E6, E7, E1, E2, E4, and E5. The HPV16 E2 protein accumulates during cell differentiation and shuts down the HPV early promoter p97. Consequently, HPV16 E6 and E7 expressions are shut off, and the HPV16 differentiation-dependent late promoter p670 is activated (Shin et al., 2007). The late promoter p670 mediates several early viral gene expressions, including E1, E2, and E4. However, when the HPV replication cycle is switched to the late stage, p670 mediates only late gene expression, including L1, L1i, and L2. All the early mRNAs produced by promoter p97 and p670 are polyadenylated at pAE, but high levels of HPV E2 protein inhibit HPV pAE, leading to the activation of HPV late L1 and L2 gene expression. The late mRNAs produced by HPV late promoter p670 are polyadenylated at pAL (Johansson and Schwartz, 2013) (**Figure 1**).

Although most of the HPV infections will be cleared by the host immune system within 2 years, some rare cases may become persistent infections and eventually progress to cancer if left untreated. It is worth mentioning that the occurrence of cancer needs to repress or mute the E2 expression since E2 plays a vital role in early gene expression (Hanahan and Weinberg, 2000). In the majority of cervical cancer cases, high-risk HPVs' genome integrates into the host genome, and the early viral gene expression regulator E2 protein is disrupted (Pyeon et al., 2009). Within this stage, progeny virion is not produced anymore, and the overexpression of E6 and E7 is induced (Groves and Coleman, 2015). A high level and constant E6 and E7 oncoprotein expression are required for tumor formation and maintenance. However, some cervical cancer cases demonstrated that high-risk-HPV DNA keeps in an episomal

form. In these cases, the differential methylation of E2 binding sites 1, 3, and 4 in long control region is related to the overexpression of the E6 and E7 proteins, further suggesting that HPV genome integration into the host genome may not be essential for malignant cellular transformation (Matsukura et al., 1989; Goodwin and DiMaio, 2000; Vinokurova et al., 2008).

HPVS E6/E7 SPLICING ISOFORM AND FUNCTIONS

HPV expresses multiple viral proteins to maintain the viral replication cycle. The most common high-risk subtypes HPV16 and HPV18 have been studied thoroughly. In HPV16, various splicing donors and acceptors have been characterized (**Figure 2**). The selection of different splicing donors and splicing acceptors decides the generation of viral proteins. The splicing donor SD226 alternatively spliced to SA409, SA526, and SA742 generates the E6*I, E6*II, and E6^E7 variants (**Figure 2**) (Guccione et al., 2004; Ajiro et al., 2012; Manzo-Merino et al., 2014; Olmedo-Nieva et al., 2018; Zheng et al., 2020). In HPV18, splicing donor SD233 and splicing acceptors SA416, SA3434, SA2779, and SA791 are involved in producing E6/E7 transcript variants (Vazquez-Vega et al., 2013; Cerasuolo et al., 2017; Olmedo-Nieva et al., 2018). The coding region of high-risk HPV E6/E7 contains at least one splicing donor, and an acceptor is used for alternative splicing to generate one or several isoforms termed E6*, but low-risk HPVs and beta-papillomavirus do not have such splicing sites in the E6/E7-
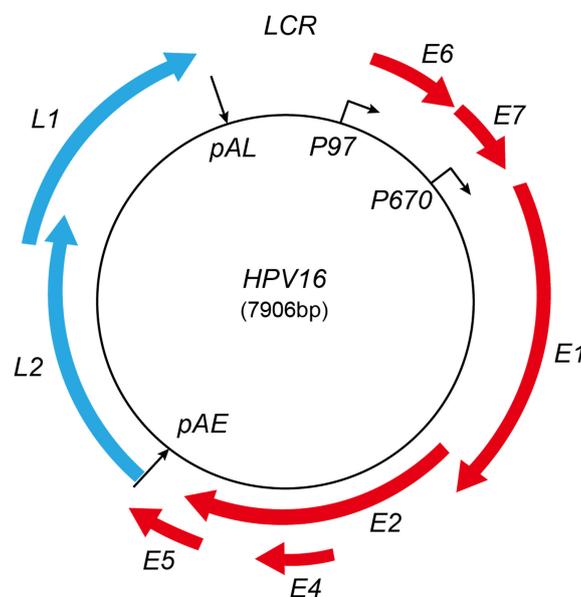
FIGURE 1 | Schematic map of HPV16 genome. Red represents early proteins; blue represents late proteins. LCR, long control region; pAE, early polyadenylation signal; pAL, late polyadenylation signal. The early and late promoters are indicated. Red boxes represent viral early-protein-encoding regions; blue boxes represent late-protein-encoding regions.

FIGURE 2 | Schematic presentation of E6/E7 transcripts. The unspliced mRNA in E6 coding region generates E6 (Zheng et al., 2020), SD226^SA409 generates E6*I and E7, SD226^SA526 generates E6*II and E7, and SD226^742 generates E6^E7. All the transcripts are detected in primary head and neck cancer cell line HN26 and subgenomic plasmid pC97ElsL transfected HeLa cells.
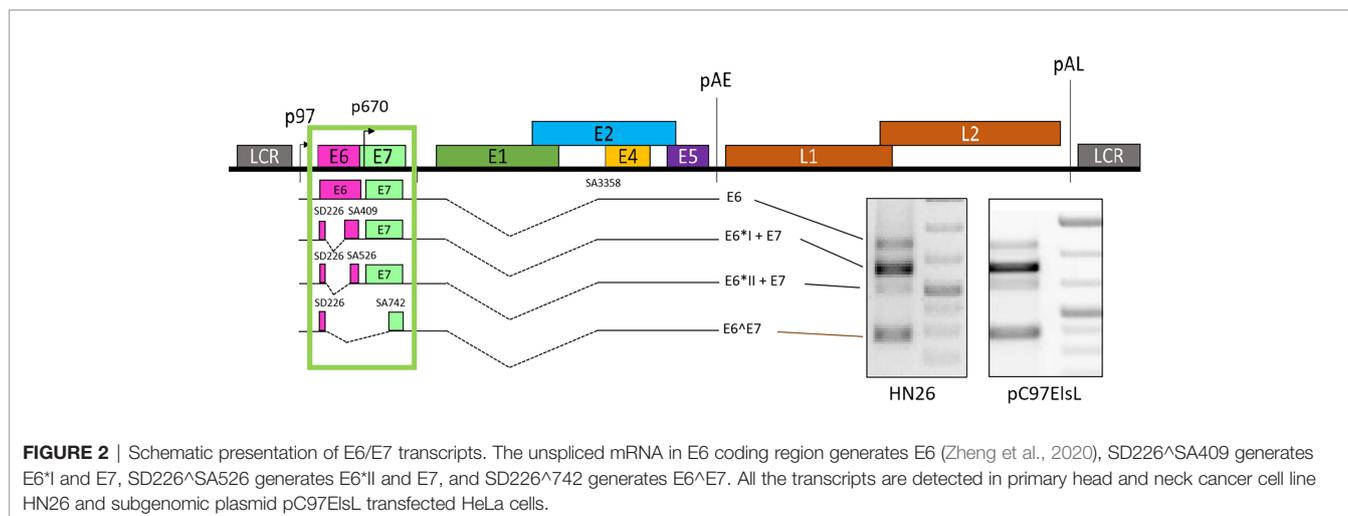
coding region to conduct alternative splicing. We concluded the known splice sites and E6/E7 mRNA variants of high-risk HPV as those shown in **Table 1**. Since E6 and E7 are oncoproteins and play a vital role in high-risk HPV-induced cancers, the splicing mechanism of E6 and E7 is of great interest for study (Graham and Faizo, 2017; Cerasuolo et al., 2020). HPV16 oncoprotein E6 and E7 are generated from the same pre-mRNA. It is disputable that the unspliced mRNA produces an early viral E6 protein since the unspliced mRNA contains the full-length E6-encoding gene. The mRNA spliced from SD226 to SA409 produces the E6*I transcript, used for E7 production (Bernard et al., 1987). In high-risk HPV, the splicing within the E6-coding region to produce E6*I mRNA is a landmark to distinguish high- and low-risk HPV (Guccione et al., 2004; Mesplede et al., 2012; Filippova et al., 2014) replication cycle. A previous study demonstrates that the most abundant transcript isoform in HPV16-positive cervical malignant and oropharyngeal lesion is E6*I. In agreement with this observation, consistent results were identified in various cancer cell lines (Guccione et al., 2004) (**Figure 2**).

