



Review

# HPV16 and HPV18 Genome Structure, Expression, and Post-Transcriptional Regulation

Lulu Yu , Vladimir Majerciak and Zhi-Ming Zheng \*

Tumor Virus RNA Biology Section, HIV Dynamics and Replication Program, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Frederick, MD 21702, USA; lulu.yu@nih.gov (L.Y.); majerciv@mail.nih.gov (V.M.)

\* Correspondence: zhengt@exchange.nih.gov; Tel.: +1-301-846-7634

**Abstract:** Human papillomaviruses (HPV) are a group of small non-enveloped DNA viruses whose infection causes benign tumors or cancers. HPV16 and HPV18, the two most common high-risk HPVs, are responsible for ~70% of all HPV-related cervical cancers and head and neck cancers. The expression of the HPV genome is highly dependent on cell differentiation and is strictly regulated at the transcriptional and post-transcriptional levels. Both HPV early and late transcripts differentially expressed in the infected cells are intron-containing bicistronic or polycistronic RNAs bearing more than one open reading frame (ORF), because of usage of alternative viral promoters and two alternative viral RNA polyadenylation signals. Papillomaviruses proficiently engage alternative RNA splicing to express individual ORFs from the bicistronic or polycistronic RNA transcripts. In this review, we discuss the genome structures and the updated transcription maps of HPV16 and HPV18, and the latest research advances in understanding RNA cis-elements, intron branch point sequences, and RNA-binding proteins in the regulation of viral RNA processing. Moreover, we briefly discuss the epigenetic modifications, including DNA methylation and possible APOBEC-mediated genome editing in HPV infections and carcinogenesis.



**Citation:** Yu, L.; Majerciak, V.; Zheng, Z.-M. HPV16 and HPV18 Genome Structure, Expression, and Post-Transcriptional Regulation. *Int. J. Mol. Sci.* **2022**, *23*, 4943. <https://doi.org/10.3390/ijms23094943>

Academic Editor: Marta del Pino

Received: 22 March 2022

Accepted: 15 April 2022

Published: 29 April 2022

Corrected: 18 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** papillomaviruses; genome structure; RNA splicing; RNA polyadenylation; epigenetic modification

## 1. Introduction

Human papillomaviruses (HPV) are a group of small, non-enveloped DNA viruses, 52–55 nm in diameter. To date, the group consists of more than 450 genotypes (types) [1–3], in which DNA sequences in the most conserved L1 open reading frame (ORF) differ by more than 10%, even from the closest known HPV type. About 40 HPV types can infect the genital tract and ~80% of the sexually active population will contract a genital HPV infection at some point in their lives. Based on the association with cancer, genital HPV types can be divided into low-risk types, causing mainly genital warts, and high-risk types, which may cause invasive cancer. Almost 100% of cervical cancers, 25–100% of other types of genital cancers, including anus, penis, vagina and vulva, as well as ~40–70% of head and neck cancers, can be attributed to high-risk HPV infection [4,5]. Among the ~15 types of high-risk HPVs that include HPV16, 18, 21, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 73, and 82, HPV16 and HPV18 are the two most common types, accounting for ~70% of all HPV-related cervical cancers [6]. In contrast to low-risk HPVs, high-risk HPVs express two potent oncoproteins, E6 and E7, that mediate, respectively, degradation of cellular p53 and pRb, two tumor suppressor proteins essential for cell cycle control [7,8] and genome stability [9–11], leading to HPV-induced carcinogenesis.

HPV infection is cell-type-specific. HPV infects keratinocytes on the basal lamina through micro-wounding induced by scratching or sexual intercourse. Initial attachment of HPV virions to the host keratinocytes takes place by the interaction of viral capsid proteins with host cell receptors [12]. The virus in the infected keratinocytes replicates in a

differentiation-dependent manner [13,14], with its early genes, including viral oncogenes E6 and E7, being expressed in undifferentiated or intermediately differentiated keratinocytes, while the expression of the late genes occurs only in the keratinocytes undergoing high or terminal differentiation. This highly complex, cell-differentiation-dependent expression of viral genes results from using alternative viral promoters, RNA splicing, and RNA polyadenylation to ensure a proper time and expression level of viral proteins at each viral replication stage. This review will focus on HPV16 and HPV18, the two major oncogenic HPVs, for their genome structures and expression. We will also update the current knowledge of their transcriptomes and the roles of cellular factors in regulating their RNA processing.

## 2. Papillomavirus Genome Structure and Expression

Papillomaviruses contain a double-stranded, circular DNA genome of ~8 kb in size. All papillomavirus genomes can be separated by two polyadenylation (pA) signals, viral early (pA<sub>E</sub>) and viral late (pA<sub>L</sub>), into three distinct regions: an upstream regulatory region (URR), an early (E), and a late (L) gene regions (Figure 1a). The URR region does not encode any protein and is also known as a long control region (LCR). The size of the URR varies among papillomaviruses (Table 1), being 853 bp in HPV16 and 825 bp in HPV18. The HPV URR contains the binding sites for viral E1 and E2 proteins for initiating HPV replication, and several cellular transcription factors, including NFI, Oct-1 (OCT), AP-1, TEF-1 (TF1), and SP1, for transcription initiation [15]. The URR also contains an origin of replication (Ori) which is crucial for papillomavirus replication and is highly homologous to the mammalian autonomous replicating conserved sequences (ACS) [16] (Figure 1b). Each papillomavirus genome contains a single Ori responsible for viral DNA replication.

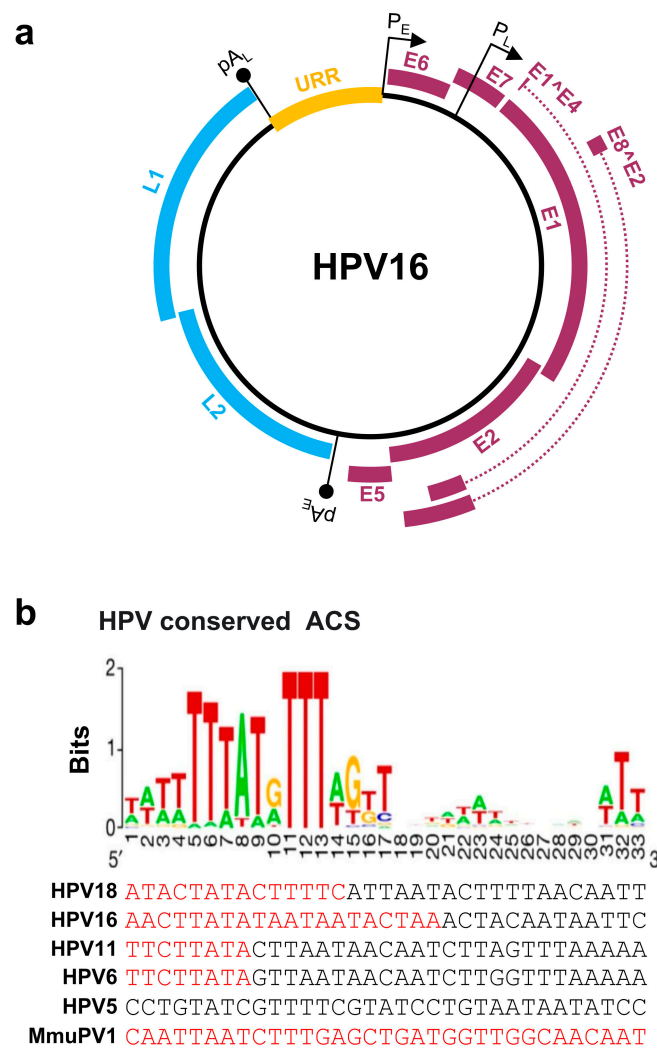
**Table 1.** The genomes and transcription regulatory elements of selected HPVs and animal papillomaviruses.

Species	Genome Size (bp)	Size of URR (bp)	No. of Major Promoters	No. of Splice Sites	No. of pA Sites	No. of CG Boxes	No. A3 Box TTC
HPV18	7857	825	2	11	2	171	73
HPV16	7906	853	2	11	2	112	73
HPV11	7931	756	3	8	2	154	56
HPV6	7996	806	3	7	2	160	77
HPV5	7746	478	4	10?	2	186	96
MmuPV1	7510	609	5	8	2	197	95
CRPV	7871	677	5	8	2	232	98
BPV-1	7946	940	7	11	2	181	89

MmuPV1—Mus musculus papillomavirus 1 or mouse papillomavirus type 1, CRPV—cotton tail rabbit papillomavirus, BPV-1—bovine papillomavirus type 1, URR—upstream regulatory region, pA—polyadenylation signal, and A3—APOBEC3. A3-binding sites were counted based on the studies [17,18]. All the reference genomes analyzed were downloaded from the Papillomavirus Episteme website (PaVE) (<https://pave.niaid.nih.gov/>, accessed on 1 March 2022).

The HPV early region encodes seven viral non-structural proteins: E1, E2, E1^E4, E5, E6, E7, and E8^E2 (Figure 1a). E1 is a viral Ori-binding DNA helicase with E2 being an E1 accessory protein. Thus, both E1 and E2 are essential for papillomavirus DNA replication. E2 also acts as a viral transcriptional activator or repressor, depending on where it binds, and participates in virus genome segregation during mitosis [19,20]. E5, E6, and E7 are viral oncogenes inducing cell immortalization and transformation. Although the HPV E5 function in cell immortalization or transformation is not well defined, viral E6 and E7 target several oncogenic and tumor suppressor proteins [8,21] and miRNAs [22] to abrogate the cell cycle checkpoints. Specifically, E6 interacts with cellular tumor suppressor p53 and E7 with pRb [23]. E8^E2 protein inhibits viral gene expression and viral genome replication [24]. Interestingly, E1^E4 encoded from the viral early region is a viral non-structural late protein expressed in the late stage of infection to facilitate the virus egress by disrupting the host cytokeratin network [25–27]. The late region of the HPV genome encodes two viral capsid

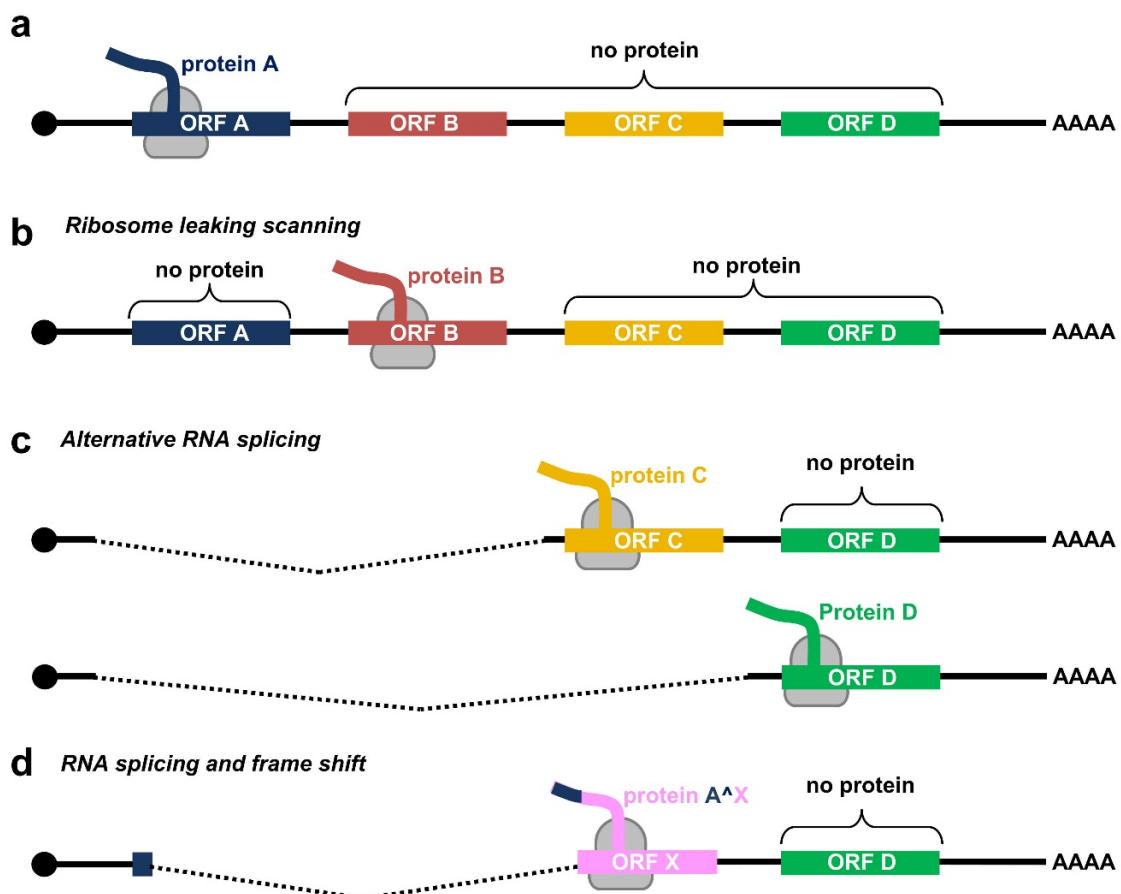
proteins: L1 and L2 (Figure 1a). Expression of L1 and L2 at the late stage of infection is restricted to highly, or terminally, differentiated epithelial cells and is triggered by viral vegetative DNA replication [28]. As all ORFs are located on the same strand of the double-stranded viral genome, papillomavirus transcription is unidirectional, occurring only from one strand of the viral genome. There is no antisense transcription during virus infection even though some background (noise) antisense transcripts may be detectable [29,30]. All viral promoters, named by their transcription start site, are in the URR and early regions of the virus genome. The number of the identified promoters varies among papillomaviruses (Table 1). Papillomaviruses do not encode their own RNA polymerase and utilize host RNA polymerase II [31] for their transcription. Thus, all viral promoter activities are subject to regulation by host and viral transcription factors and chromatin modifications. The viral transcripts derived from each promoter are then polyadenylated, either at an early pA<sub>E</sub> or a late pA<sub>L</sub> site, using host polyadenylation machinery [32,33]. HPV early promoters and pA<sub>E</sub> are active during the early stage of the viral life cycle, whereas viral late promoters and pA<sub>L</sub> become active during the late stage of the viral life cycle. E1^E4 transcripts are initiated from the late promoter P<sub>L</sub> but are normally polyadenylated at the pA<sub>E</sub> site to encode the E1^E4 protein.



**Figure 1.** Papillomavirus genome structure, annotated ORFs, and conserved ACS in the origin of replication (Ori). (a) HPV16 genome and annotated ORFs. The non-structural proteins from the early region of the genome, including E1, E2, E1^E4, E5, E6, E7, and E8^E2, are shown in purple. The viral capsid proteins, L1 and L2, from the late region of the genome are shown in blue. URR (yellow),

upstream regulatory region.  $P_E$  and  $P_L$  mark the early and late promoters, and  $pA_E$  and  $pA_L$  stand for the early and late polyadenylation sites. (b) Conservation of the mammalian ARS consensus sequence (ACS) [16] in the Ori of selected papillomavirus genomes. Shown in this panel are conserved bases with position weight matrix (sequence logo bits) of the 167 predicted ACS elements [16] in comparison with the viral Ori sequences from the selected viral URR tail (red)-head (black) regions. Adapted with permission from Ref. [16]. Copyright 2018 Springer Nature. ARS—autonomously replicating sequence.

Due to a relatively small genome and the compacted genomic organization, papillomaviruses transcribe their RNAs as bicistronic or polycistronic mRNAs with a 5' m<sup>7</sup>G cap to encode multiple viral proteins. However, based on the current dogma of 5' cap-dependent eukaryotic translation starting from the RNA 5' to 3' direction and in the absence of an internal ribosome entry site (IRES) in papillomaviral RNA transcripts, only the first ORF located at the transcript's 5' end is efficiently translatable (Figure 2a), whereas the downstream ORFs not translatable become a part of the transcript's 3' untranslated region (UTR). In some cases, a leaky scanning 40S ribosome might skip the first ORF AUG with weak Kozak content [34,35] and thus allow recognition of the downstream ORF AUG (Figure 2b). All the first three ORFs (papillomaviral E6, E7, and E1 ORFs) of papillomavirus polycistronic transcripts have good Kozak content; therefore, a leaky scanning model might not be a plausible mechanism for viral protein translation from papillomavirus polycistronic RNAs.



**Figure 2.** Translation modes of HPV proteins from polycistronic viral transcripts. (a) Diagram of a polycistronic mRNA harboring multiple ORFs (ORF A–D), of which only the first ORF A is translated by eukaryotic ribosomes (in grey) into a functional protein. (b–d) Strategies utilized by papillomaviruses to express viral proteins from polycistronic transcripts by possible ribosome leaking scanning (b), alternative splicing (c), or usage of alternative frames created by RNA splicing (d).

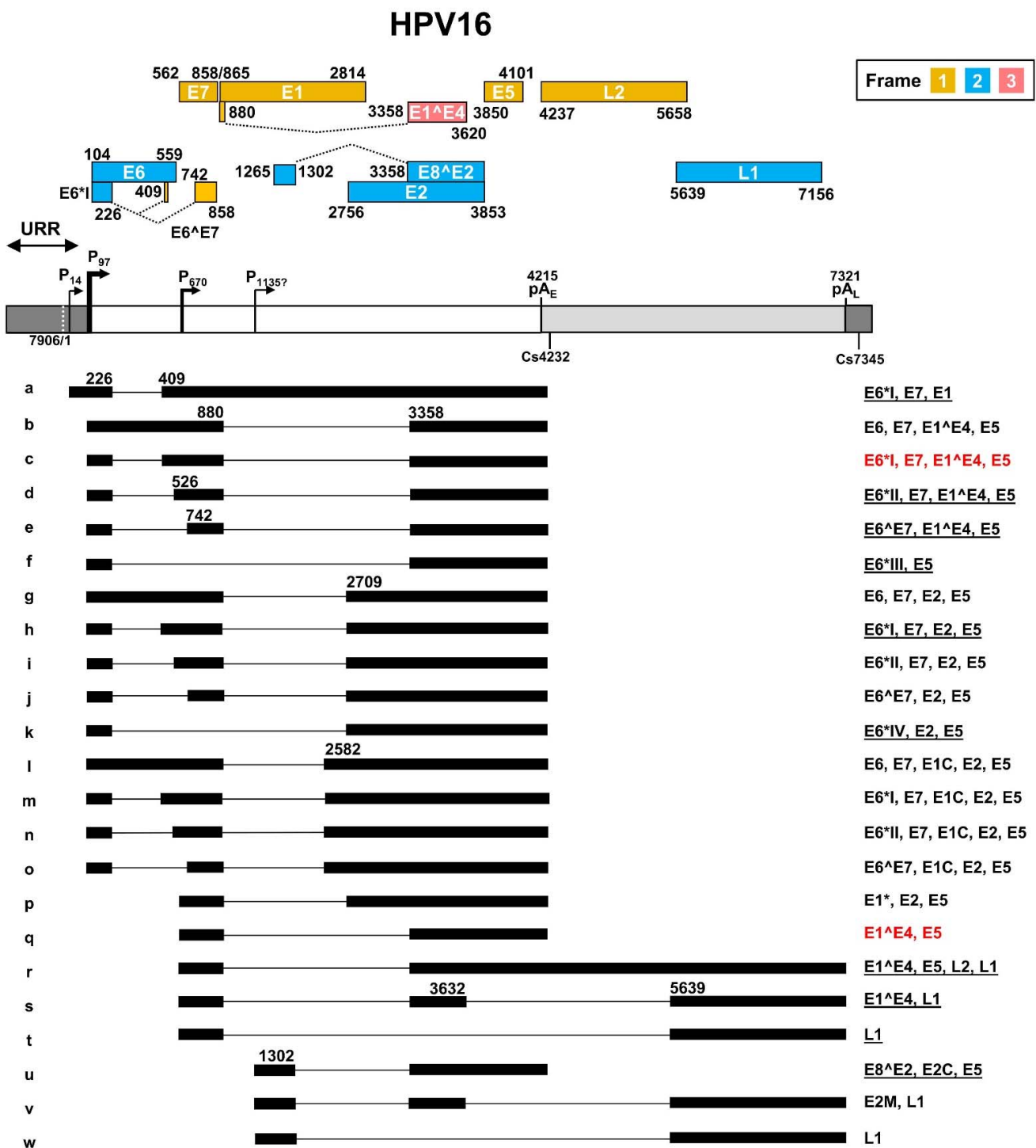
Evolutionarily, papillomaviruses have developed the use of alternative promoters to transcribe viral RNAs with an alternative first ORF. An excellent example of this strategy is using alternative promoters to express E6 RNA from one promoter and E7 RNA from another promoter by animal papillomaviruses [29] and low-risk HPVs [36,37]. However, high-risk HPVs have developed several alternative RNA processing mechanisms to express their proteins from a single polycistronic RNA transcribed from a single promoter (Figure 2c). First, high-risk HPVs utilize alternative RNA splicing for the RNA transcripts derived from a single early promoter to express E6 and E7. Second, utilization of alternative RNA polyadenylation sites, in combination with alternative RNA splicing, allows the expression of L1 and L2 from a late promoter located in the early region. Third, use of alternative translational frames generated by alternative RNA splicing facilitates the expression of viral E1^E4 or E8^E2 from the RNA transcripts with the coding regions partially overlapping with E2 or E1 ORF, respectively (Figure 2d). As a result, the HPV transcriptome consists of a large number of alternatively processed RNAs. This highly complex regulation of viral RNA processing ensures the desired level of individual viral proteins at a proper time for a complete productive cycle of virus infection. A dysregulation of these RNA processing steps is a hallmark of HPV-induced carcinogenesis. To date, we have comprehensively mapped the transcriptome of only a handful of human and animal papillomaviruses.

### 3. Genome Expression and Transcription Maps of HPV16 and HPV18

As more than 90% of HPV16 and HPV18 genomes encode ORFs, alternative RNA splicing of intron-containing bicistronic or polycistronic RNAs bearing more than one ORF to generate mRNA transcripts is essential for the favorable expression of a downstream ORF. Similar to all other papillomaviruses, all introns in HPV16 and HPV18 bicistronic or polycistronic transcripts have been characterized as major U2-type GU-AG, not minor U12-type AU-AC, introns [38–40].

#### 3.1. HPV16 Genome Expression and Transcription Map

As a leading causative agent for cervical cancer, HPV16 genome transcription and viral oncogene E6 and E7 expression were the major focus of HPV studies in the past ~40 years. The HPV16 transcription map was initially generated by multiple laboratories based on viral RNAs extracted from cervical cancer tissues, HPV DNA-transfected cells, and HPV-infected tumor cell lines containing both extrachromosomal and integrated HPV16 genomes [41–43]. The full HPV16 transcription map illustrating sixteen transcripts, currently at the Papillomavirus Episteme website (PaVE) (<https://pave.niaid.nih.gov/>, accessed on 1 March 2022), is an old, outdated version from the Human Papillomaviruses Compendium-1996 composed by Carl Baker and Charles Calef (<https://pave.niaid.nih.gov/lanl-archives/HPVcompintro4.html>, accessed on 1 March 2022). This old version was then updated in 2006 [44] to include an additional four viral transcripts. In this review, new features discovered in the past 16 years since 2006 are collected in an updated transcription map which includes a total of twenty-three viral transcripts and two newly identified promoters P14 and the putative P1135 (Figure 3).



**Figure 3.** The transcription map of HPV16 updated from Ref. [44]. The HPV16 genome in a linear form diagrammed as a bracket line consists of the upstream regulatory region (URR, dark grey box), early region (white box) and late region (light grey box) with nucleotide positions of viral promoters (P), early (pA<sub>E</sub>) and late (pA<sub>L</sub>) polyadenylation signals, and polyadenylation cleavage sites (Cs). The diagram is not to scale. The predicted viral ORFs with nucleotide positions in the viral genome and the corresponding frame used for their translation are shown above. Below the linearized genome are the mapped viral transcripts with identified exons (thick lines) and introns (thin lines). The numbers represent the nucleotide positions of splicing donor sites (5' splice site) and splicing acceptor sites (3' splice site). On the right are coding potentials for each transcript, with the most abundant (>60%) transcripts labeled in red color, and less abundant, but detectable ones (<5%), being underlined.



In general, HPV16 genome expression is driven mainly by two major promoters. The early promoter P97 upstream of the E6 ORF contains a TATA box (a eukaryotic core promoter motif) 35-nt upstream of its transcription start site and is used for transcription of nearly all viral early gene transcripts, while the late promoter P670 within the E7 ORF is a TATA-less promoter responsible for the expression of HPV late genes. The HPV16 E2 protein regulates the promoter P97 activity through binding to the cis-elements in the URR region [45–47]. HPV16 P97-derived early pre-mRNA transcripts undergo alternative RNA splicing to generate various isoforms of mRNAs (b–o) to produce viral early proteins. Eight major RNA splicing sites, including two 5' splice donor sites (5' ss or SD) at nt 226 (SD226) and nt 880 (SD880), and six 3' splice acceptor sites (3' ss or SA) at nt 409 (SA409), nt 526 (SA526), nt 742 (SA742), nt 2582 (SA2582), nt 2709 (SA2709), and nt 3358 (SA3358) are used for alternative RNA splicing to generate viral early mRNA isoforms (Figure 3, RNA species a–o). The viral late RNA transcripts are transcribed from a late promoter P670, highly activated by viral vegetative DNA replication in highly differentiated keratinocytes. These late viral transcripts are also spliced to produce five isoforms of mRNAs (Figure 3, species p–t). The majority of the late pre-mRNAs are spliced from SD880 to SA3358 to produce E1^E4 transcripts (Figure 3, RNA species q), but only a small fraction of these late transcripts is double-spliced from SD880 to SA3358 and then from SD3632 to SA5639 to produce L1 transcripts (Figure 3, RNA species s). Occasionally, the SD880 may be spliced to SA2709 (Figure 3, RNA species p) or directly to SA5639 (Figure 3, RNA species t). The SD880 in the P97 transcripts might also splice back to the SA409 crossover the E7 ORF, producing a circular RNA circE7 with extremely low abundance (<0.4 copies per cell) as a side product of conventional RNA splicing [48]. This scarce amount of circE7 is not thought to have any claimed biological function [48]. Transcripts from a P14 promoter [49] and a putative P1135 promoter [50] are also detectable. The P14 promoter transcript was proposed to encode E1 after splicing from SD226 to SA409 [49] (Figure 3, RNA species a), whereas the P1135 promoter transcript was proposed to encode E8^E2 by splicing from SD1302 to SA3358 [50] (Figure 3, RNA species u). Other unusual RNA splicing patterns described in HPV16 quasivirion-infected keratinocytes [51], and rare viral promoters described in differentiated W12 subclone cells [52], remain to be further verified for possible future inclusion.

The HPV16 genome contains two canonical pA signals (AAUAAA) to guide host RNA polyadenylation machinery to cleave the viral RNA transcripts for the addition of a poly A tail and termination of viral RNA transcription. The pA<sub>E</sub> at nt 4215 is mainly used for polyadenylation of viral early RNA transcripts, and the pA<sub>L</sub> at nt 7321 is responsible for the polyadenylation of viral late RNA transcripts. However, the late RNA transcript E1^E4 uses the pA<sub>E</sub> for its polyadenylation. The cleavage sites for RNA polyadenylation of HPV16 early transcripts have been mapped mainly to nt 4232 [53] and of late transcripts to nt 7345 [52].