The splicing between HPV16 splice site SD226 and SA409 results in the shortening of the E6-coding region. The shortened coding region produces a small protein E6*I; the E6*I open reading frame (Bernard et al., 1987) contains a weak ATG located upstream of the E7 ORF. Therefore, the ribosome may reinitiate translation from E7 ATG by leaky scanning. Zheng ZM et al. demonstrated that most E7 proteins are translated from E6*I mRNA in HPV16-positive cervical cancer cells (CaSki) and HPV18-positive HeLa cells (Schafer et al., 2015). Although increasing evidence supports this conclusion, several individual studies also presented different opinions that unspliced E6E7 RNA expresses more E7 than spliced variants. It is necessary to point out that these findings were drawn from a less stringent experimental system (Filippova et al., 2007; Manzo-Merino et al., 2014). E6*I is a multi-functional protein that has been detected in a HPV16-positive CaSki cell line and functions as an antagonist of the full-length E6 protein, further suggesting its antitumorigenic role. In HPV18, E6*I protein indirectly

promotes the expression of the potential cervical cancer diagnostic marker p14ARF through p53 degradation. The overexpression of E6*I results in a slight increase of p14ARF. This result suggests that HPV18 E6*I protein may interact with p53 to prevent p53 from regulating p14ARF (Vazquez-Vega et al., 2013). Several earlier studies demonstrate that E6*I interacts with p53 to impact p53 degradation. Moreover, it has been shown that the E6*I protein does not drive keratinocyte immortalization and proliferation (Sedman et al., 1991; Pim et al., 1997). However, HPV16 produces other E6/E7 splicing isoforms, the mRNA SD226^SA526 encodes E6*II protein, while the mRNA SD226^SA742 produces E6^E7 protein. The ratio of unspliced E6, E6*I, E6*II, and E6^E7 mRNAs plays a vital role in HPV-related cancer progression. The variant E6^E7 is generally present at the lower level, while E6*I is the most abundant mRNA expressed in HPV16-infected cells or HPV16-related cancers (Ajiro et al., 2012; Ajiro et al., 2016; Olmedo-Nieva et al., 2018; Zheng et al., 2020). In addition, the splicing variant E6*II mRNA expression level is regularly higher than that of the unspliced E6 mRNA but lower than the major spliced product E6*I. However, the E6*I and E6*II mRNA levels vary in different cancers; cervical cancer samples express higher levels of E6*I and E6*II mRNAs than those in oropharyngeal cancer (Cerasuolo et al., 2017). E6*II is a suboptimal E7 producer (Schmitt et al., 2010; Schmitt and Pawlita, 2011) that has been associated with the grade of cervical lesions, but this association could not be confirmed. Very little is known about the functions of E6^E7 protein, but it has been reported to assist in the stabilization of E6, E7, and E6*I (Ajiro and Zheng, 2015). Furthermore, several suboptimal splice sites have been identified in HPV16 to generate E6/E7 mRNA variants within the E6 ORF, including SD191^SA409 (E6*VI), SD221^SA409 (E6*V), SD174^SA718 (E6^E7*I), and SD221^SA850 (E6^E7*II) (Ajiro et al., 2012; Islam et al., 2017). A previous study indicates that HPV16-positive breast cancer samples detect various E6/E7 transcript isoforms, including E6*I, E6*II, E6^E7, E6^E7*I, and E6^E7*II. This finding does not elucidate the role of these isoforms in breast cancer progression but suggests that HPV16 is

TABLE 1 | Summary of identified E6/E7 splicing variants in 20 high-risk human papillomaviruses (HPVs) (Olmedo-Nieva et al., 2018).

HPV type	E6/E7 transcripts	5′ ss–3′ ss (nucleotide positions)
HPV16	E6*I	226–409 (Doorbar et al., 1990)
	E6*II	226–526 (McFarlane et al., 2015)
	E6*III	226–3,358 (Doorbar et al., 1990)
	E6*IV	226–2,709 (Schmitt et al., 2010)
	E6*V	221–409 (Ajiro et al., 2012)
	E6*VI	191–409 (Ajiro et al., 2012)
	E6^E7 (E6*X)	226–742 (Ajiro and Zheng, 2015)
	E6^E7*I	174–718 (Islam et al., 2017)
	E6^E7*II	221–850 (Islam et al., 2017)
HPV18	E6*I	233–416 (Toots et al., 2014; Ajiro et al., 2016)
	E6*II	233–3,434 (Toots et al., 2014; Ajiro et al., 2016)
	E6*III	233–2,779 (Ajiro et al., 2016)
	E6^E7	233–791 (Ajiro and Zheng, 2015; Ajiro et al., 2016)
HPV26	E6*I	173–406 (Mesplede et al., 2012)
HPV31	E6*I	210–413 (Hummel et al., 1992; Ozbun, 2002)
	E6*III	210–3,295 (Ozbun, 2002)
HPV33	E6*I	231–509 (Snijders et al., 1992)
	E6*II	231–785 (Snijders et al., 1992)
	E6*III	231–3,351 (Snijders et al., 1992)
HPV35	E6*I	228–419 (Mesplede et al., 2012)
HPV39	E6*I	231–420 (Mesplede et al., 2012)
HPV45	E6*I	230–414 (Halec et al., 2013)
HPV51	E6*I	173–406 (Mesplede et al., 2012)
HPV52	E6*I	224–502 (Halec et al., 2013)
HPV53	E6*I	236–419 (Halec et al., 2013)
HPV56	E6*I	157–420 (Mesplede et al., 2012)
HPV58	E6*I	232–510 (Li et al., 2013)
	E6*II	232–3,355 (Li et al., 2013)
HPV59	E6*I	183–582 (Halec et al., 2013)
HPV66	E6*I	157–420 (Mesplede et al., 2012)
HPV67	E6*I	224–502 (Halec et al., 2013)
HPV70	E6*I	231–422 (Halec et al., 2013)
HPV73	E6*I	227–410 (Halec et al., 2013)
HPV82	E6*I	178–411 (Mesplede et al., 2012)

transcriptionally active (Islam et al., 2017). However, the functional feature of these mRNA variants in HPV16-related cancer progression is largely unknown.

HPV E6/E7 SPLICING REGULATION

The splicing regulation is tightly related to the splicing site selection regulated by the interaction of cis-acting elements and trans-acting factors. cis-acting elements provide binding motifs to interact with cellular splicing factors resulting in splicing enhancement or splicing inhibition. The cis-acting elements located in exons enhancing splicing are termed exonic splicing enhancers (ESE), while those of inhibiting splicing are termed exonic splicing silencers (ESS). Similarly, the cis-acting elements located in introns are also termed intronic splicing enhancers (Garland et al., 2007) and intronic splicing silencers (ISS). The trans-acting factors interact with cis-regulatory elements to facilitate the composition of spliceosomes, further conducting splicing from splice sites. The expression level of trans-acting factors varies in different tissues, and they may compete for binding to the cis-acting elements to promote

splicing or inhibit splicing (Manley and Tacke, 1996; Wang et al., 2006; Zhou and Fu, 2013). The families of serine/arginine-rich proteins (SR proteins) and heterogeneous nuclear ribonucleoproteins (hnRNPs) are known splicing factors. The SR proteins usually function as splicing activators binding to ESE or ISE, while hnRNPs function as splicing silencers binding to ESS or ISS (Graveley et al., 1999; Hertel and Maniatis, 1999; Han et al., 2010). However, this classification is not strict; increasing evidence indicates that the role of SR and hnRNP proteins in splicing regulation may alter (Lim et al., 2011). Their opposite function in splicing depends on the positions of the splicing regulatory elements relative to the splice site. Besides SR and hnRNP proteins, various splicing factors participate in the pre-mRNA splicing process (Elliott et al., 1998).