HPV16 uses all three coding frames to translate viral proteins. Viral E6, E2, E8^E2, and L1 are expressed from the frame 2 and viral E7, E1, E5, and L2 from the frame 1. Interestingly, the first 5 aa of E1^E4 are translated from the frame 1 and are identical to the first 5 aa of E1 protein, but the rest of the E1^E4 protein is translated from the late RNA exon 2 in the frame 3 generated by viral RNA splicing (Figure 3). The transcript a derived from the P14 promoter has an in-frame AUG start codon at position nt 83, upstream of the annotated E6 start codon at nt 104, but the AUG at nt 83 is weaker than the AUG at nt 104 in the Kozak sequence context (ANNAUGN or GNNNAUGG) [54,55], lacking a purine in position -3 of the AUG and a G in position +4 of the AUG for ribosome binding. Thus, this upstream AUG is unlikely to be useful for translation initiation. As shown in Figure 3, the annotated viral ORFs encoding viral E6, E1, and L2 overlap an intron, and thus, their ORF integrity is disrupted by nuclear RNA splicing in the infected host cells. How these three viral transcripts escape host RNA splicing machinery remains a mystery. In particular, the viral L2 expression requires the late transcript r both to escape the recognition of the early pA<sub>E</sub> signal and the intron 2 splicing.

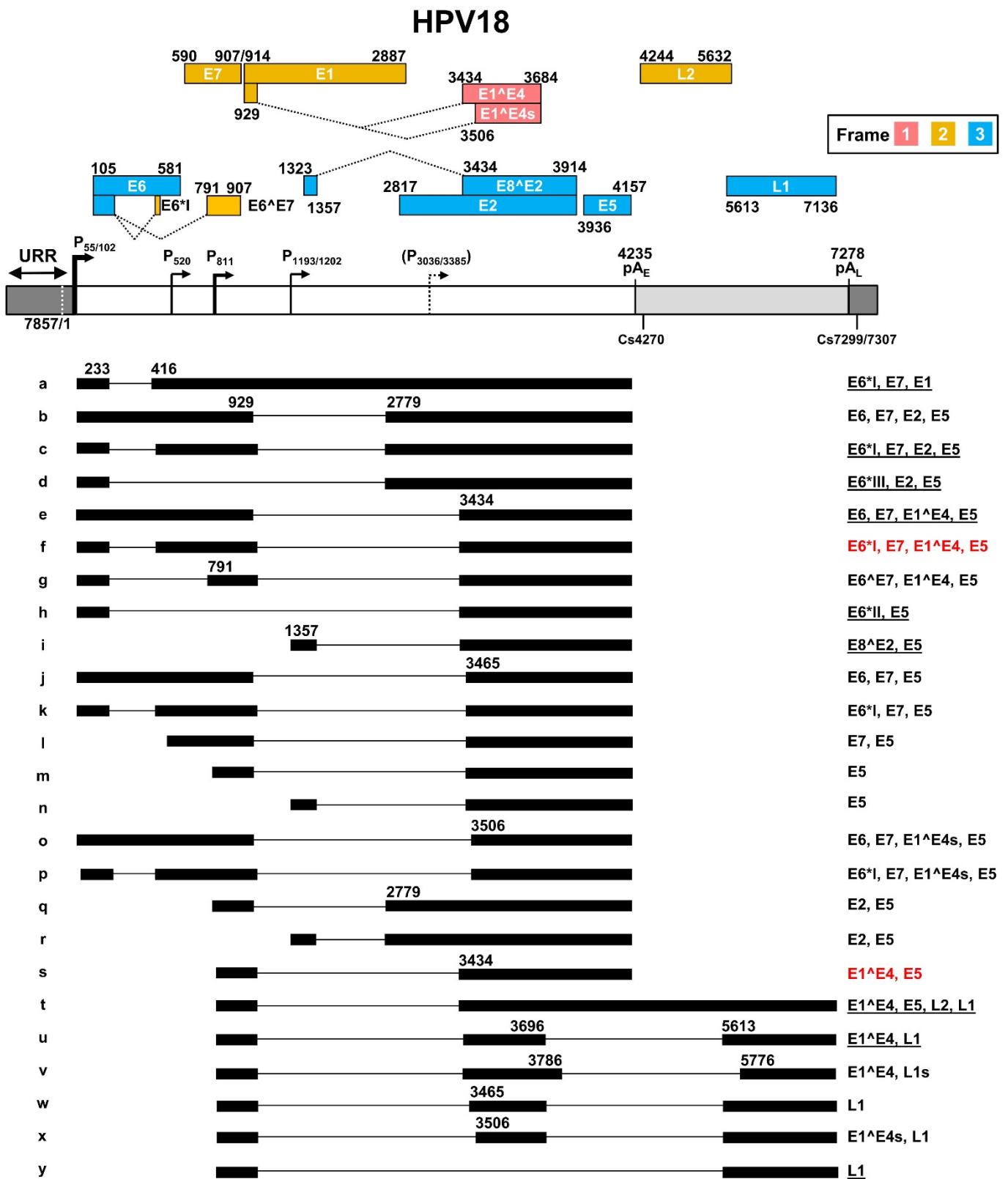
### 3.2. HPV18 Genome Expression and Transcription Map

The most recent updated HPV18 transcription map [56] is based on transcripts generated from productive HPV18 infection in primary human keratinocytes [38] and from U2OS cells transfected with an HPV18 plasmid [57] (Figure 4). In this transcription map, HPV18 shares many features with HPV16 in terms of transcription, alternative RNA splicing, and polyadenylation to express viral early and late genes. HPV18 has two major early promoters P55 and P102 (previously called P105) in the URR region, with a TATA box 27-bp and 25-bp upstream of transcriptional start sites, respectively. As the P55 TATA-box is located in the core Ori region and surrounded by three E2-binding sites, the P55 promoter activity can be inhibited by viral DNA replication, thus leading to a switch to the promoter P102 of which the TATA box lies outside of the Ori core [58]. In contrast to viral early promoters, the HPV18 late promoter lacking a consensus TATA box exhibits heterogeneous transcription start sites ranging from nt 590 to nt 907 in the virus genome. The late promoter P811 within the E7 ORF has been identified as a major HPV18 late promoter [38]. Its activity depends on vegetative viral DNA replication in highly differentiated keratinocytes, and it can be regulated by host factors binding to viral cis-acting DNA elements [28,59,60]. Two minor promoters might exist in the E1 (P1193/1202) [57] and E2 (P3036/3385) [57,61] ORF regions, but remain to be further verified. The P3036/3385 promoter can be activated in the differentiated keratinocytes [60].

HPV18 early transcripts originated from P55 or P102 are polyadenylated at the pA<sub>E</sub> at nt 4235 with a predominant cleavage site at nt 4270 [38]. Similar to HPV16 early transcripts derived from the promoter P97, HPV18 early transcripts derived from the promoter P55/P102 undergo alternative RNA splicing using eight major splice sites, including two 5' donor sites (SD233 and SD929) and six 3' acceptor sites (SA416, SA791, SA2779, SA3434, SA3465, and SA3506), to generate an array of single- or double-spliced mRNAs for expression of various viral early proteins (Figure 4, RNA species a–h, j–k, and o–p). Among these splice acceptor sites, the SA416 and SA3434 are preferably used. The removal of intron one within the E6 ORF through alternative RNA splicing leads to disruption of E6 ORF integrity and promotes expression of E6\*I, E6\*II, E6\*III, E7, and E6^E7 proteins. Similarly, alternative RNA splicing of the intron 2 in these transcripts disrupts the E1 ORF integrity but promotes expression of E2 and E5 in addition to E6 and E7. In this case, only the RNA transcript (Figure 4, RNA species a) that retains the intron 2 by escaping RNA splicing could express viral E1 protein. The promoter P1202 is a minor promoter [38,57] and its RNA transcripts bear only a single intron, which appears to be responsible for the expression of viral E8^E2, E5, and E2, through alternative RNA splicing from SD1357 to SA3434, SA3465, or SA2779 (Figure 4, RNA species i, n, and r).

HPV18 late transcripts initiated from the late promoter P811 can be polyadenylated using an early pA<sub>E</sub> signal at nt 4235 for production of viral late protein E1^E4 (Figure 4, RNA species s), or a late pA<sub>L</sub> signal at nt 7278 for expression of viral capsid proteins L1 (Figure 4, RNA species u, w–y) and L2 (Figure 4, RNA species t). The cleavage sites for polyadenylation of HPV18 L1 and L2 RNAs have been mapped to either nt 7299 or nt 7307 [38]. Thus, HPV18 late RNA transcripts could be both single intron- and double intron-containing RNA transcripts depending on which pA site is selected for RNA polyadenylation and transcription termination. In total, HPV18 late transcripts utilize eight RNA splicing sites, including three splice donor sites SD929, SD3696, and SD3786, and five splice acceptor sites SA3434, SA3465, SA3506, SA5613, and SA5776, to produce nine species of the viral late RNA transcripts (Figure 3, RNA species m, q, s–y), with the RNA species s, u, and y being three predominant late RNA isoforms. Intron 1 in HPV18 late transcripts spans partially the E1 and E2 ORFs and its efficient removal by constitutive RNA splicing during the late stage of virus infection leaves no detectable intron 1-containing late transcripts. However, intron 2 covers the early pA<sub>E</sub> site and almost the entire L2 ORF. Intron 2 retention (Figure 3, RNA species t) by escaping both pA<sub>E</sub> recognition and RNA splicing, which mechanism remains to be understood, is required for the expression of L2 protein.





**Figure 4.** The transcription map of HPV18 updated from Ref. [56]. See the HPV16 transcription map for other description details.

#### 4. Characteristics of Papillomavirus Introns and RNA Cis-Elements in the Regulation of Viral Alternative RNA Splicing

Eukaryotic RNA splicing [62] is carried out in the nucleus by a spliceosomal complex composed of five uridine-rich small nuclear RNAs (snRNA U1, U2, U4, U5, and U6) and ~200 associated proteins [63,64]. All introns are characterized by a 5' splice site (5' ss) interacting with U1, a branch point sequence (BPS) interacting with U2, and a 3' splice site (3' ss) interacting with U2AF [62]. A polypyrimidine tract (PPT), 15–40 nts in size, lies between the BPS and 3' ss and binds PPT-binding proteins, such as PTB. RNA splicing begins with intron recognition in a stepwise interaction, leading to the removal of the recognized intron by two transesterification reactions and joining two exons by the U4/U6. U5 tri-snRNP complex [65,66]. The removed intron in a circular lariat structure is then debranched for degradation [67–69]. When a splice site or BPS is suboptimal, the splicing efficiency of this intron is subject to regulation by many cellular splicing factors, including serine/arginine (SR)-rich protein family members and heterogeneous nuclear ribonucleoproteins (hnRNPs) [70].

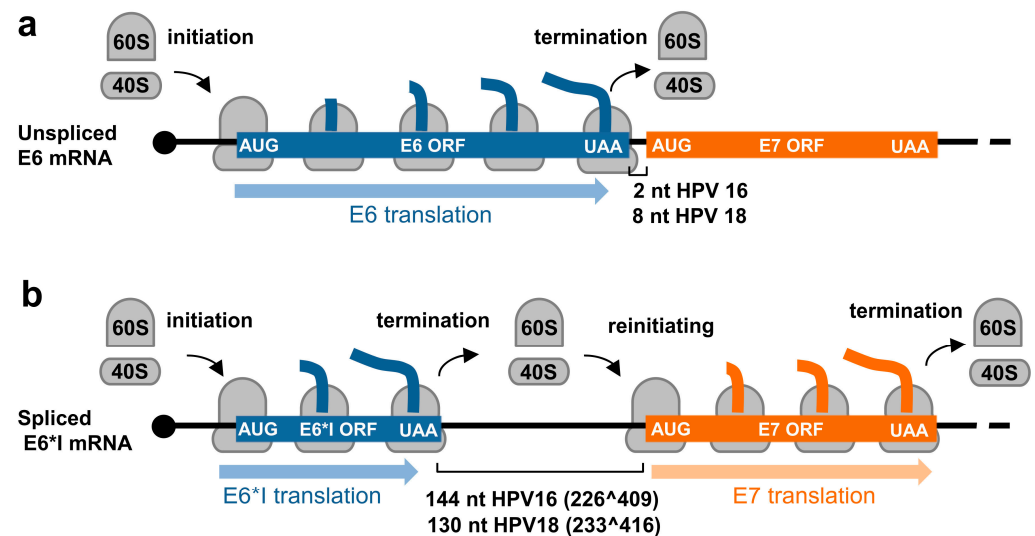
Almost all introns in papillomavirus transcripts are suboptimal [38,71–73]. RNA 5' ss in both HPV16 and HPV18 transcripts miss base-pairing of a few nucleotides to the six core nucleotides of the U1 5' end in the first step of RNA splicing [38,72]. Their 3' ss are also weak, with multiple purines in the PPT, and are accompanied by a non-consensus BPS [74,75], and are therefore prone to an alternative selection of different 3' ss. To understand how this alternative 3' ss selection could occur in the course of papillomavirus infection, early studies using bovine papillomavirus type 1 (BPV-1) late pre-mRNAs, which contain a late leader 5' ss and two alternative 3' ss, a proximal 3' ss at nt 3225 and a distal 3' ss at nt 3605, showed that the alternative selection of the two suboptimal 3' ss is regulated both by exonic splicing enhancers (ESE) and an exonic splicing suppressor or silencer (ESS) [76,77]. Selection of the proximal 3' ss is controlled by an ESE-ESS bipartite regulator, SE1-ESS1 and SE2, upstream of the distal 3' ss for expression of L2 [71,73,76], but usage of the distal 3' ss, also with a suboptimal BPS, is under control by another ESE-ESS bipartite regulator SE4-ESS2, mainly confined to highly differentiated keratinocytes [77,78]. Selection of the distal 3' ss splicing triggers RNA splicing of a downstream intron which covers the early pA site and partial L2 ORF, leading to the production of viral L1 mRNA [78]. Subsequent studies demonstrated that these RNA cis-elements (SE1, SE2, SE4, ESS1, and ESS2) function by interacting with host SR proteins and hnRNPs, leading to a viral early-to-late switch of the splice sites [79]. These interactions include AG-rich SE1 and SE2 binding SRSF1 (ASF/SF2, SRp30a), SRSF2 (SC35, SRp30b), SRSF4 (SRp75), and SRSF6 (SRp55) [80], the AC-rich SE4 interacting with SRSF3 (SRp20) [79], and pyrimidine-rich ESS1 interacting with PTB, U2AF, and SR proteins [81]. These groundbreaking discoveries set up a foundation for further explorations of how human papillomavirus transcription and post-transcriptional RNA processing in the infected keratinocytes are regulated for the translation of viral proteins in a cell differentiation manner.

Subsequently, a subset of RNA cis-elements has been discovered in the corresponding and other regions of HPV16 and HPV18 genomes that contribute to the regulation of HPV16 and HPV18 alternative RNA splicing [56,79,82,83]. Similar to BPV-1, these viral RNA cis-elements in HPV16 and HPV18 function by binding various host splicing factors, of which SR proteins bind to ESEs and hnRNPs to ESS in general [70]. In the following paragraphs, we will specifically discuss their roles in selecting individual alternative 3' ss.

##### 4.1. Intron 1 in the E6 ORF and Its Splicing Regulation of HPV16 and HPV18 Early RNA Transcripts

In contrast to animal and human low-risk HPVs, which express E6 and E7 from two separate promoters and the E6 ORF region which has no intron, high-risk HPVs transcribe E6 and E7 from a single major promoter as a polycistronic E6E7 pre-mRNA (Figures 3 and 4). This E6E7 pre-mRNA contains three exons and two introns, with intron 1 being in the E6 ORF and intron 2 overlapping partially with the E1 and E2 ORFs. This feature of each intron leads to the alternative selection of different 3' ss and thereby produces

distinct RNA splicing isoforms. Only the intron 1-retained E6E7 RNA has an intact E6 ORF and thus translates a full-length E6 protein. The limited intercistronic space between the E6 stop codon and the E7 start codon only 2 nt apart in HPV16 and 8 nt in HPV18, prevents the translating ribosomes from translation re-initiation of the downstream E7 ORF after finalizing E6 translation (Figure 5a). In contrast, removal of intron 1 leads to a frameshift to create a premature stop codon, thus increasing the intercistronic space between newly formed E6\*I ORF and E7 ORF to 144 nt for HPV16 and 130 nt for HPV18, of which intercistronic space is sufficient for the translating ribosomes to re-initiate translation of the downstream E7 ORF (Figure 5b). HPV16 E6\*I ORF encodes a labile polypeptide with 43 amino acid residues [74] of which function remains to be investigated.



**Figure 5.** Regulation of E6 and E7 translation from HPV16 and HPV18 polycistronic E6E7 transcripts by alternative RNA splicing. (a) Diagrams showing E6 translation from the unspliced E6E7 RNA and (b) E7 translation from the spliced E6\*I RNA by translation re-initiation.

Intron 1 splicing and production of E6\*I RNA responsible for efficient expression of viral E7 protein was repeatedly observed experimentally [72,84]. Ample evidence supporting these initial observations was provided in 2006 by using small interfering RNAs (siRNAs) specifically targeting the intron 1 in the E6 ORF in cervical cancer cell lines and further confirmed by E6E7 minigene assays [85]. There is no evidence to date whether E7 is translated through ribosome leaky scanning in cervical cancer cells or high-risk HPV immortalized cells, although an *in vitro* study was reported [86].

In addition to its suboptimal nature of splice sites, HPV16 E6E7 intron 1 has a BPS heptamer AACAAAC with single adenosine (underlined) at nt 385 to serve as a branch site [74] during the first step of RNA splicing reactions (Table S1). HPV18 E6E7 intron 1 exhibits two overlapped, identical BPS heptamer AACUAAC, where one heptamer has a branch site adenosine (underlined) at nt 384 and the other at nt 388 [75], for E6\*I splicing (Table S1). In exploring the regulation of E6\*I RNA splicing, recent studies demonstrated that an RNA cis-element ESS functions as a negative RNA splicing regulator. This ESS was first discovered in the HPV18 E7 ORF in 2016 [56] and later in 2020 in the corresponding region of HPV16 E7 ORF [87]. It is positioned between nt 612–639 in the HPV18 genome and between nt 594–604 in the HPV16 genome. It bears a UUAGA core motif that binds hnRNP A1 to inhibit the intron 1 (233^416) splicing of HPV18 E6E7 pre-mRNAs [56] or the intron 1 (226^409) splicing of HPV16 E6E7 pre-mRNAs [87], thus allowing expression of the viral E6 protein. Knockdown of hnRNP A1 expression in the cells increases E6\*I RNA splicing and thereby the production of E7 proteins [56,87] (Table S1). Additionally, EGF signaling via Erk1/2 activation appears to inhibit HPV16 226^409 splicing of E6E7 pre-mRNAs in the HPV16 E6E7-immortalized keratinocytes and HPV16-positive SiHa

but not CaSki cells [88], suggesting a possible cell-type regulation of E6\*1 RNA splicing. However, the downstream protein target(s) of Erk1/2 in the immortalized or SiHa cells remain to be carefully investigated.

#### 4.2. A Common Intron in the E1 ORF of Both Early and Late Transcripts and Its Splicing Regulation in HPV16 and HPV18 Gene Expression

Intron 2 in the HPV16 early transcripts has one 5' ss at nt 880 and three alternative 3' ss, respectively, at nt 3358, 2709, and 2582 in the order of their frequency of usage in RNA splicing, with 880^3358 splicing being the most common splicing event [89] (Figure 3). In HPV18, intron 2 of the early transcripts has one 5' ss at nt 929 and four alternative 3' ss, respectively at nt 3434, 3465, 3506, and 2779, with 929^3434 splicing being the most common splicing event [38,56] (Figure 4). Because of its suboptimal features with an unknown BPS, this intron splicing is subject to extensive regulation by alternative RNA splicing. It is evident that intron 2 removal in the early transcripts by RNA splicing between 880^3358 of HPV16 or 929^3434 of HPV18 interrupts the integrity of both E1 and E2 ORFs and blocks the production of E1 and E2 proteins but creates an E1^E4 ORF in the spliced early transcripts. However, this newly created E1^E4 ORF is presumably not translatable from the early transcripts due to limited intercistronic space (only 7 nts) between the E7 stop codon and the E1^E4 start codon. However, it might serve for encoding the E5 protein in addition to the spliced E5 RNA isoform f in HPV16 (Figure 3) and h in HPV18 (Figure 4). To date, we know little about how the viral early transcripts escape intron 2 splicing to express the E1 protein in HPV-infected cells. However, the RNAs with 880^2709 splicing in HPV16 and 929^2779 splicing in HPV18 have been predicted to be the most likely to encode the viral E2 protein.

In the viral late transcripts, intron 2 of the viral early transcript is referred as intron 1 and is removed by RNA splicing from most late mRNAs transcribed from a late promoter P670 in HPV16 and P811 in HPV18. The viral late transcripts with 880^3358 splicing in HPV16 and 929^3434 splicing in HPV18 are predominant and necessary to encode viral late proteins of E1^E4 and L2, although alternative splicing of 880^5639 in HPV16 and 929^5613 in HPV18 can be detected for L1 production (Figures 3 and 4). In addition, it has been well-documented that intron 1 splicing in the viral late transcripts commonly triggers late RNA intron 2 splicing between 3632^5639 in HPV16 and 3696^5613 in HPV18. The intron 2 splicing deletes the unwanted pA<sub>E</sub> signal and disrupts L2 ORF integrity to produce the L1 protein. Notably, the generation of an L2 mRNA needs to bypass the pA<sub>E</sub> recognition and retain intron 2 to reach the pA<sub>L</sub> for L2 RNA polyadenylation.

As described above, alternative selection of 3225 3' ss and 3605 3' ss in BPV-1 is mediated through five RNA cis-elements (SE1, ESS1, SE2, SE4, and ESS2) by interaction with various RNA splicing factors [71,73,76,77,79–81,90]. Similar splicing regulation of an equivalent 3358 3' ss in HPV16 and 3434 3' ss in HPV18 has been observed and extensively studied. A 65-nt AC-rich splicing enhancer element at nt 3462–3527, promoting the selection of 3358 3' ss usage for 880^3358 splicing in HPV16 [83], was first described in 2005 and further characterized as a BPV-1 SE4-equivalent 29-nt long ESE at nt 3488–3516 region [79], which binds SRp20 (SRSF3), SRp30s including SRSF1 (ASF/SF2 or SRp30a), SRSF2 (SC35 or SRp30b), 9G8 (SRSF7), and SRp30c (SRSF9), SRp55 (SRSF6), SRp75 (SRSF4), YB-1, and hnRNP L [79] (Table S1). Later, an additional 9-nt AG-rich ESE (ACCGAAGAA) at nt 3446–3454 also enhances 880^3358 splicing by competitive binding of SRSF1, hnRNP G, and Tra2B [91,92] (Table S1).

Although the AC-rich ESE binding SRSF3 and AG-rich ESE interacting with SRSF1, hnRNP G, and Tra2B are important for the usage of the 3358 3' ss splicing of the HPV16 early transcripts to express viral E6 and E7 [79,91,93], the binding of SRSF3 to this AC-rich splicing enhancer also inhibits intron 2 3632^5639 splicing of the viral late transcripts to promote RNA polyadenylation at the pA<sub>E</sub> for the viral E1^E4 expression, thus decreasing HPV16 L1 expression [79]. Inhibition of intron 2 3632^5639 splicing of the HPV16 late RNAs by SRSF3 appears mechanistically different from suppression by an hnRNP D and hnRNP

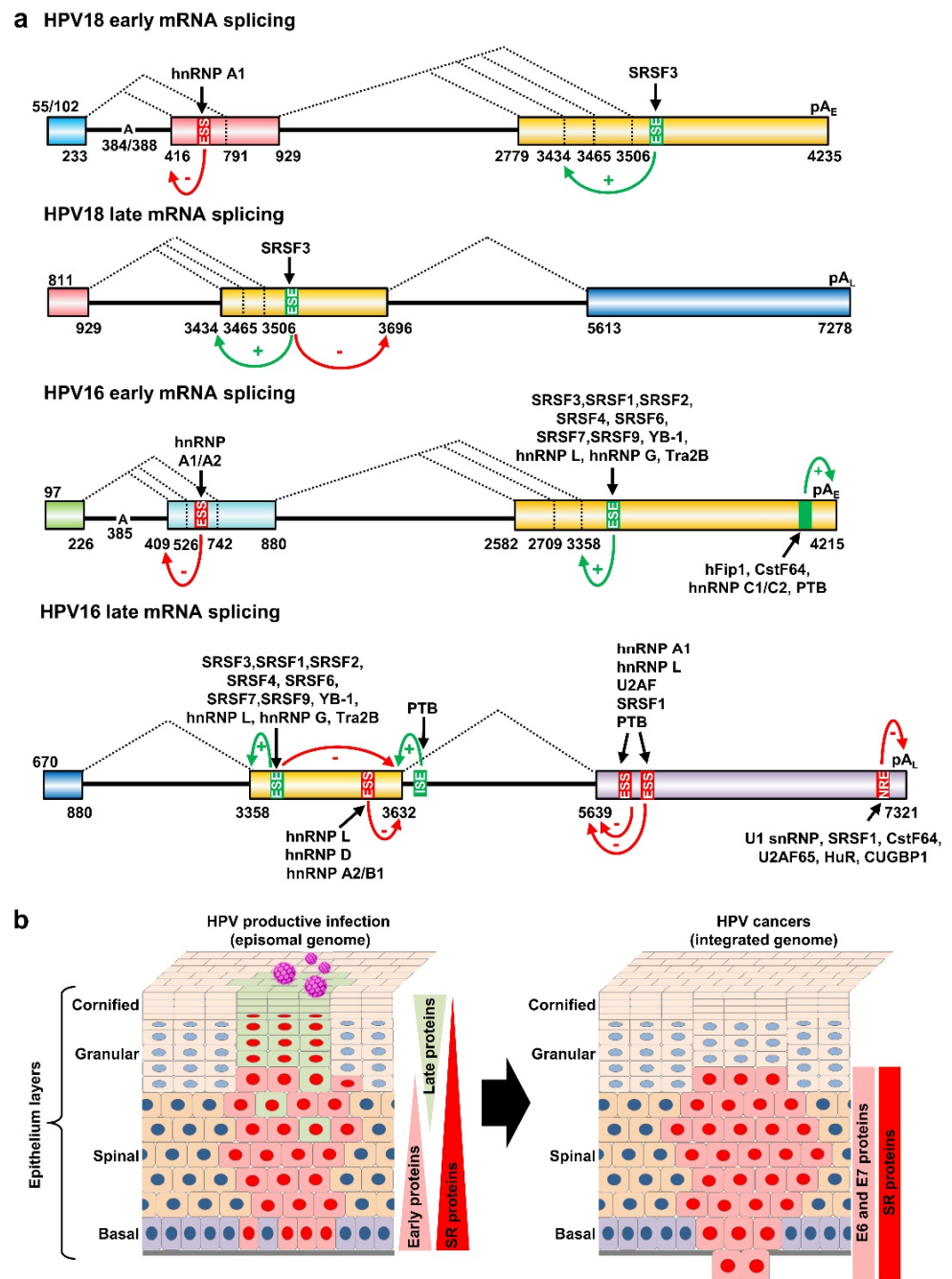
A2/B1-binding inhibitory cis-element (AUAGUA) at nt 3600–3605 upstream of the 3632 5' ss [94] (Table S1). Like the 3358 3' ss in HPV16, the 3434 3' ss in HPV18 is the major 3' ss for both early and late pre-mRNAs [38]. SRSF3 promotes HPV18 929^3434 splicing and E1^E4 production by interaction with the ESE from nt 3520–3550 in the HPV18 genome [56] (Table S1).