Cis-Regulatory Elements in E6/E7-Coding Region

The study on the mechanism of HPV splicing was well performed on the most common subtypes, including HPV16 and HPV18. Several cis-regulatory elements have been mapped, including BPS and splicing silencers. Moreover, various splicing factors have been identified to impact the production of expression of major splicing variants E6*I and E7.

Researchers have identified the sequence of HPV16 BPS as AACAAAC located within the E6-coding region upstream of 3′ ssSA409. The adenosine at nucleotide 385 (underlined) is the branch site to dominate the splicing to 3′ ss SA409, affecting the E7 expression. Once the point mutation was introduced to the branch site in BPS, the BPS binding activity to U2 protein was interrupted, resulting in 3′ ss SA409 splicing inhibition to 3′ ss SA409. Therefore, the pre-mRNAs contain the mutated BPS, causing inefficient splicing activity to produce very little E7 protein. Moreover, since the BPS locates in the E6-coding region, the point mutation at BPS generates the mutated E6 protein, which has little effect on p53 degradation (Ajiro et al., 2012). An earlier study identified a suboptimal BPS (AGUGAGU) on HPV16, from nucleotide 323 to 329 in the E6-coding region. The underlined guanosine at nucleotide 328 is a cryptic branch site. This cryptic BPS may play its role in keeping the E6/E7 mRNA level, but no information about E6/E7 pre-mRNA splicing pattern change has been reported (De la Rosa-Rios et al., 2006) (**Figure 3A**). In HPV18, researchers have identified two alternative branch sites for E6*I generation and E7 production. The mapped BPS is AACUAAC (from nucleotide 383 to 389); both branch sites are adenosine: one adenosine locates at nucleotide 384, and the other locates four nucleotides downstream at 388 (underlined). The selection of two alternative branch sites decides the efficiency of the E6*I splice, further affecting the production of HPV18 E6 and E7. The study has demonstrated that the E6*I splice favors the branch site at nucleotide 388 rather than the adenosine at nucleotide 384. If the preferred branch site is inactivated or mutated, cryptic splicing to 3′ ss SA636 would be activated (Brant et al., 2019) (**Figure 3B**). In general, the identified HPV18 BPS in the E6-coding region is similar to the BPS in HPV16. However, the fourth nucleotide in the BPS is differed—uridine for HPV18 and adenosine for HPV16.

In addition to BPS identification, a splicing silencer in the HPV18 E7-coding region at nucleotide 612 to 639 has been
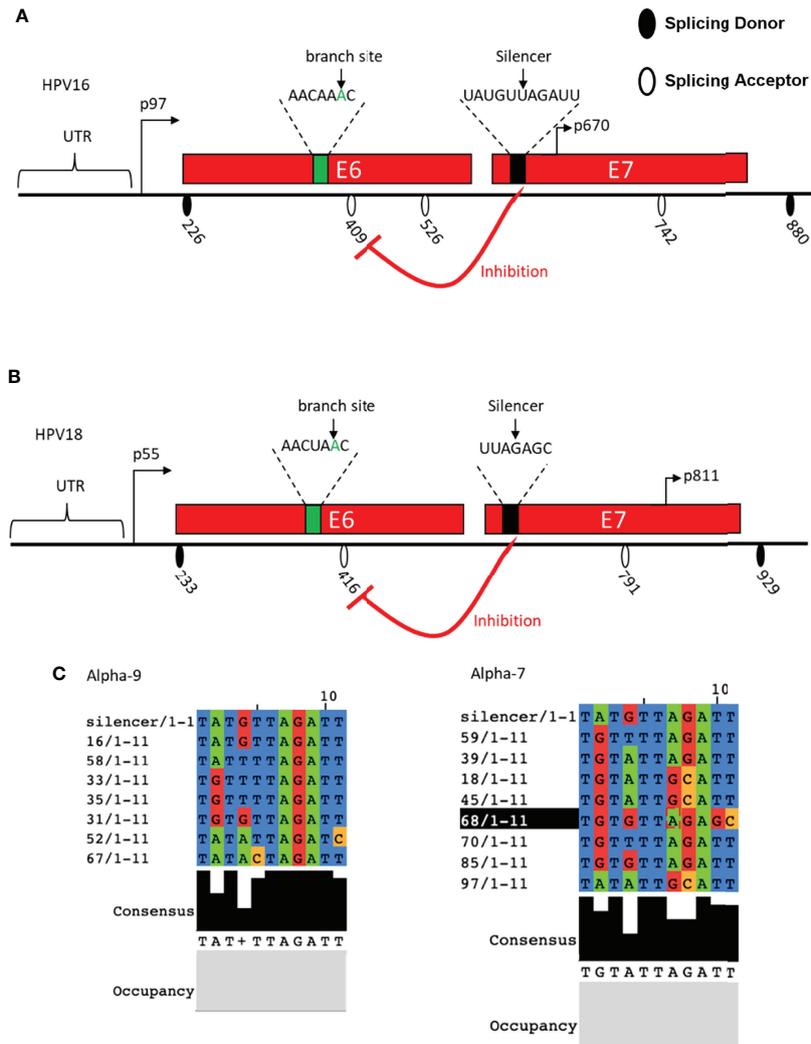
FIGURE 3 | Schematic presentation of cis-regulatory elements identified in HPV16 and HPV18. **(A)** The HPV16 E6- and E7-encoding regions and the splice sites are indicated. Identified branch site and SA409 silencer are indicated. **(B)** HPV18 E6- and E7-encoding regions and the splice sites are indicated. Identified branch site and SA416 silencer are indicated. **(C)** HPV16 silencer sequence alignment with alpha-9 and alpha-7 groups.

mapped to inhibit the splicing from 5′ ss SD233 to 3′ ss SA416 by binding to cellular splicing factor hnRNP A1. The interaction between the mapped silencer and hnRNP A1 inhibits HPV18 E6*I production and E7 expression. These 27 nucleotides in the length silencer contains two hnRNP A1 binding motifs: the sequence of motif one is AAGACA, and that of the other is UUAGAGC. The mutation of the motif AAGACA did not affect the SD233^SA416 splicing efficiency, but the mutation of UUAGAGC resulted in an increase in E6*I expression (Ajiro et al., 2016). This outcome suggested that the sequence UUAGAGC contributes to the interaction with hnRNP A1. Similarly, the HPV16 E7-coding region contains a splicing silencer interacting with hnRNP A1/A2 to inhibit splicing to 3′ ss SA409, thereby reducing the expression of E6*I and E7 production. This silencer locates at nucleotide 594 to 604,

consisting of UAUGUUAGAUU. Since HPV16 E7 genetic conservation is essential to carcinogenesis, sequence alignment was therefore performed, and it indicated that the HPV16 11 nucleotide in the length silencer is not well conserved in all high-risk HPVs (Zheng et al., 2020). However, the UAGAU is completely conserved in HPV16 belonging to the alpha-9 subgroup. In alpha-7 group, UAGAU is conserved with HPV18, HPV39, HPV59, HPV70, and HPV85 but less conserved with HPV45, HPV68, and HPV97. Interestingly, all high-risk HPVs have a "UA"-rich region close to the E7 start codon, and one can speculate that the "UA"-rich region is necessary to the silencer's inhibitory activity, which interacted with hnRNP A1/A2 to manipulate the balancing of the E6/E7 protein level and is required for malignant transformation (Zheng et al., 2020) (**Figure 3C**).