#### 4.3. Intron 2 in the L2 ORF and Its Splicing Regulation of HPV16 and HPV18 Late RNA Transcripts

Papillomavirus late RNA transcripts derived from the viral late promoter have an intron 2 spanning over a large part of the L2 ORF. This intron, featuring suboptimal splice sites and an unmapped BPS, needs to be retained to express L2 in the late stage of papillomavirus infection. Otherwise, removal of the intron 2 by RNA splicing of the late transcripts would not only delete the pA<sub>E</sub> signal from the late mRNAs but also disrupt the integrity of the L2 ORF. Intron 2 3632^5639 splicing in the HPV16 genome or 3696^5613 splicing in the HPV18 genome is exclusively used for L1 production even though a few minor splice sites were occasionally detected in the virus-infected cells. The HPV16 late pre-mRNAs transcribed from its late promoter P670 are usually spliced once from nt 880^3358 for E1^E4 and L2 expression or twice from nt 880^3358, and then from nt 3632^5639 before reaching the late pA<sub>L</sub> at nt 7321 for HPV16 L1 expression (Figure 3). Similarly, the HPV18 late pre-mRNAs transcribed from its late promoter P811 are spliced once from nt 929^3434 for E1^E4 and L2 expression, or twice from nt 929^3434, and then from nt 3696^5613 before reaching the late pA<sub>L</sub> at nt 7278 for HPV18 L1 expression (Figure 4). In both HPV16- and HPV18-infected cells, only a small proportion of the viral late transcripts undergo double RNA splicing. Most of the late transcripts are spliced once and polyadenylated at the pA<sub>E</sub> for E1^E4 production, with only a few transcripts capable of reading through the intron 2 5' ss and the pA<sub>E</sub> site for L2 expression (Figures 3 and 4).

Regulation of intron 2 splicing in the HPV16 late transcripts is a complicated process and involves both in RNA cis-elements adjacent to each splice site and host RNA-binding proteins (RBP). As shown in Figure 6a, the expression of HPV16 L1 mRNAs or intron 2 3632^5639 splicing could be inhibited by hnRNP L, hnRNP D, and hnRNP A2/B1 through binding to AUAGUA RNA motifs (nt 3600–3605) located 32 nt upstream of the 3632 5' ss (Figure 6a) [94,95] (Table S1). Interestingly, this suppression of the 3632 5' ss usage could be relieved by transient overexpression of hnRNP C1, or its family member RALYL, which also bind to the same AUAGUA silencer element [96], or by an Akt-kinase-inhibitor-mediated reduction in hnRNP L-binding to this silencer element [95] (Table S1). PTB binds to a 191-nt AU-rich intronic element at a nt 3653–3850 region downstream of the 3632 5' ss, and overexpression of PTB was found to promote the 3632 5' ss splicing and L1 expression [97] (Table S1). Compared to the nt 3632 5' ss, regulation of the nt 5639 3' ss is much simpler. Downstream of this 3' ss reside two ESS elements, a 55-nt ESS1 at nt 5639–5693 [95,98] and a 49-nt ESS2 at nt 5816–5864 in the HPV16 genome [82] (Table S1). Both ESS1 and ESS2 bind hnRNP L, hnRNP A1, PTB, U2AF, and SRSF1 [95] to inhibit the selection of nt 5639 3' ss for L1 expression (Table S1). As all these RBPs are abundant in cells, in vivo binding of the HPV16 late RNAs to specific RBPs and knockout of individual RBPs in the testing cells for the L1 expression and productive HPV infection remain to be further investigated.





**Figure 6.** Regulation of HPV18 and HPV16 RNA splicing and polyadenylation by RNA cis-elements and host RBPs. (a) Diagrams (not to scale) showing the exonic (colored boxes) and intronic (thick lines) regions of HPV18 and HPV16 transcripts. The major alternative splicing events are depicted by dashed lines with the nucleotide positions of splicing sites shown below. Exonic splicing enhancers (ESE), exonic splicing silencers (ESS), intronic splicing enhancers (ISE), and negative regulatory elements (NRE) are shown in green for positive (+) and in red (–) for negative regulation. The black arrows show the host RBPs binding to individual cis-elements. The mapped branch site A in the intron branch point sequence is indicated by its genome position in the viral early transcript. (b) Depiction of HPV-infected epithelium with productive HPV infection (on the left) or with HPV-transformed cells invasive of the basement membrane (on the right). HPV-infected cells expressing viral early proteins are shown in red and the cells expressing HPV late proteins in light green. The changes in the expression of viral proteins and host splicing factors (SR proteins) are on the right of each diagram.

## 5. Papillomavirus RNA Polyadenylation and Its Regulation

The 3' end processing of pre-mRNAs transcribed by RNA polymerase II is essential for RNA stability, export, and translation. RNA polyadenylation initiated by cleavage of the nascent transcript is coupled with the addition of non-templated 150–200 adenylate residues to an RNA transcript to form a poly-A tail. Four core cis-elements in the nascent transcripts, including a UGAU cis-element for a downstream pA signal (PAS) definition, a highly conserved AAUAAA hexamer representing PAS, a cleavage site generally positioned 10–30 nt downstream of PAS, and a G/U- or U-rich element (downstream element, DSE) that is ~10–30 nt further downstream of the cleavage site, are recognized by cellular polyadenylation machinery to initiate the process of RNA cleavage and polyadenylation [99,100]. Numerous cellular factors are involved in this process, including the cleavage and polyadenylation specificity factor (CPSF) that binds to the AAUAAA hexamer, the cleavage stimulation factor (CstF) that binds to the G/U- or U-rich sequence, the cleavage factors I and II (CFI and CFII) that bind to the cleavage site, poly-A polymerase (PAP) for poly A addition, and the poly A-binding protein (PABP) that associates with the nascent poly A tail to facilitate RNA export and translation [101–104].

### 5.1. Regulated RNA Polyadenylation at the Viral pA<sub>E</sub> Site

All papillomaviruses have two polyadenylation signals, pA<sub>E</sub> and pA<sub>L</sub>, for their gene expression. As shown in Figures 3 and 4, HPV16 has a pA<sub>E</sub> at nt 4215 for cleavage and polyadenylation at nt 4232 for all viral early and late E1<sup>E</sup>E4 RNA transcripts and a pA<sub>L</sub> at nt 7321 for cleavage and polyadenylation at nt 7345 for viral late L1 and L2 RNAs. In parallel, HPV18 utilizes a pA<sub>E</sub> at nt 4235 for cleavage and polyadenylation at the nt 4270 cleavage site and a pA<sub>L</sub> at nt 7278 for cleavage and polyadenylation at the nt 7299/7307 site. Recognition of individual viral PAS sites is intimately coupled to viral promoter activation and alternative RNA splicing and thus linked to keratinocyte differentiation, with the pA<sub>E</sub> being recognized in both undifferentiated and differentiated keratinocytes while the pA<sub>L</sub> is recognized only in highly differentiated cells. Although not yet fully understood, additional RNA cis-elements upstream and downstream of the pA<sub>E</sub> are involved in the regulation of viral pA<sub>E</sub> recognition in HPV16 and HPV31 [105,106] (Figure 6a). More specifically, HPV16 applies two RNA cis-elements, a 57-nt U-rich element at nt 4155–4212 upstream of the pA<sub>E</sub> [53] and a 96-nt triple-G-rich element at nt 4383–4479 downstream of the pA<sub>E</sub> [107], to promote the pA<sub>E</sub> activity (Table S1). This enhancement of the pA<sub>E</sub> usage is carried out by the 57-nt U-rich element interacting with hFip1, CstF64, hnRNP C1/C2, and PTB [53] (Figure 6a), and the 96-nt triple G-rich element-binding hnRNP H [107] (Table S1). Interestingly, both the N-terminal and hinge regions of HPV16 E2 contribute to inhibition of pA<sub>E</sub> recognition by the CPSF complex and were shown to cause a read-through at the pA<sub>E</sub> into the late region to induce the expression of L1 and L2 mRNAs in an in vitro reporter study [108].

HPV18 contains six UGUA motifs in the 80-nt sequences upstream of its pA<sub>E</sub> at nt 4235. These UGUA motifs are responsible for CFIm binding and polyadenylation efficiency [109,110]. It would be interesting to know whether the U-rich sequences upstream, and the G-triplets downstream of the HPV18 pA<sub>E</sub>, play similar roles as described for the polyadenylation of the HPV16 early transcripts.

### 5.2. Regulated RNA Polyadenylation at the Viral pA<sub>L</sub> Site

Viral late pA<sub>L</sub> usage is also regulated by RNA cis-element(s) and RBPs. The cis-elements controlling the activity of the viral late pA<sub>L</sub> were initially discovered in BPV-1 late transcripts. These studies revealed a 53-nt BPV-1 inhibitory cis-element at the late 3' UTR, which contains a consensus cryptic 5' ss and inhibits late RNA polyadenylation at the late pA<sub>L</sub>. Through binding U1 snRNAs, this cryptic 5' ss-containing cis-element inhibits late RNA polyadenylation by a direct interaction of U1-70K with poly(A) polymerase [111,112]. The 3' UTR of HPV16 late RNAs also contains a 79-nt negative regulatory element (NRE) spanning from the last few nucleotides of the L1 ORF into the viral late 3' UTR [113]

(Table S1). The HPV16 NRE proximal half contains four weak 5' ss binding U1 snRNP-like complex. Its distal half is GU-rich and binds cellular factors U2AF65, HuR, SRSF1, CstF64 and CUG-binding protein 1 (CUGBP1) [114–119] (Figure 6a) (Table S1). As the NRE is highly conserved in all other papillomaviral L1 RNAs, the functional consequence of U1 snRNP binding to these late 3' UTRs is most likely to protect viral L1 RNAs from premature cleavage and polyadenylation as described [120,121].

Analyses of the 80-nt sequences 5' to the HPV18 pA<sub>L</sub> at nt 7278 showed five UGUA motifs in this region [38]. Similar to the HPV16 late 3' UTR, the HPV18 late 3' UTR contains multiple U1-binding sites (cryptic 5' ss), including one from nt 7217 to 7225 upstream of the pA<sub>L</sub> and four overlapped U1-binding sites (nt 7292–7302, nt 7300–7310, nt 7309–7319, nt 7321–7331) downstream of the late cleavage sites [38]. Collectively, they may suppress RNA polyadenylation of HPV18 late transcripts via their interaction with the factors identified for the HPV16 late NRE [114–119].

## 6. RBPs and Their Dysregulation by HPV Infections

The human genome encodes ~1500 RBPs [122–124]. As stated above, many RBPs involved in RNA splicing, polyadenylation, stability, and translation have been characterized in the regulation of papillomaviral gene expression at the post-transcriptional levels, thus contributing to the control of viral early and late stages of infection. However, these regulations are not unidirectional; the expression levels of RBPs are reciprocally affected by HPV infection.

### 6.1. Host RBPs in HPV Infections

One group of the common RBPs consists of abundant multifunctional hnRNPs [125,126]. Multiple hnRNPs, including hnRNP A1/A2, hnRNP C1/C2, hnRNP H, and hnRNP I (also referred to as polypyrimidine tract-binding protein 1, PTBP1), are significantly upregulated in HPV-induced high-grade lesions and cervical cancer tissues [127–131]. Their contributions to HPV carcinogenesis were assumed, as knockdown of hnRNP A1/A2 [132] and hnRNP I (PTBP1) [133] expression was found to prevent HeLa cell proliferation. In addition, the expression of hnRNP E1 (also referred to as poly C-binding protein 1, PCBP1) is inversely related to the progression from low-grade lesions to cervical cancer [134]. As well as hnRNPs, HPV-positive cervical cancer tissues show the altered expression of many other RBP genes, including increased expression of CDKN2A, ELAVL2, GRB7, HSPB1, KHSRP, and RNASEH2A and decreased expression of NOVA1 [130]. The increased expression of CDKN2A and RNASEH2A, and the decreased expression of NOVA1, are easily detectable from HPV16- and HPV18-infected keratinocytes [130]. RNASEH2A interacts with the proliferation cell nuclear antigen (PCNA) to remove the RNA primers and promote Okazaki fragment maturation during DNA replication [135]. NOVA1 (neuro-oncological ventral antigen 1) is a neuron-specific RNA-binding splicing regulator, which binds to the YCA<sub>Y</sub> elements of the target pre-mRNAs to selectively enhance or suppress exon exclusion [136].

Host RNA splicing factor SR proteins, including SRSF1 (ASF/SF2), SRSF2 (SC35), SRSF3 (SRp20), and SRSF10 (SRp38/SRp40), are also upregulated in HPV-positive cervical precancer lesions and cancer tissues [127,137–139] (Figure 6b). These upregulations could be partially attributed to viral E2's DNA binding and transactivation activities [140] and E6 and E7 protein activities [130]. Among them, SRSF1 (ASF/SF2) and SRSF3 (SRp20) are two proto-oncoprotein splicing factors [138,141,142], but their expressions are mutually dependent by reciprocal regulation of RNA splicing [143]. In cervical basal cells, SRSF1 and SRSF3 are co-expressed along with other splicing factors including SRSF2, hnRNP K, hnRNP L, hnRNP A1, and YB-1 [143]. Consistent with the observation that SRSF3 promotes the E1<sup>+</sup>E4 880<sup>+</sup>3358 splicing in HPV16 and 929<sup>+</sup>3434 splicing in HPV18, but suppresses viral L1 3632<sup>+</sup>5639 splicing in HPV16 and 3696<sup>+</sup>5613 in HPV18 [56,79], the expression level of SRSF3 in keratinocytes in raft cultures supporting L1 expression and virion production is much lower than in the keratinocytes in monolayer cultures which do not normally support L1 expression and production of infectious virions [79]. Importantly, the reduced

expression of SRSF3 also leads to increase keratinocyte differentiation and L1 expression in human foreskin keratinocytes [56] and in papillomavirus-infected tissues [79]. Therefore, differentially expressed SRSF3 controls the HPV early-to-late switch (Figure 6b, left panel).

### 6.2. Viral Proteins Function as RNA-Binding Proteins in the Regulation of RNA Splicing and Polyadenylation

To investigate which viral protein(s) in HPV16 and HPV18 infections causes the increased expression of host RBPs, HPV16 E2, which binds DNA and regulates viral and host gene transcription, has been demonstrated to enhance the expression of SRSF3 (SRp20), SRSF1 (ASF/SF2), SRSF2 (SC35) and SRSF4 (SRp75) and SR proteins kinase 1 (SRPK1) [137,144–146]. HPV16- and HPV18-positive cancers display increased SRSF10 (SRp38/SRrp40) expression which could be upregulated by E6E7 via E2F1 transcriptional activation [139]. As a result, the increased RNA splicing factors in HPV-infected cells not only regulate viral RNA splicing [140], but also RNA splicing of host gene transcripts [147].