Splicing Factors Involved in E6/E7 mRNA Expression

According to the effect of splicing regulation, splicing factors are classified as splicing activator and splicing repressor. However, this division is not strict; the effect of splicing factors can switch depending on the position of the splicing regulatory elements and their activity. The regulation of both HPVs E6/E7 gene expression is tightly regulated by several splicing factors, thereby affecting the splicing pattern change (**Figure 4**).

hnRNP A1/A2/G/D

The splicing factor hnRNP A1 was identified as a splicing repressor interacting with the mapped splicing silencer (UUAGAGC) located at the HPV18 E7-coding region to inhibit splicing to 3′ ss SA416, resulting in a reduction of E6*I and E7 protein (Ajiro et al., 2016). Nevertheless, this study did not reveal how hnRNP A1 interacts with the silencer. In a recent study, the researchers reported that hnRNP A1/A2 inhibits the splicing to 3′ ss SA409 on HPV16 but that this resulted in different consequences. The overexpression of hnRNP A1 leads to an increase of unspliced E6 mRNAs at the expense of E7 mRNAs, while the overexpression of hnRNP A2 leads to alternative splicing to downstream 3′ ss SA742. The splice site SA742 is used to produce E6^E7, E1, and E4 mRNAs (Cerasuolo et al., 2017). Moreover, researchers revealed that the hnRNP A1 inhibitory effect is contributed by the interaction of hnRNP A1 C-terminus with the splicing silencer (UAGAU) in the E7-coding region (Zheng et al., 2020). Taken together, as one of the most abundant splicing factors, hnRNP A1 is reported to participate in the regulation of HPV16 early and late gene expressions, causing an E6/E7 splicing pattern change and demonstrating its vital role in the HPV replication cycle as well as malignant transformation.

A recent study demonstrates that hnRNP G inhibits splicing from HPV16 3′ ss SA409 to prevent HPV16 E7 production. The reduction of E7 levels restores the pRB levels, while the absence of hnRNP G causes a reduction of pRB. The inhibitory effect of hnRNP G may be due to the direct interaction of hnRNP G with the previously reported splicing silencer in the E7-coding region.

Meanwhile, the overexpression of hnRNP G promotes the splicing from HPV16 3′ ss SA2709 to increase the production of E2 protein. The different effects of hnRNP G in the regulation of E6, E7, and E2 splicing are contributed by the different domains of hnRNP G. The hnRNP G N-terminal RGG domain and the NTD domain play an essential role in the inhibition of E6/E7 mRNA splicing, while the C-terminal region is required for E2 production. Since E2 protein shuts down the HPV16 early promoter, it thereby repressed the expression of E6 and E7 oncoproteins to pave the way for differentiation and late gene expression. The high levels of pro-apoptotic E2 protein caused by the overexpression of hnRNP G would counteract cell transformation and HPV16-related cancer progression, suggesting that hnRNP G may be used for anticancer or antiviral treatment (Hao et al., 2022).

hnRNP D has four variants, including hnRNP D37, D40, D42, and D45. All four variants have been shown to promote the intron retention of HPV16 E6 mRNAs, resulting in increasing E6 mRNA levels at the expense of E7 levels. The RRM1 and RRM2 domains of hnRNP D40 are necessary for the interaction of hnRNP D40 with HPV16 mRNAs, while the inhibitory effect of hnRNP D40 is contributed by its N-terminal region and C-terminal RGG domain. Besides splicing inhibition, hnRNP D40 promotes the intron-retained E6 mRNA expression level in the cytoplasm (Cui et al., 2022).

Epidermal Growth Factor

The epidermal growth factor (EGF) regulates HPV16 E6/E7 splicing via inducing the activation of the ERK1–ERK2 pathway, which inhibits splicing to 3′ ss SA409. The inhibitory effect of EGF results in the reduction of E6*I but gives rise to unspliced E6 mRNA generation (Rosenberger et al., 2010). However, the mechanism of EGF in the regulation of E6/E7 splicing remains a subject for further study.

5′ Cap-Binding Factors

The previous study has reported that the splicing on 5′ capped E6/E7 mRNAs prefers the 5′ ss SD226 on HPV16. The cap-

FIGURE 4 | Schematic presentation of identified splicing factors involved in the regulation of human papillomaviruses early oncoprotein E6/E7 RNA splicing. HPV16-encoded viral genes are represented by a "red" box. The splice sites are indicated.
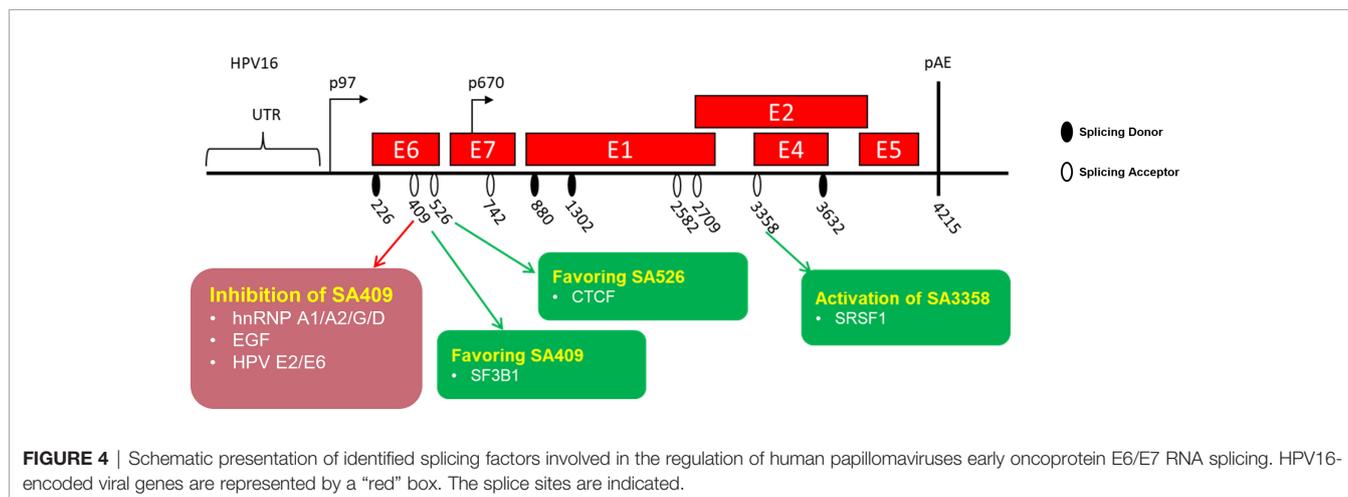
dependent HPV16 E6/E7 splicing in cervical cancer-derived cells is particularly efficient and supposed to be conducted by interacting with cap binding complex. Furthermore, the E6/E7 splicing efficiency from 5′ ss SD226 is impacted by the distance of the cap-proximal intron to the 5′ cap. Further studies identified that the optimal distance is less than 307 nucleotides, which promotes the association of 5′ splice site with U1 snRNP. Once the distance of the 5′ cap to the SD226 increases, splicing is inhibited, promoting unspliced E6 mRNA expression and decreasing E6*I production (Zheng et al., 2004).

SRSF1/SRSF2

The HPV16 3′ ss SA3358 is the most commonly used splice site in the HPV16 genome. SA3358 generates early mRNAs encoding early proteins, including E6 and E7. The splicing factor serine/arginine-rich (SR) protein SRSF1 has been reported to enhance splicing to 3′ ss SA3358 by interacting with a splicing enhancer (ACCGAAGAA) located downstream of the splice site SA3358. Since SA3358 is used by the majority of the E6 and E7 mRNAs, high levels of SRSF1 function as proto-oncoprotein to upregulate the expression of E6/E7, further resulting in hyperproliferation and, possibly, cancer occurrence (Somberg and Schwartz, 2010). Another SR protein SRSF2 has been shown to be required for E6E7 mRNA production only in cervical cancer-derived cells. The role of SRSF2 is to maintain E6/E7 mRNA stability and inhibit their decay. The knock-down of SRSF2 in cervical tumor cells results in cell apoptosis, suggesting its oncogenic role in cervical cancer progression (McFarlane et al., 2015).

HPVs E2 and E6 Proteins

HPV16 early E2 protein and early oncoprotein E6 have been shown to have an inhibitory effect on HPV16 E6/E7 pre-mRNA splicing (Gomes and Espinosa, 2010). HPV16 E2 and E6 directly bind to the intron located in between SD226 and SA409 through their RNA interaction domain in the C-terminus, resulting in the reduction of the E6*I isoform in HPV16-infected cells (Gomes and Espinosa, 2010). This inhibitory effect may also be due to E2 and E6 interfering with several SR proteins, including SRSF9, SRSF6, and SRSF5. Taken together, these findings demonstrate that HPV16 viral proteins E2 and E6 function as splicing regulatory proteins which affect E6/E7 pre-mRNA splicing through their interaction with SR proteins (Bodaghi et al., 2009).

CTCF and SF3B1

The CCCTC-binding factor known as CTCF has been firstly identified as a zinc finger DNA-binding transcription factor. Over the past decade, the new tricks of CTCF has been elucidated, especially its new role as a RNA-binding alternative splicing factor (Gomes and Espinosa, 2010; Shukla et al., 2011; Guo et al., 2012; Monahan et al., 2012; Marina et al., 2016). A previous study has shown that CTCF binds to the HPV18 E2-coding region to increase the production of the E6*II isoform. The mutation of CTCF results in the increased production of unspliced E6 mRNA and spliced E6*I mRNA (Paris et al., 2015). It is reasonable to speculate that the increase of E6*II at the expense of unspliced E6 and E6*I mRNA has interrupted the balance of the E6/E7 ratio, causing infected cell apoptosis. However, the mechanism of CTCF in the regulation of E6/E7 expression was not well understood until recently. Researchers have revealed that the inhibitory effect of CTCF is dependent on CTCF and YY-1 loop formation in the HPV18 genome. The downregulation of YY-1 causes an interruption of loop formation and a reversal of epigenetic silencing (Pentland et al., 2018).

The splicing factor 3B1 (SF3B1) was reported to modulate HPV16 E6/E7 pre-mRNA splicing, resulting in increased E6*I in HPV16-positive head and neck cancer cells. The treatment of SF3B1 inhibitor meayamycin B causes the downregulation of the transcript E6*I but promotes the level of the unspliced E6 mRNAs (Gao et al., 2014). These findings further suggest the anti-oncogenic role of SF3B1 in HPV16-derived tumorigenesis. However, the exact mechanism of how SF3B1 affects HPV16 splicing has not been elucidated.

CONCLUSION

At present, there is no effective treatment for HPV infection and HPV-related cancer. Although there are vaccines against several high-risk HPV types, the vaccines have no effect on infected patients, and vaccines are not universally accessible due to the high price. Therefore, continuous research on the carcinogenic mechanism of HPV is of great significance for the prevention and treatment of HPV infections and their related cancers. As to the oncogenic proteins expressed in the early replication cycle of high-risk HPVs, the splicing regulation mechanism of E6 and E7 pre-mRNA has a profound impact on the entire replication cycle of HPVs. Moreover, the differentiation, proliferation, and immortalization transformation of HPV-infected cells are tightly related to the level of E6 and E7. Transcripts encoding E6 and E7 proteins are derived from the same pre-mRNA; the selection of splice site SA409 in HPV16 yields the most abundant isoform E6*I, which is further translated into the oncoprotein E7; the unspliced mRNA is translated into the oncoprotein E6 since it contains the complete E6-coding region. Such a mechanism indicates that a perfect balance between E6 and E7needs is required. In HPV16, if the splicing from SA409 becomes too efficient, the production of E6*I will be upregulated, resulting in increased E7 level, but E6 level is decreased, thereby causing apoptosis of the infected cells. Conversely, if the unspliced E6 mRNA is upregulated, the level of E7 will decrease, and the fate of the infected cells will go toward apoptosis. This balancing mechanism between E6 and E7 is regulated by the interaction of cellular splicing factors with cis-regulatory RNA elements on HPV pre-mRNA. A series of RNA regulatory elements and splicing factors are identified

to regulate E6/E7 mRNA splicing, indicating that the E6/E7 expression levels are tightly regulated during posttranscription. The role of various splicing factors in HPV-related cancer progression has also been characterized. Overexpression or knockdown of splicing factors will interrupt the ratio of E6/E7 proteins, thereby causing infected cell apoptosis. This weak point of high-risk HPV provides us with an opportunity to identify small molecules that can be used as a splicing factor inhibitor, paving the way to explore more novel antiviral and anticancer drugs.

REFERENCES

Ajiro, M., Jia, R., Zhang, L., Liu, X., and Zheng, Z. M. (2012). Intron Definition and a Branch Site Adenosine at Nt 385 Control RNA Splicing of HPV16 E6*I and E7 Expression. *PloS One* 7 (10), e46412. doi: 10.1371/journal.pone.0046412

Ajiro, M., Tang, S., Doorbar, J., and Zheng, Z. M. (2016). 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 (20), 9138–9152. doi: 10.1128/JVI.00965-16

Ajiro, M., and Zheng, Z. M. (2014). Oncogenes and RNA Splicing of Human Tumor Viruses. *Emerg. Microbes Infect.* 3 (9), e63. doi: 10.1038/emi.2014.62

Ajiro, M., and Zheng, Z. M. (2015). E6^E7, a Novel Splice Isoform Protein of Human Papillomavirus 16, Stabilizes Viral E6 and E7 Oncoproteins *via* HSP90 and GRP78. *mBio.* 6 (1), e02068–e02014. doi: 10.1128/mBio.02068-14

Arbyn, M., de Sanjose, S., Saraiya, M., Sideri, M., Palefsky, J., Lacey, C., et al. (2012). EUROGIN 2011 Roadmap on Prevention and Treatment of HPV-Related Disease. *Int. J. Cancer.* 131 (9), 1969–1982. doi: 10.1002/ijc.27650

Beaudoing, E., Freier, S., Wyatt, J. R., Claverie, J. M., and Gautheret, D. (2000). Patterns of Variant Polyadenylation Signal Usage in Human Genes. *Genome Res.* 10 (7), 1001–1010. doi: 10.1101/gr.10.7.1001

Bernard, H. U., Oltersdorf, T., and Seedorf, K. (1987). Expression of the Human Papillomavirus Type 18 E7 Gene by a Cassette-Vector System for the Transcription and Translation of Open Reading Frames in Eukaryotic Cells. *EMBO J.* 6 (1), 133–138. doi: 10.1002/j.1460-2075.1987.tb04730.x

Bienroth, S., Keller, W., and Wahle, E. (1993). Assembly of a Processive Messenger RNA Polyadenylation Complex. *EMBO J.* 12 (2), 585–594. doi: 10.1002/j.1460-2075.1993.tb05690.x

Black, D. L. (2003). Mechanisms of Alternative Pre-Messenger RNA Splicing. *Annu. Rev. Biochem.* 72, 291–336. doi: 10.1146/annurev.biochem.72.121801.161720

Bodaghi, S., Jia, R., and Zheng, Z. M. (2009). Human Papillomavirus Type 16 E2 and E6 are RNA-Binding Proteins and Inhibit *In Vitro* Splicing of pre-mRNAs With Suboptimal Splice Sites. *Virology.* 386 (1), 32–43. doi: 10.1016/j.virol.2008.12.037

Borgogna, C., Zavattaro, E., De Andrea, M., Griffin, H. M., Dell'Oste, V., Azzimonti, B., et al. (2012). Characterization of Beta Papillomavirus E4 Expression in Tumours From Epidermodysplasia Verruciformis Patients and in Experimental Models. *Virology.* 423 (2), 195–204. doi: 10.1016/j.virol.2011.11.029

Brant, A. C., Majerciak, V., Moreira, M. A. M., and Zheng, Z. M. (2019). HPV18 Utilizes Two Alternative Branch Sites for E6*I Splicing to Produce E7 Protein. *Virol. Sin.* 34 (2), 211–221. doi: 10.1007/s12250-019-00098-0

Bzhalava, D., Guan, P., Franceschi, S., Dillner, J., and Clifford, G. (2013). A Systematic Review of the Prevalence of Mucosal and Cutaneous Human Papillomavirus Types. *Virology.* 445 (1-2), 224–231. doi: 10.1016/j.virol.2013.07.015