In addition to their DNA-binding activities, purified HPV16 E2 and E6 are two viral RNA-binding proteins which suppress *in vitro* RNA splicing via the E2 N-terminal half and hinge regions and the E6 N-terminal and central portions, respectively [148]. Both E2 and E6 preferentially bind to the intron region of pre-mRNAs via the E2 C-terminal DNA-binding domain and E6 nuclear localization signal-3 (NLS-3) in the E6 C-terminal domain [148]. HPV16 E2 interacts with SRp30s, SRp40 and SRp75, whereas HPV16 E6 interacts with SRp30s and SRp75. Both E2 and E6 bind only weakly with SRp55 [148]. Although HPV5 E2 interacts with ASF/SF2, SC35, U1-70K, and U5-100K, through its hinge region containing multiple arginine-serine (RS) repeats and regulates RNA splicing of a reporter RNA [149], it has no effect on RNA splicing *in vitro* [148]. E2 proteins from HPV1, HPV8, HPV11, HPV16, and HPV18 interact with SRPK1 and E2 proteins from HPV1 and HPV8 also strongly bind SRPK2 [150]. SRPK1 and SRPK2 are two SR protein kinases important for SR protein phosphorylation to function [151,152]. Interestingly, HPV1 E1^E4 also interacts with SRPK1 and inhibits phosphorylation of SR proteins and HPV1 E2 [153]. PI3K/Akt3 and protein kinase C (PKC) also phosphorylate SR proteins and regulate papillomavirus RNA splicing and late gene expression [154,155].

HPV16 E2 inhibits polyadenylation *in vitro* by preventing the assembly of the CPSF complex at the pA<sub>E</sub> site, which is reversible by high levels of HPV16 E1 protein or by the polyadenylation factor CPSF30. Both the N-terminal and hinge regions of HPV16 E2 contribute to this inhibition and cause a read-through at the pA<sub>E</sub> into the late region, leading to induction of the expression of L1 and L2 mRNAs with increased L1 RNA 3632^5639 splicing [108].

## 7. DNA Methylation and APOBEC-Mediated Genome Editing in HPV Infections

### 7.1. DNA Methylation in HPV Infections

DNA methylation, catalyzed by the enzymatic activity of DNA methyltransferases (DNMTs), occurs predominantly on cytosines preceding guanine nucleotides (CpG dinucleotides) and maintains CpG methylation patterns in the cells following DNA replication [156]. Methylation of DNA is involved in the regulation of gene expression, and this regulation is context-dependent. In general, methylation in regulatory sequences of genes, such as promoters and enhancers, often leads to transcriptional silencing of genes, whereas transcription-active genes often have hypomethylated or unmethylated regulatory sequences. However, methylation within transcriptional silencers in gene bodies can enhance transcription and RNA splicing [157]. Infection with oncogenic HPVs enhances the expression and enzymatic activity of DNMT1, a major mammalian DNA methylation enzyme, leading to aberrant methylation patterns of both viral and host genomes after DNA replication. As p53 is capable of binding and repressing the DNMT1 promoter in cooperation with SP1 [158], HPV16 E6-mediated p53 degradation results in upregulation of DNMT1 expression [159]. E2F transcription factors activate the DNMT1 promoter [160]. HPV16 E7-mediated free E2Fs from associated pRb enhance DNMT1 expression [161].



HPV16 E7 also promotes DNMT1 activity by binding DNMT1 directly in vitro and in vivo via the E7 zinc-finger CR3 motif [162]. Thus, CpG methylation patterns of host genes in HR-HPV infections are actively investigated as possible specific biomarkers for HPV-related pre-cancer progression [163–165]. In HR-HPV-infected cervical tissues, EPB41L3, EDNRB, LMX1, DPYS, MAL, PAX1, and CADM1 genes were found to be highly methylated in CIN2 and CIN3 (CIN2/3) [166,167]. To date, EPB41L3 [164,168,169] and PAX1 [167,170] have been further used in a few large population studies.

All papillomavirus genomes contain hundreds of CpG dinucleotides, as shown in Table 1, and the correlation between DNA methylation and HPV genome expression has been extensively studied [171]. Methylation of the E2-binding sites (E2BSs) in the URR region reduces E2 binding and deregulates E6 and E7 expression [172,173]. CpG methylation in HPV L1 and L2 genes is associated with HPV pathogenesis and disease stage [174–177]. Methylation of HPV16, HPV18, HPV31, and HPV45 L1 and L2 genes and the cellular DAPK gene appears lowest in asymptomatic infection and increases successively along with cervical lesion progression to cancer [178]. Currently, methylation patterns of the high-risk HPV late regions and one of the selected host genes EPB41L3, PAX1, or DAPK, which is sensitive to high-risk HPV infection, are under active evaluation as possible biomarkers for diagnosis and prognosis of HPV-related pre-cancer progression [163,164,169,170].

### 7.2. Genome Editing and APOBEC3 Expression in HPV Infections

As well as DNA methylation on both host and viral genomes, host RNA-editing by deamination enzymes also plays an important role in genome modifications by introducing point mutations into RNA or single-stranded DNA (ssDNA). Adenosine-to-inosine (A-to-I) editing on double-stranded RNA (dsRNA) is catalyzed by dsRNA-specific adenosine deaminase ADAR1 or ADAR2 [179,180] of which ADAR1 exhibits an increased expression in cervical lesions and cancer tissues [181,182], whereas cytidine-to-uridine (C-to-U) editing on a single-stranded RNA (ssRNA) and single-stranded DNA involves 11 members of the mammalian cytidine deaminase family, the apolipoprotein B messenger RNA-editing enzyme catalytic polypeptide (APOBEC) family [183,184]. Among seven members of APOBEC3 (A3) proteins (A3A, A3B, A3C, A3D, A3F, A3G, and A3H), a correlation between increased HPV DNA editing in the viral Ori-promoter region and expression of A3A, A3C, and A3H was reported in 2008 using a differential DNA denaturation polymerase chain reaction (3D-PCR) to selectively amplify AT-rich edited genome sequences at a lower denaturation temperature 82 °C [185]. Subsequent studies showed that A3A and A3B expression levels and their activities are highly increased in HPV-positive keratinocytes, cervical lesions, and cancer tissues [186–188], most likely due to the expression of viral E6 and E7. As p53 restricts A3B expression [189], viral E6-mediated degradation of p53 leads to the increased expression of A3B. HPV16 E7 stabilizes A3A protein by inhibition of cullin 2-dependent protein degradation [190]. Notably, the increased A3s induce HPV hypermutations throughout the viral genomes of HPV16, 52, and 58, and almost evenly on both sense and antisense strands of each analyzed viral genome [191], which appears predominantly in CIN1, less in CIN2/3 and even below the detection level in cervical cancer tissues [191]. However, other studies show A3-mediated C-to-T substitutions in the E7 [192] and E2 ORF regions [193–195], but more often in the URR region [196].

The increased expression of A3s in HR-HPV cancers has been considered as a source of oncogenic drivers to introduce point mutations in the host genome [186,197–199] and may contribute to intratumor heterogeneity [200]. On the other hand, A3s act as a restriction factor to reduce HPV infection [187,201,202] and promote viral clearance from the infection [203]. In contrast to A3s, A2, a transcriptional repressor [204], displays a decreased expression in CRPV-induced tumor tissues [205]. As A3s facilitate antiviral activities [206,207] and are highly elevated in HPV-associated cancers, the A3 levels in HPV-positive tissues might serve as a prognostic biomarker [195,202,208,209].

As stated above, A3s utilize ssRNA and sometimes ssDNA as a substrate for genome editing. The papillomavirus genome is a double-stranded DNA (dsDNA) and should not



normally be editable by A3s [210,211]. It is speculated that the enriched A3 editing in the HPV16 URR could be a result of the ssDNA stage at the origin of replication and/or transcription from the URR region. A3A- and A3B-mediated mutations are mainly caused by the deamination of the lagging strand template during DNA replication [212,213]. The question of how an A3 editing complex competes in a very brief time with replicating or transcription machinery on a ssDNA for editing without affecting viral DNA replication or transcription remains to be understood. Further hard evidence or careful studies are needed to draw a conclusion about the likelihood that A3 is involved in HPV genome editing. Since the frequency of detected HPV editing is significantly lower than that detected for HIV-1 and HBV to be edited by A3s [196], and highly sensitive techniques, such as 3D-PCR or NGS, had to be used for detection, future studies are required to elucidate how A3s mechanistically affect HPV genome editing during viral DNA replication.

## 8. Remarks and Perspectives

It has been sixteen years since our past review [44] on papillomavirus genome structure, expression, and post-transcriptional regulation. We have made significant progress in dissecting HPV genome structure and function, finding more proteins and noncoding RNAs in the regulation of genome transcription and post-transcriptional process, and enabling keratinocytes' long-term expansion in culture [214]. Since 2006, several HPV vaccines have been implemented by regulatory agencies, and, in 2018, at the 32nd International Papillomavirus Conference in Sydney, WHO called for action toward the elimination of HPV-induced cervical cancer by the year 2050 (<https://www.who.int/reproductivehealth/call-to-action-elimination-cervical-cancer/en/>, accessed on 1 March 2022). For HPV16, these advances include the characterization of a promoter P14 possible for E1 transcription [49] and a potential P1135 promoter for encoding E8'E2 [50]. Additional RNA cis-elements and RBPs required for regulation of RNA splicing and polyadenylation of HPV16 polycistronic RNAs were also identified [215]. The advances in HPV18 are the construction of a full HPV18 transcription map [38], subsequent observation of transcriptional repression and activation by host transcription factors [59,60,216], and the identification of new RNA cis-elements and protein splicing factors to begin our understanding of the regulated HPV18 RNA splicing [56]. The discovery of an ESS element in the E7 ORF region involved in the inhibition of HPV18 E6\*I RNA splicing via binding host hnRNP A1 [56] opens up a new chapter to explore how and when the E6\*I might be produced for translation of viral E7 protein during productive papillomavirus infection and in the course of high-risk HPV-induced oncogenesis. Similarly, this finding [56] led to the identification of a similar ESS element in the HPV16 E7 ORF region involved in the inhibition of HPV16 E6\*I splicing and E7 protein expression [87]. Conceivably, the reduction in hnRNP A1 level, or blocking of hnRNP A1-ESS interaction in the HPV-infected cells, increases E6\*I RNA splicing and E7 protein production, but decreases the expression of full-length E6 mRNAs and thus E6 protein production. To date, the expression level of E6\*I RNA has been found to affect cervical Cancer Response to chemoradiation treatment (CRT) and might be viewed as a predictive biomarker of cervical cancer CRT [217].

The link between alternative RNA splicing in cancer development and progression has been one of the focal areas of research in the papillomavirus field since the 1990s. Alternative RNA splicing of viral polycistronic pre-mRNAs is essential for the expression of individual viral proteins from properly spliced RNA isoforms. Multiple RBPs, especially SR proteins and hnRNPs, have been identified as being involved in the regulation of HPV gene expression by binding to cis-elements in HPV pre-mRNAs. Among the host splicing factors facilitating the regulation of HPV16 and HPV18 RNA splicing, SRSF3 (SRp20) is perhaps the most remarkable host splicing factor, which not only regulates the early-to-late switch of HPV16 and HPV18 productive infection, but also prevents keratinocyte differentiation [56,79]. High SRSF3 expression in basal or undifferentiated keratinocytes prevents keratinocyte differentiation and promotes expression of viral E6 and E7, but not viral L1 (Figure 6b). Raft tissues and terminally differentiated cells in the papilloma and cervical

tissues express very little or no SRSF3, thus supporting viral L1 expression and production of infectious virions [79,218]. Similarly, knockdown of SRSF3 expression in keratinocytes increases L1 production and keratinocyte differentiation [79]. Moreover, SRSF3 regulates the expression of 60 host genes and 182 splicing events in 164 genes [142,143] and functions as a proto-oncogene of which increased expression promotes cell proliferation, transformation, and tumorigenesis [138]. SRSF3 also regulates the expression of at least 20 cellular miRNAs and its expression is mutually co-dependent on SRSF1 [143], another proto-oncogenic host splicing factor [141] also involved in papillomavirus RNA splicing [91,93]. To date, many drugs targeting the splicing factors or oncogenic spliced RNA isoforms are under active development [219–221]. Seeking for splicing-targeted small molecules or chemicals specific to HPV-related cancers would be a future research focus.

While recognizing our past achievements, we remain puzzled by our lack of knowledge of how viral E1, E2, E5, and L2 are produced in the context of the HPV16 or HPV18 genomes, although their functions have been studied extensively by artificial overexpression of each protein from a mammalian expression vector with a foreign promoter. In the context of the HPV16 or HPV18 genomes during HPV infection, each of these genes is expressed along with other genes from a shared early or late viral promoter. Thus, the corresponding ORF of each gene is in the middle or last part of a polycistronic RNA. Considering that RNA 5' cap-mediated ribosome scanning is from the RNA 5' to 3' direction and HPV RNA transcripts lack a ribosomal internal entry site, RNA splicing of the polycistronic RNAs removing or disrupting the upstream ORF(s) has been a plausible mechanism for the expression of E1, E2, E5, and L2 (Figures 3 and 4). However, the E1 ORF and L2 ORF span over a respective intron (Figures 3 and 4), of which splicing will disrupt their ORF integrity and abolish the production of the full-length protein. So far, we have no clue as to how these two introns could escape recognition by cellular splicing machinery, and, if they do, under what circumstances? In the context of the HPV16 and HPV18 genomes, L2 expression is more challenging, as it has to read-through the pA<sub>E</sub> site on the intron 2 retained late transcripts. Whether viral E2 inhibits recognition of the pA<sub>E</sub> site on the late transcripts by host polyadenylation machinery and promotes the read-through into the late region during HPV late infection [108] remains to be carefully explored.