Cerasuolo, A., Annunziata, C., Tortora, M., Starita, N., Stellato, G., Greggi, S., et al. (2017). Comparative Analysis of HPV16 Gene Expression Profiles in Cervical and in Oropharyngeal Squamous Cell Carcinoma. *Oncotarget* 8 (21), 34070–34081. doi: 10.18632/oncotarget.15977

Cerasuolo, A., Buonaguro, L., Buonaguro, F. M., and Tornesello, M. L. (2020). 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, 474. doi: 10.3389/fcell.2020.00474

Cui, X., Hao, C., Gong, L., Kajitani, N., and Schwartz, S. (2022). HnRNP D Activates Production of HPV16 E1 and E6 mRNAs by Promoting Intron Retention. *Nucleic Acids Res.* 50 (5), 2782–2806. doi: 10.1093/nar/gkac132

De la Rosa-Rios, M. A., Martinez-Salazar, M., Martinez-Garcia, M., Gonzalez-Bonilla, C., and Villegas-Sepulveda, N. (2006). The Intron 1 of HPV 16 has a Suboptimal Branch Point at a Guanosine. *Virus Res.* 118 (1-2), 46–54. doi: 10.1016/j.virusres.2005.11.010

de Sanjose, S., Quint, W. G., Alemany, L., Geraets, D. T., Klaustermeier, J. E., Lloveras, B., et al. (2010). Human Papillomavirus Genotype Attribution in Invasive Cervical Cancer: A Retrospective Cross-Sectional Worldwide Study. *Lancet Oncol.* 11 (11), 1048–1056. doi: 10.1016/S1470-2045(10)70230-8

de Villiers, E. M. (2013). Cross-Roads in the Classification of Papillomaviruses. *Virology* 445 (1-2), 2–10. doi: 10.1016/j.virol.2013.04.023

de Villiers, E. M., Fauquet, C., Broker, T. R., Bernard, H. U., and Zur Hausen, H. (2004). Classification of papillomaviruses. *Virology* 324 (1), 17–27. doi: 10.1016/j.virol.2004.03.033

Doorbar, J. (2005). The Papillomavirus Life Cycle. *J. Clin. Virol.* 32 Suppl 1, S7–15. doi: 10.1016/j.jcv.2004.12.006

Doorbar, J., Parton, A., Hartley, K., Banks, L., Crook, T., Stanley, M., et al. (1990). Detection of Novel Splicing Patterns in a HPV16-Containing Keratinocyte Cell Line. *Virology* 178 (1), 254–262. doi: 10.1016/0042-6822(90)90401-C

Egawa, K. (2003). Do Human Papillomaviruses Target Epidermal Stem Cells? *Dermatology* 207 (3), 251–254. doi: 10.1159/000073085

Elliott, D. J., Oghene, K., Makarov, G., Makarova, O., Hargreave, T. B., Chandley, A. C., et al. (1998). Dynamic Changes in the Subnuclear Organisation of pre-mRNA Splicing Proteins and RBM During Human Germ Cell Development. *J. Cell Sci.* 111 (Pt 9), 1255–1265. doi: 10.1242/jcs.111.9.1255

Filippova, M., Evans, W., Aragon, R., Filippov, V., Williams, V. M., Hong, L., et al. (2014). The Small Splice Variant of HPV16 E6, E6, Reduces Tumor Formation in Cervical Carcinoma Xenografts. *Virology* 450-451, 153–164. doi: 10.1016/j.virol.2013.12.011

Filippova, M., Johnson, M. M., Bautista, M., Filippov, V., Fodor, N., Tungteakkhun, S. S., et al. (2007). The Large and Small Isoforms of Human Papillomavirus Type 16 E6 Bind to and Differentially Affect Procaspase 8 Stability and Activity. *J. Virol.* 81 (8), 4116–4129. doi: 10.1128/JVI.01924-06

Gao, Y., Trivedi, S., Ferris, R. L., and Koide, K. (2014). Regulation of HPV16 E6 and MCL1 by SF3B1 Inhibitor in Head and Neck Cancer Cells. *Sci. Rep.* 4, 6098. doi: 10.1038/srep06098

Garland, S. M., Hernandez-Avila, M., Wheeler, C. M., Perez, G., Harper, D. M., Leodolter, S., et al. (2007). Quadrivalent Vaccine Against Human Papillomavirus to Prevent Anogenital Diseases. *N. Engl. J. Med.* 356 (19), 1928–1943. doi: 10.1056/NEJMoa061760

Gomes, N. P., and Espinosa, J. M. (2010). Gene-Specific Repression of the P53 Target Gene PUMA *via* Intragenic CTCF-Cohesin Binding. *Genes Dev.* 24 (10), 1022–1034. doi: 10.1101/gad.1881010

Goodwin, E. C., and DiMaio, D. (2000). Repression of Human Papillomavirus Oncogenes in HeLa Cervical Carcinoma Cells Causes the Orderly Reactivation of Dormant Tumor Suppressor Pathways. *Proc. Natl. Acad. Sci. U.S.A.* 97 (23), 12513–12518. doi: 10.1073/pnas.97.23.12513

Graham, S. V., and Faizo, A. A. A. (2017). Control of Human Papillomavirus Gene Expression by Alternative Splicing. *Virus Res.* 231, 83–95. doi: 10.1016/j.virusres.2016.11.016

Graveley, B. R., Hertel, K. J., and Maniatis, T. (1999). SR Proteins are 'Locators' of the RNA Splicing Machinery. *Curr. Biol.* 9 (1), R6–R7. doi: 10.1016/S0960-9822(99)80032-3

Groves, I. J., and Coleman, N. (2015). Pathogenesis of Human Papillomavirus-Associated Mucosal Disease. *J. Pathol.* 235 (4), 527–538. doi: 10.1002/path.4496

Guccione, E., Pim, D., and Banks, L. (2004). HPV-18 E6*I Modulates HPV-18 Full-Length E6 Functions in a Cell Cycle Dependent Manner. *Int. J. Cancer.* 110 (6), 928–933. doi: 10.1002/ijc.20184

Guo, Y., Monahan, K., Wu, H., Gertz, J., Varley, K. E., Li, W., et al. (2012). CTCF/Cohesin-Mediated DNA Looping is Required for Protocadherin Alpha Promoter Choice. *Proc. Natl. Acad. Sci. U.S.A.* 109 (51), 21081–21086. doi: 10.1073/pnas.1219280110

Halec, G., Schmitt, M., Dondog, B., Sharkhuu, E., Wentzensen, N., Gheit, T., et al. (2013). Biological Activity of Probable/Possible High-Risk Human Papillomavirus Types in Cervical Cancer. *Int. J. Cancer.* 132 (1), 63–71. doi: 10.1002/ijc.27605

Hallegger, M., Llorian, M., and Smith, C. W. (2010). Alternative Splicing: Global Insights. *FEBS J.* 277 (4), 856–866. doi: 10.1111/j.1742-4658.2009.07521.x

Hanahan, D., and Weinberg, R. A. (2000). The Hallmarks of Cancer. *Cell.* 100 (1), 57–70. doi: 10.1016/S0092-8674(00)81683-9

Han, S. P., Tang, Y. H., and Smith, R. (2010). Functional Diversity of the hnRNPs: Past, Present and Perspectives. *Biochem. J.* 430 (3), 379–392. doi: 10.1042/BJ20100396

Hao, C., Zheng, Y., Jonsson, J., Cui, X., Yu, H., Wu, C., et al. (2022). hnRNP G/RBMX Enhances HPV16 E2 mRNA Splicing Through a Novel Splicing Enhancer and Inhibits Production of Spliced E7 Oncogene mRNAs. *Nucleic Acids Res.* 50 (7), 3867–3891. doi: 10.1093/nar/gkac213

Hertel, K. J., and Maniatis, T. (1999). Serine-Arginine (SR)-Rich Splicing Factors Have an Exon-Independent Function in pre-mRNA Splicing. *Proc. Natl. Acad. Sci. U.S.A.* 96 (6), 2651–2655. doi: 10.1073/pnas.96.6.2651

Howie, H. L., Katzenellenbogen, R. A., and Galloway, D. A. (2009). Papillomavirus E6 Proteins. *Virology* 384 (2), 324–334. doi: 10.1016/j.virol.2008.11.017

Howley, P. M., and Pfister, H. J. (2015). Beta Genus Papillomaviruses and Skin Cancer. *Virology.* 479-480, 290–296. doi: 10.1016/j.virol.2015.02.004

Hummel, M., Hudson, J. B., and Laimins, L. A. (1992). Differentiation-Induced and Constitutive Transcription of Human Papillomavirus Type 31b in Cell Lines Containing Viral Episomes. *J. Virol.* 66 (10), 6070–6080. doi: 10.1128/jvi.66.10.6070-6080.1992

Islam, S., Dasgupta, H., Roychowdhury, A., Bhattacharya, R., Mukherjee, N., Roy, A., et al. (2017). Study of Association and Molecular Analysis of Human Papillomavirus in Breast Cancer of Indian Patients: Clinical and Prognostic Implication. *PloS One* 12 (2), e0172760. doi: 10.1371/journal.pone.0172760

Johansson, C., and Schwartz, S. (2013). Regulation of Human Papillomavirus Gene Expression by Splicing and Polyadenylation. *Nat. Rev. Microbiol.* 11 (4), 239–251. doi: 10.1038/nrmicro2984

Kato, J., Matsushime, H., Hiebert, S. W., Ewen, M. E., and Sherr, C. J. (1993). Direct Binding of Cyclin D to the Retinoblastoma Gene Product (pRb) and pRb Phosphorylation by the Cyclin D-Dependent Kinase CDK4. *Genes Dev.* 7 (3), 331–342. doi: 10.1101/gad.7.3.331

Li, X., and Coffino, P. (1996). High-Risk Human Papillomavirus E6 Protein has Two Distinct Binding Sites Within P53, of Which Only One Determines Degradation. *J. Virol.* 70 (7), 4509–4516. doi: 10.1128/jvi.70.7.4509-4516.1996

Li, N., Franceschi, S., Howell-Jones, R., Snijders, P. J., and Clifford, G. M. (2011). Human Papillomavirus Type Distribution in 30,848 Invasive Cervical Cancers Worldwide: Variation by Geographical Region, Histological Type and Year of Publication. *Int. J. Cancer.* 128 (4), 927–935. doi: 10.1002/ijc.25396

Lim, K. H., Ferraris, L., Filloux, M. E., Raphael, B. J., and Fairbrother, W. G. (2011). Using Positional Distribution to Identify Splicing Elements and Predict pre-mRNA Processing Defects in Human Genes. *Proc. Natl. Acad. Sci. U.S.A.* 108 (27), 11093–11098. doi: 10.1073/pnas.1101135108

Liu, N., and Pan, T. (2016). N6-Methyladenosine-Encoded Epitranscriptomics. *Nat. Struct. Mol. Biol.* 23 (2), 98–102. doi: 10.1038/nsmb.3162

Li, Y., Wang, X., Ni, T., Wang, F., Lu, W., Zhu, J., et al. (2013). Human Papillomavirus Type 58 Genome Variations and RNA Expression in Cervical Lesions. *J. Virol.* 87 (16), 9313–9322. doi: 10.1128/JVI.01154-13

Lorincz, A. T., Reid, R., Jenson, A. B., Greenberg, M. D., Lancaster, W., and Kurman, R. J. (1992). Human Papillomavirus Infection of the Cervix: Relative Risk Associations of 15 Common Anogenital Types. *Obstet Gynecol.* 79 (3), 328–337. doi: 10.1097/00006250-199203000-00002

Manley, J. L., and Tacke, R. (1996). SR Proteins and Splicing Control. *Genes Dev.* 10 (13), 1569–1579. doi: 10.1101/gad.10.13.1569

Manzo-Merino, J., Massimi, P., Lizano, M., and Banks, L. (2014). The Human Papillomavirus (HPV) E6 Oncoproteins Promotes Nuclear Localization of Active Caspase 8. *Virology.* 450-451, 146–152. doi: 10.1016/j.virol.2013.12.013

Marina, R. J., Sturgill, D., Bailly, M. A., Thenoz, M., Varma, G., Prigge, M. F., et al. (2016). TET-Catalyzed Oxidation of Intragenic 5-Methylcytosine Regulates CTCF-Dependent Alternative Splicing. *EMBO J.* 35 (3), 335–355. doi: 10.15252/embj.201593235

Martinez-Zapien, D., Ruiz, F. X., Poirson, J., Mitschler, A., Ramirez, J., Forster, A., et al. (2016). Structure of the E6/E6AP/p53 Complex Required for HPV-Mediated Degradation of P53. *Nature.* 529 (7587), 541–545. doi: 10.1038/nature16481

Matsukura, T., Koi, S., and Sugase, M. (1989). Both Episomal and Integrated Forms of Human Papillomavirus Type 16 are Involved in Invasive Cervical Cancers. *Virology.* 172 (1), 63–72. doi: 10.1016/0042-6822(89)90107-4

McFarlane, M., MacDonald, A. I., Stevenson, A., and Graham, S. V. (2015). Human Papillomavirus 16 Oncoprotein Expression Is Controlled by the Cellular Splicing Factor SRSF2 (Sc35). *J. Virol.* 89 (10), 5276–5287. doi: 10.1128/JVI.03434-14

Mesplede, T., Gagnon, D., Bergeron-Labrecque, F., Azar, I., Senechal, H., Coutlee, F., et al. (2012). P53 Degradation Activity, Expression, and Subcellular Localization of E6 Proteins From 29 Human Papillomavirus Genotypes. *J. Virol.* 86 (1), 94–107. doi: 10.1128/JVI.00751-11

Monahan, K., Rudnick, N. D., Kehayova, P. D., Pauli, F., Newberry, K. M., Myers, R. M., et al. (2012). Role of CCCTC Binding Factor (CTCF) and Cohesin in the Generation of Single-Cell Diversity of Protocadherin-Alpha Gene Expression. *Proc. Natl. Acad. Sci. U S A.* 109 (23), 9125–9130. doi: 10.1073/pnas.1205074109

Munger, K., Basile, J. R., Duensing, S., Eichten, A., Gonzalez, S. L., Grace, M., et al. (2001). Biological Activities and Molecular Targets of the Human Papillomavirus E7 Oncoprotein. *Oncogene.* 20 (54), 7888–7898. doi: 10.1038/sj.onc.1204860

Olmedo-Nieva, L., Munoz-Bello, J. O., Contreras-Paredes, A., and Lizano, M. (2018). The Role of E6 Spliced Isoforms (E6*) in Human Papillomavirus-Induced Carcinogenesis. *Viruses.* 10 (1), 1–20. doi: 10.3390/v10010045

Ozbun, M. A. (2002). Human Papillomavirus Type 31b Infection of Human Keratinocytes and the Onset of Early Transcription. *J. Virol.* 76 (22), 11291–11300. doi: 10.1128/JVI.76.22.11291-11300.2002

Paris, C., Pentland, I., Groves, I., Roberts, D. C., Powis, S. J., Coleman, N., et al. (2015). CCCTC-Binding Factor Recruitment to the Early Region of the Human Papillomavirus 18 Genome Regulates Viral Oncogene Expression. *J. Virol.* 89 (9), 4770–4785. doi: 10.1128/JVI.00097-15

Parkin, D. M., and Bray, F. (2006). Chapter 2: The Burden of HPV-Related Cancers. *Vaccine.* 24 Suppl 3, S3/11–S3/25. doi: 10.1016/j.vaccine.2006.05.111

Patel, D., Huang, S. M., Baglia, L. A., and McCance, D. J. (1999). The E6 Protein of Human Papillomavirus Type 16 Binds to and Inhibits Co-Activation by CBP and P300. *EMBO J.* 18 (5), 5061–5072. doi: 10.1093/emboj/18.18.5061

Pentland, I., Campos-Leon, K., Cotic, M., Davies, K. J., Wood, C. D., Groves, I. J., et al. (2018). Disruption of CTCF-YY1-Dependent Looping of the Human Papillomavirus Genome Activates Differentiation-Induced Viral Oncogene Transcription. *PloS Biol.* 16 (10), e2005752. doi: 10.1371/journal.pbio.2005752

Pim, D., Massimi, P., and Banks, L. (1997). Alternatively Spliced HPV-18 E6* Protein Inhibits E6 Mediated Degradation of P53 and Suppresses Transformed Cell Growth. *Oncogene.* 5 (3), 257–264. doi: 10.1038/sj.onc.1201202

Pyeon, D., Pearce, S. M., Lank, S. M., Ahlquist, P., and Lambert, P. F. (2009). Establishment of Human Papillomavirus Infection Requires Cell Cycle Progression. *PloS Pathog.* 5 (2), e1000318. doi: 10.1371/journal.ppat.1000318

Richardson, H., Kelsall, G., Tellier, P., Voyer, H., Abrahamowicz, M., Ferenczy, A., et al. (2003). The Natural History of Type-Specific Human Papillomavirus Infections in Female University Students. *Cancer Epidemiol. Biomarkers Prev.* 12 (6), 485–490.