Protein translation is the last step in gene expression, involving 5' cap-mediated ribosome binding to a translation initiation codon AUG of the ORF. The canonical ribosome footprint is ~20–30 nt long from RNA [222–225]. A striking question in the HPV field, which remains to be addressed, is how such a short 5' UTR upstream of the E6 ORF could provide enough space to fit ribosome occupancy and enable translation initiation. Because the majority of the HPV16 viral early transcripts are transcribed from the P97 promoter with a 5' end at nt 97, and the HPV16 E6 translational start codon AUG is at nt 104, this leaves a very short 7-nt 5' UTR for the HPV16 E6 to recruit a ribosome to initiate E6 translation. Similarly, the 5' UTR of the HPV18 E6 ORF, of which the AUG starts at nt 105, is only 3 nt long if the viral early transcripts start from the promoter P102. This puzzle has been solved partially by recent identification of the HPV18 P55 promoter in transcription of the viral early transcripts in raft cultures and HeLa cells [38,58]. These findings indicate that the P55 promoter-derived early transcripts would be useful for translation of viral E6 and other proteins without any ribosome occupancy problem, whereas the P102 promoter-derived early transcripts are most likely responsible for translation of E7 or other downstream ORFs. Conceivably, seeking such a promoter upstream of the P97 promoter in HPV16 genome expression would be an important priority.

Non-coding RNAs (miRNAs and other small RNAs, lncRNAs, circular RNAs, etc.) have revolutionized our understanding of the complexity of the transcriptome involving every part of modern biology [226–228]. Although non-coding RNAs are not part of this focused review and the HPV genome does not encode any authentic viral miRNAs [229], and produces only a scarce amount of circE7 RNA (0.4 copies per CaSki cell) [48], high-risk HPV infections induce aberrant expression of host miRNAs and lncRNAs with oncogenic or tumor-suppressive functions that may contribute to HPV infection-induced

carcinogenesis [22,182,229–237]. The altered expression changes in miRNAs and lncRNAs during HPV infections are induced mainly by viral oncoproteins E6 and E7 [22,229,231,236]. The altered expression of miRNAs and lncRNAs may affect the expression and function of RBPs, and thus HPV RNA processing, as each miRNA may have an average of ~300 conserved targets [226,238], and many lncRNAs function as RNA-binding protein sponges [239,240]. Moreover, the miRNAs and lncRNAs with altered expression may serve as possible biomarkers for diagnosis and prognosis of HR-HPV infection-induced lesion progression [229,234,236,241,242]. Looking forward, the study of non-coding RNAs and HPV interactions in HPV-infected cells holds great promise for further understanding the life cycle and pathogenesis of this important virus.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms23094943/s1>.

**Author Contributions:** L.Y. and Z.-M.Z. contributed to writing, editing, and reviewing this work. V.M. contributed to composing figures and reviewing this work. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported by the Intramural Research Program of the National Institutes of Health, National Cancer Institute, Center for Cancer Research (ZIA5010357 to Z.-M.Z.).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. McBride, A.A. Human papillomaviruses: Diversity, infection and host interactions. *Nat. Rev. Microbiol.* **2022**, *20*, 95–108. [CrossRef] [PubMed]
2. Tirosh, O.; Conlan, S.; Deming, C.; Lee-Lin, S.Q.; Huang, X.; Su, H.C.; Freeman, A.F.; Segre, J.A.; Kong, H.H. Expanded skin virome in DOCK8-deficient patients. *Nat. Med.* **2018**, *24*, 1815–1821. [CrossRef] [PubMed]
3. Pastrana, D.V.; Peretti, A.; Welch, N.L.; Borgogna, C.; Olivero, C.; Badolato, R.; Notarangelo, L.D.; Gariglio, M.; FitzGerald, P.C.; McIntosh, C.E.; et al. Metagenomic Discovery of 83 New Human Papillomavirus Types in Patients with Immunodeficiency. *mSphere* **2018**, *3*, e00645-18. [CrossRef] [PubMed]
4. De Martel, C.; Georges, D.; Bray, F.; Ferlay, J.; Clifford, G.M. Global burden of cancer attributable to infections in 2018: A worldwide incidence analysis. *Lancet Glob. Health* **2020**, *8*, e180–e190. [CrossRef]
5. Sabatini, M.E.; Chiocca, S. Human papillomavirus as a driver of head and neck cancers. *Br. J. Cancer* **2020**, *122*, 306–314. [CrossRef]
6. Muñoz, N.; Bosch, F.X.; de Sanjosé, S.; Herrero, R.; Castellsagué, X.; Shah, K.V.; Snijders, P.J.; Meijer, C.J. Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N. Engl. J. Med.* **2003**, *348*, 518–527. [CrossRef]
7. Münger, K.; Baldwin, A.; Edwards, K.M.; Hayakawa, H.; Nguyen, C.L.; Owens, M.; Grace, M.; Huh, K. Mechanisms of human papillomavirus-induced oncogenesis. *J. Virol.* **2004**, *78*, 11451–11460. [CrossRef]
8. Roman, A.; Munger, K. The papillomavirus E7 proteins. *Virology* **2013**, *445*, 138–168. [CrossRef]
9. Spriggs, C.C.; Laimins, L.A. Human Papillomavirus and the DNA Damage Response: Exploiting Host Repair Pathways for Viral Replication. *Viruses* **2017**, *9*, 232. [CrossRef]
10. Kadaja, M.; Isok-Paas, H.; Laos, T.; Ustav, E.; Ustav, M. Mechanism of genomic instability in cells infected with the high-risk human papillomaviruses. *PLoS Pathog.* **2009**, *5*, e1000397. [CrossRef]
11. Bester, A.C.; Roniger, M.; Oren, Y.S.; Im, M.M.; Sarni, D.; Chaoat, M.; Bensimon, A.; Zamir, G.; Shewach, D.S.; Kerem, B. Nucleotide deficiency promotes genomic instability in early stages of cancer development. *Cell* **2011**, *145*, 435–446. [CrossRef] [PubMed]
12. Ozbun, M.A. Extracellular events impacting human papillomavirus infections: Epithelial wounding to cell signaling involved in virus entry. *Papillomavirus Res.* **2019**, *7*, 188–192. [CrossRef] [PubMed]
13. Longworth, M.S.; Laimins, L.A. Pathogenesis of human papillomaviruses in differentiating epithelia. *Microbiol. Mol. Biol. Rev.* **2004**, *68*, 362–372. [CrossRef] [PubMed]
14. Moody, C. Mechanisms by which HPV Induces a Replication Competent Environment in Differentiating Keratinocytes. *Viruses* **2017**, *9*, 261. [CrossRef]
15. O'Connor, M.; Chan, S.-Y.; Bernard, H.-U. Transcription factor binding sites in the long control region of genital HPVs. In *The Human Papillomaviruses Compendium III-21-40*; Los Alamos National Laboratory: Los Alamos, NM, USA, 1995.
16. Li, N.; Lam, W.H.; Zhai, Y.; Cheng, J.; Cheng, E.; Zhao, Y.; Gao, N.; Tye, B.K. Structure of the origin recognition complex bound to DNA replication origin. *Nature* **2018**, *559*, 217–222. [CrossRef]

17. Bohn, M.F.; Shandilya, S.M.D.; Silvas, T.V.; Nalivaika, E.A.; Kouno, T.; Kelch, B.A.; Ryder, S.P.; Kurt-Yilmaz, N.; Somasundaran, M.; Schiffer, C.A. The ssDNA Mutator APOBEC3A Is Regulated by Cooperative Dimerization. *Structure* **2015**, *23*, 903–911. [[CrossRef](#)]
18. Silvas, T.V.; Hou, S.; Myint, W.; Nalivaika, E.; Somasundaran, M.; Kelch, B.A.; Matsuo, H.; Kurt Yilmaz, N.; Schiffer, C.A. Substrate sequence selectivity of APOBEC3A implicates intra-DNA interactions. *Sci. Rep.* **2018**, *8*, 7511. [[CrossRef](#)]
19. Zur Hausen, H. Papillomaviruses and cancer: From basic studies to clinical application. *Nat. Rev. Cancer* **2002**, *2*, 342–350. [[CrossRef](#)]
20. McBride, A.A.; Oliveira, J.G.; McPhillips, M.G. Partitioning viral genomes in mitosis: Same idea, different targets. *Cell Cycle* **2006**, *5*, 1499–1502. [[CrossRef](#)]
21. Vande Pol, S.B.; Klingelutz, A.J. Papillomavirus E6 oncoproteins. *Virology* **2013**, *445*, 115–137. [[CrossRef](#)]
22. Zheng, Z.M.; Wang, X. Regulation of cellular miRNA expression by human papillomaviruses. *Biochim. Biophys. Acta* **2011**, *1809*, 668–677. [[CrossRef](#)] [[PubMed](#)]
23. Münger, K.; Howley, P.M. Human papillomavirus immortalization and transformation functions. *Virus Res.* **2002**, *89*, 213–228. [[CrossRef](#)]
24. Dreer, M.; van de Poel, S.; Stubenrauch, F. Control of viral replication and transcription by the papillomavirus E8<sup>+</sup>E2 protein. *Virus Res.* **2017**, *231*, 96–102. [[CrossRef](#)] [[PubMed](#)]
25. Doorbar, J.; Ely, S.; Sterling, J.; McLean, C.; Crawford, L. Specific interaction between HPV-16 E1-E4 and cytokeratins results in collapse of the epithelial cell intermediate filament network. *Nature* **1991**, *352*, 824–827. [[CrossRef](#)]
26. Wang, Q.; Kennedy, A.; Das, P.; McIntosh, P.B.; Howell, S.A.; Isaacson, E.R.; Hinz, S.A.; Davy, C.; Doorbar, J. Phosphorylation of the human papillomavirus type 16 E1-E4 protein at T57 by ERK triggers a structural change that enhances keratin binding and protein stability. *J. Virol.* **2009**, *83*, 3668–3683. [[CrossRef](#)] [[PubMed](#)]
27. McIntosh, P.B.; Laskey, P.; Sullivan, K.; Davy, C.; Wang, Q.; Jackson, D.J.; Griffin, H.M.; Doorbar, J. E1-E4-mediated keratin phosphorylation and ubiquitylation: A mechanism for keratin depletion in HPV16-infected epithelium. *J. Cell Sci.* **2010**, *123*, 2810–2822. [[CrossRef](#)]
28. Wang, X.; Liu, H.; Ge, H.; Ajiro, M.; Sharma, N.R.; Meyers, C.; Morozov, P.; Tuschl, T.; Klar, A.; Court, D.; et al. Viral DNA Replication Orientation and hnRNPs Regulate Transcription of the Human Papillomavirus 18 Late Promoter. *mBio* **2017**, *8*, e00713-17. [[CrossRef](#)]
29. Xue, X.Y.; Majerciak, V.; Uberoi, A.; Kim, B.H.; Gotte, D.; Chen, X.; Cam, M.; Lambert, P.F.; Zheng, Z.M. The full transcription map of mouse papillomavirus type 1 (MmuPV1) in mouse wart tissues. *PLoS Pathog.* **2017**, *13*, e1006715. [[CrossRef](#)]
30. Taguchi, A.; Nagasaka, K.; Kawana, K.; Hashimoto, K.; Kusumoto-Matsuo, R.; Plessy, C.; Thomas, M.; Nakamura, H.; Bonetti, A.; Oda, K.; et al. Characterization of novel transcripts of human papillomavirus type 16 using cap analysis gene expression technology. *J. Virol.* **2015**, *89*, 2448–2452. [[CrossRef](#)]
31. Proudfoot, N.J. Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut. *Science* **2016**, *352*, aad9926. [[CrossRef](#)]
32. Manley, J.L. Polyadenylation of mRNA precursors. *Biochim. Biophys. Acta* **1988**, *950*, 1–12. [[CrossRef](#)]
33. Tian, B.; Manley, J.L. Alternative polyadenylation of mRNA precursors. *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 18–30. [[CrossRef](#)]
34. Kozak, M. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **1986**, *44*, 283–292. [[CrossRef](#)]
35. Kozak, M. Initiation of translation in prokaryotes and eukaryotes. *Gene* **1999**, *234*, 187–208. [[CrossRef](#)]
36. Chow, L.T.; Nasser, M.; Wolinsky, S.M.; Broker, T.R. Human papillomavirus types 6 and 11 mRNAs from genital condylomata acuminata. *J. Virol.* **1987**, *61*, 2581–2588. [[CrossRef](#)] [[PubMed](#)]
37. DiLorenzo, T.P.; Steinberg, B.M. Differential regulation of human papillomavirus type 6 and 11 early promoters in cultured cells derived from laryngeal papillomas. *J. Virol.* **1995**, *69*, 6865–6872. [[CrossRef](#)] [[PubMed](#)]
38. Wang, X.; Meyers, C.; Wang, H.K.; Chow, L.T.; Zheng, Z.M. Construction of a full transcription map of human papillomavirus type 18 during productive viral infection. *J. Virol.* **2011**, *85*, 8080–8092. [[CrossRef](#)]
39. Wu, Q.; Krainer, A.R. AT-AC pre-mRNA splicing mechanisms and conservation of minor introns in voltage-gated ion channel genes. *Mol. Cell. Biol.* **1999**, *19*, 3225–3236. [[CrossRef](#)]
40. Turunen, J.J.; Niemelä, E.H.; Verma, B.; Frilander, M.J. The significant other: Splicing by the minor spliceosome. *Wiley Interdiscip. Rev. RNA* **2013**, *4*, 61–76. [[CrossRef](#)]
41. Doorbar, J.; Parton, A.; Hartley, K.; Banks, L.; Crook, T.; Stanley, M.; Crawford, L. Detection of novel splicing patterns in a HPV16-containing keratinocyte cell line. *Virology* **1990**, *178*, 254–262. [[CrossRef](#)]
42. Grassmann, K.; Rapp, B.; Maschek, H.; Petry, K.U.; Iftner, T. 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.* **1996**, *70*, 2339–2349. [[CrossRef](#)] [[PubMed](#)]
43. Rohlfs, M.; Winkenbach, S.; Meyer, S.; Rupp, T.; Dürst, M. Viral transcription in human keratinocyte cell lines immortalized by human papillomavirus type-16. *Virology* **1991**, *183*, 331–342. [[CrossRef](#)]
44. Zheng, Z.M.; Baker, C.C. Papillomavirus genome structure, expression, and post-transcriptional regulation. *Front. Biosci.* **2006**, *11*, 2286–2302. [[CrossRef](#)] [[PubMed](#)]
45. Hawley-Nelson, P.; Androphy, E.J.; Lowy, D.R.; Schiller, J.T. The specific DNA recognition sequence of the bovine papillomavirus E2 protein is an E2-dependent enhancer. *EMBO J.* **1988**, *7*, 525–531. [[CrossRef](#)] [[PubMed](#)]