Rosenberger, S., De-Castro Arce, J., Langbein, L., Steenbergen, R. D., and Rosl, F. (2010). 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 (15), 7006–7011. doi: 10.1073/pnas.1002620107

Schafer, G., Blumenthal, M. J., and Katz, A. A. (2015). Interaction of Human Tumor Viruses With Host Cell Surface Receptors and Cell Entry. *Viruses.* 7 (5), 2592–2617. doi: 10.3390/v7052592

Scheffner, M., Huibregtse, J. M., Vierstra, R. D., and Howley, P. M. (1993). The HPV-16 E6 and E6-AP Complex Functions as a Ubiquitin-Protein Ligase in the Ubiquitination of P53. *Cell.* 75 (3), 495–505. doi: 10.1016/0092-8674(93)90384-3

Schmitt, M., Dalstein, V., Waterboer, T., Clavel, C., Gissmann, L., and Pawlita, M. (2010). Diagnosing Cervical Cancer and High-Grade Precursors by HPV16 Transcription Patterns. *Cancer Res.* 70 (1), 249–256. doi: 10.1158/0008-5472.CAN-09-2514

Schmitt, M., and Pawlita, M. (2011). The HPV Transcriptome in HPV16 Positive Cell Lines. *Mol. Cell Probes* 25 (2-3), 108–113. doi: 10.1016/j.mcp.2011.03.003

Sedman, S. A., Barbosa, M. S., Vass, W. C., Hubbert, N. L., Haas, J. A., Lowy, D. R., et al. (1991). The Full-Length E6 Protein of Human Papillomavirus Type 16 has Transforming and Trans-Activating Activities and Cooperates With E7 to Immortalize Keratinocytes in Culture. *J. Virol.* 65 (9), 4860–4866. doi: 10.1128/jvi.65.9.4860-4866.1991

Shin, K. H., Kim, R. H., Kang, M. K., Kim, R. H., Kim, S. G., Lim, P. K., et al. (2007). P53 Promotes the Fidelity of DNA End-Joining Activity by, in Part, Enhancing the Expression of Heterogeneous Nuclear Ribonucleoprotein G. *DNA Repair (Amst).* 6 (6), 830–840. doi: 10.1016/j.dnarep.2007.01.013

Shukla, S., Kavak, E., Gregory, M., Imashimizu, M., Shutinoski, B., Kashlev, M., et al. (2011). CTCF-Promoted RNA Polymerase II Pausing Links DNA Methylation to Splicing. *Nature.* 479 (7371), 74–79. doi: 10.1038/nature10442

Snijders, P. J., van den Brule, A. J., Schrijnemakers, H. F., Raaphorst, P. M., Meijer, C. J., and Walboomers, J. M. (1992). Human Papillomavirus Type 33 in a Tonsillar Carcinoma Generates its Putative E7 mRNA *via* Two E6* Transcript Species Which are Terminated at Different Early Region Poly(A) Sites. *J. Virol.* 66 (5), 3172–3178. doi: 10.1128/jvi.66.5.3172-3178.1992

Somberg, M., and Schwartz, S. (2010). 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 (16), 8219–8230. doi: 10.1128/JVI.00462-10

Stanley, M. (2008). Immunobiology of HPV and HPV Vaccines. *Gynecol Oncol.* 109 (2 Suppl), S15–S21. doi: 10.1016/j.ygyno.2008.02.003

Taggart, A. J., DeSimone, A. M., Shih, J. S., Filloux, M. E., and Fairbrother, W. G. (2012). Large-Scale Mapping of Branchpoints in Human pre-mRNA Transcripts *In Vivo. Nat. Struct. Mol. Biol.* 19 (7), 719–721. doi: 10.1038/nsmb.2327

Toots, M., Mannik, A., Kivi, G., Ustav, M.Jr., Ustav, E., and Ustav, M. (2014). The Transcription Map of Human Papillomavirus Type 18 During Genome Replication in U2OS Cells. *PloS One* 9 (12), e116151. doi: 10.1371/journal.pone.0116151

Topisirovic, I., Svitkin, Y. V., Sonenberg, N., and Shatkin, A. J. (2011). Cap and Cap-Binding Proteins in the Control of Gene Expression. *Wiley Interdiscip Rev. RNA.* 2 (2), 277–298. doi: 10.1002/wrna.52

Vazquez-Vega, S., Sanchez-Suarez, L. P., Andrade-Cruz, R., Castellanos-Juarez, E., Contreras-Paredes, A., Lizano-Soberon, M., et al. (2013). Regulation of P14arf Expression by HPV-18 E6 Variants. *J. Med. Virol.* 85 (7), 1215–1221. doi: 10.1002/jmv.23568

Vinokurova, S., Wentzensen, N., Kraus, I., Klaes, R., Driesch, C., Melsheimer, P., et al. (2008). Type-Dependent Integration Frequency of Human Papillomavirus Genomes in Cervical Lesions. *Cancer Res.* 68 (1), 307–313. doi: 10.1158/0008-5472.CAN-07-2754

Wang, Z., Xiao, X., Van Nostrand, E., and Burge, C. B. (2006). General and Specific Functions of Exonic Splicing Silencers in Splicing Control. *Mol. Cell.* 23 (1), 61–70. doi: 10.1016/j.molcel.2006.05.018

Zheng, Z. M., and Baker, C. C. (2006). Papillomavirus Genome Structure, Expression, and Post-Transcriptional Regulation. *Front. Biosci.* 11, 2286–2302. doi: 10.2741/1971

Zheng, Y., Jonsson, J., Hao, C., Shoja Chaghervand, S., Cui, X., Kajitani, N., et al. (2020). 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 (20), e01509–20. doi: 10.1128/JVI.01509-20

Zheng, Z. M., Tao, M., Yamanegi, K., Bodaghi, S., and Xiao, W. (2004). 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 (5), 1091–1108. doi: 10.1016/j.jmb.2004.02.023

Zhou, Z., and Fu, X. D. (2013). Regulation of Splicing by SR Proteins and SR Protein-Specific Kinases. *Chromosoma.* 122 (3), 191–207. doi: 10.1007/s00412-013-0407-z

Zimmermann, H., Koh, C. H., Degenkolbe, R., O'Connor, M. J., Muller, A., Steger, G., et al. (2000). Interaction With CBP/p300 Enables the Bovine Papillomavirus Type 1 E6 Oncoprotein to Downregulate CBP/p300-Mediated Transactivation by P53. *J. Gen. Virol.* 81 (Pt 11), 2617–2623. doi: 10.1099/0022-1317-81-11-2617

zur Hausen, H. (2002). Papillomaviruses and Cancer: From Basic Studies to Clinical Application. *Nat. Rev. Cancer.* 2 (5), 342–350. doi: 10.1038/nrc798

zur Hausen, H. (2009). Papillomaviruses in the Causation of Human Cancers - a Brief Historical Account. *Virology.* 384 (2), 260–265. doi: 10.1016/j.virol.2008.11.046