46. Hou, S.Y.; Wu, S.Y.; Zhou, T.; Thomas, M.C.; Chiang, C.M. Alleviation of human papillomavirus E2-mediated transcriptional repression via formation of a TATA binding protein (or TFIID)-TFIIB-RNA polymerase II-TFIIF preinitiation complex. *Mol. Cell. Biol.* **2000**, *20*, 113–125. [[CrossRef](#)] [[PubMed](#)]
47. Bechtold, V.; Beard, P.; Raj, K. Human papillomavirus type 16 E2 protein has no effect on transcription from episomal viral DNA. *J. Virol.* **2003**, *77*, 2021–2028. [[CrossRef](#)]
48. Yu, L.; Zheng, Z.M. Human Papillomavirus Type 16 Circular RNA Is Barely Detectable for the Claimed Biological Activity. *mBio* **2022**, *13*, e0359421. [[CrossRef](#)]
49. Lace, M.J.; Anson, J.R.; Turek, L.P.; Haugen, T.H. Functional mapping of the human papillomavirus type 16 E1 cistron. *J. Virol.* **2008**, *82*, 10724–10734. [[CrossRef](#)]
50. Straub, E.; Fertey, J.; Dreer, M.; Iftner, T.; Stubenrauch, F. Characterization of the Human Papillomavirus 16 E8 Promoter. *J. Virol.* **2015**, *89*, 7304–7313. [[CrossRef](#)]
51. Bienkowska-Haba, M.; Luszczek, W.; Myers, J.E.; Keiffer, T.R.; DiGiuseppe, S.; Polk, P.; Bodily, J.M.; Scott, R.S.; Sapp, M. A new cell culture model to genetically dissect the complete human papillomavirus life cycle. *PLoS Pathog.* **2018**, *14*, e1006846. [[CrossRef](#)]
52. Milligan, S.G.; Veerapraditsin, T.; Ahamet, B.; Mole, S.; Graham, S.V. Analysis of novel human papillomavirus type 16 late mRNAs in differentiated W12 cervical epithelial cells. *Virology* **2007**, *360*, 172–181. [[CrossRef](#)] [[PubMed](#)]
53. Zhao, X.; Oberg, D.; Rush, M.; Fay, J.; Lambkin, H.; Schwartz, S. A 57-nucleotide upstream early polyadenylation element in human papillomavirus type 16 interacts with hFip1, CstF-64, hnRNP C1/C2, and polypyrimidine tract binding protein. *J. Virol.* **2005**, *79*, 4270–4288. [[CrossRef](#)] [[PubMed](#)]
54. Kozak, M. Effects of intercistronic length on the efficiency of reinitiation by eucaryotic ribosomes. *Mol. Cell. Biol.* **1987**, *7*, 3438–3445. [[CrossRef](#)] [[PubMed](#)]
55. Kozak, M. Pushing the limits of the scanning mechanism for initiation of translation. *Gene* **2002**, *299*, 1–34. [[CrossRef](#)]
56. Ajiro, M.; Tang, S.; Doorbar, J.; Zheng, Z.M. 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.* **2016**, *90*, 9138–9152. [[CrossRef](#)]
57. Toots, M.; Männik, A.; Kivi, G.; Ustav, M., Jr.; Ustav, E.; Ustav, M. The transcription map of human papillomavirus type 18 during genome replication in U2OS cells. *PLoS ONE* **2014**, *9*, e116151. [[CrossRef](#)]
58. Wang, X.; Liu, H.; Wang, H.K.; Meyers, C.; Chow, L.; Zheng, Z.M. HPV18 DNA replication inactivates the early promoter P55 activity and prevents viral E6 expression. *Virol. Sin.* **2016**, *31*, 437–440. [[CrossRef](#)]
59. Pentland, I.; Campos-León, K.; Cotic, M.; Davies, K.J.; Wood, C.D.; Groves, I.J.; Burley, M.; Coleman, N.; Stockton, J.D.; Noyvert, B.; et al. Disruption of CTCF-YY1-dependent looping of the human papillomavirus genome activates differentiation-induced viral oncogene transcription. *PLoS Biol.* **2018**, *16*, e2005752. [[CrossRef](#)]
60. Ferguson, J.; Campos-León, K.; Pentland, I.; Stockton, J.D.; Günther, T.; Beggs, A.D.; Grundhoff, A.; Roberts, S.; Noyvert, B.; Parish, J.L. The chromatin insulator CTCF regulates HPV18 transcript splicing and differentiation-dependent late gene expression. *PLoS Pathog.* **2021**, *17*, e1010032. [[CrossRef](#)]
61. Karlen, S.; Beard, P. Identification and characterization of novel promoters in the genome of human papillomavirus type 18. *J. Virol.* **1993**, *67*, 4296–4306. [[CrossRef](#)]
62. Zheng, Z.M. Regulation of alternative RNA splicing by exon definition and exon sequences in viral and mammalian gene expression. *J. Biomed. Sci.* **2004**, *11*, 278–294. [[CrossRef](#)] [[PubMed](#)]
63. Maniatis, T.; Reed, R. The role of small nuclear ribonucleoprotein particles in pre-mRNA splicing. *Nature* **1987**, *325*, 673–678. [[CrossRef](#)] [[PubMed](#)]
64. Wahl, M.C.; Will, C.L.; Lührmann, R. The spliceosome: Design principles of a dynamic RNP machine. *Cell* **2009**, *136*, 701–718. [[CrossRef](#)] [[PubMed](#)]
65. Wan, R.; Yan, C.; Bai, R.; Wang, L.; Huang, M.; Wong, C.C.; Shi, Y. The 3.8 Å structure of the U4/U6.U5 tri-snRNP: Insights into spliceosome assembly and catalysis. *Science* **2016**, *351*, 466–475. [[CrossRef](#)] [[PubMed](#)]
66. Agafonov, D.E.; Kastner, B.; Dybkov, O.; Hofele, R.V.; Liu, W.T.; Urlaub, H.; Lührmann, R.; Stark, H. Molecular architecture of the human U4/U6.U5 tri-snRNP. *Science* **2016**, *351*, 1416–1420. [[CrossRef](#)] [[PubMed](#)]
67. Ruskin, B.; Green, M.R. An RNA processing activity that debranches RNA lariats. *Science* **1985**, *229*, 135–140. [[CrossRef](#)]
68. Arenas, J.; Hurwitz, J. Purification of a RNA debranching activity from HeLa cells. *J. Biol. Chem.* **1987**, *262*, 4274–4279. [[CrossRef](#)]
69. Chapman, K.B.; Boeke, J.D. Isolation and characterization of the gene encoding yeast debranching enzyme. *Cell* **1991**, *65*, 483–492. [[CrossRef](#)]
70. Busch, A.; Hertel, K.J. Evolution of SR protein and hnRNP splicing regulatory factors. *Wiley Interdiscip. Rev. RNA* **2012**, *3*, 1–12. [[CrossRef](#)]
71. Zheng, Z.M.; He, P.J.; Baker, C.C. Function of a bovine papillomavirus type 1 exonic splicing suppressor requires a suboptimal upstream 3' splice site. *J. Virol.* **1999**, *73*, 29–36. [[CrossRef](#)]
72. Zheng, Z.M.; Tao, M.; Yamanegi, K.; Bodaghi, S.; Xiao, W. 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.* **2004**, *337*, 1091–1108. [[CrossRef](#)]
73. Zheng, Z.M.; Quintero, J.; Reid, E.S.; Gocke, C.; Baker, C.C. Optimization of a weak 3' splice site counteracts the function of a bovine papillomavirus type 1 exonic splicing suppressor in vitro and in vivo. *J. Virol.* **2000**, *74*, 5902–5910. [[CrossRef](#)] [[PubMed](#)]



74. Ajiro, M.; Jia, R.; Zhang, L.; Liu, X.; Zheng, Z.M. Intron definition and a branch site adenosine at nt 385 control RNA splicing of HPV16 E6\*1 and E7 expression. *PLoS ONE* **2012**, *7*, e46412. [[CrossRef](#)] [[PubMed](#)]
75. Brant, A.C.; Majerciak, V.; Moreira, M.A.M.; Zheng, Z.M. HPV18 Utilizes Two Alternative Branch Sites for E6\*1 Splicing to Produce E7 Protein. *Virologica Sinica* **2019**, *34*, 211–221. [[CrossRef](#)] [[PubMed](#)]
76. Zheng, Z.M.; He, P.; Baker, C.C. Selection of the bovine papillomavirus type 1 nucleotide 3225 3' splice site is regulated through an exonic splicing enhancer and its juxtaposed exonic splicing suppressor. *J. Virol.* **1996**, *70*, 4691–4699. [[CrossRef](#)]
77. Zheng, Z.M.; Reid, E.S.; Baker, C.C. Utilization of the bovine papillomavirus type 1 late-stage-specific nucleotide 3605 3' splice site is modulated by a novel exonic bipartite regulator but not by an intronic purine-rich element. *J. Virol.* **2000**, *74*, 10612–10622. [[CrossRef](#)]
78. Barksdale, S.; Baker, C.C. Differentiation-specific alternative splicing of bovine papillomavirus late mRNAs. *J. Virol.* **1995**, *69*, 6553–6556. [[CrossRef](#)]
79. Jia, R.; Liu, X.; Tao, M.; Kruhlak, M.; Guo, M.; Meyers, C.; Baker, C.C.; Zheng, Z.M. Control of the papillomavirus early-to-late switch by differentially expressed SRp20. *J. Virol.* **2009**, *83*, 167–180. [[CrossRef](#)]
80. Zheng, Z.M.; He, P.J.; Baker, C.C. Structural, functional, and protein binding analyses of bovine papillomavirus type 1 exonic splicing enhancers. *J. Virol.* **1997**, *71*, 9096–9107. [[CrossRef](#)]
81. Zheng, Z.M.; Huynen, M.; Baker, C.C. A pyrimidine-rich exonic splicing suppressor binds multiple RNA splicing factors and inhibits spliceosome assembly. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 14088–14093. [[CrossRef](#)]
82. Zhao, X.; Rush, M.; Schwartz, S. Identification of an hnRNP A1-dependent splicing silencer in the human papillomavirus type 16 L1 coding region that prevents premature expression of the late L1 gene. *J. Virol.* **2004**, *78*, 10888–10905. [[CrossRef](#)] [[PubMed](#)]
83. Rush, M.; Zhao, X.; Schwartz, S. 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.* **2005**, *79*, 12002–12015. [[CrossRef](#)] [[PubMed](#)]
84. Sedman, S.A.; Barbosa, M.S.; Vass, W.C.; Hubbert, N.L.; Haas, J.A.; Lowy, D.R.; Schiller, J.T. 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.* **1991**, *65*, 4860–4866. [[CrossRef](#)] [[PubMed](#)]
85. Tang, S.; Tao, M.; McCoy, J.P., Jr.; Zheng, Z.M. The E7 oncoprotein is translated from spliced E6\*1 transcripts in high-risk human papillomavirus type 16- or type 18-positive cervical cancer cell lines via translation reinitiation. *J. Virol.* **2006**, *80*, 4249–4263. [[CrossRef](#)]
86. Stacey, S.N.; Jordan, D.; Williamson, A.J.; Brown, M.; Coote, J.H.; Arrand, J.R. Leaky scanning is the predominant mechanism for translation of human papillomavirus type 16 E7 oncoprotein from E6/E7 bicistronic mRNA. *J. Virol.* **2000**, *74*, 7284–7297. [[CrossRef](#)]
87. Zheng, Y.; Jönsson, J.; Hao, C.; Shoja Chaghervand, S.; Cui, X.; Kajitani, N.; Gong, L.; Wu, C.; Schwartz, S. 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.* **2020**, *94*, e01509-20. [[CrossRef](#)]
88. Rosenberger, S.; De-Castro Arce, J.; Langbein, L.; Steenbergen, R.D.; Rösl, F. Alternative splicing of human papillomavirus type-16 E6/E6\* early mRNA is coupled to EGF signaling via Erk1/2 activation. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 7006–7011. [[CrossRef](#)]
89. Schmitt, M.; Dalstein, V.; Waterboer, T.; Clavel, C.; Gissmann, L.; Pawlita, M. Diagnosing cervical cancer and high-grade precursors by HPV16 transcription patterns. *Cancer Res.* **2010**, *70*, 249–256. [[CrossRef](#)]
90. Jia, R.; Zheng, Z.M. Regulation of bovine papillomavirus type 1 gene expression by RNA processing. *Front. Biosci.* **2009**, *14*, 1270–1282. [[CrossRef](#)]
91. Li, X.; Johansson, C.; Cardoso Palacios, C.; Mossberg, A.; Dhanjal, S.; Bergvall, M.; Schwartz, S. 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* **2013**, *8*, e72776. [[CrossRef](#)]
92. Yu, H.; Gong, L.; Wu, C.; Nilsson, K.; Li-Wang, X.; Schwartz, S. hnRNP G prevents inclusion on the HPV16 L1 mRNAs of the central exon between splice sites SA3358 and SD3632. *J. Gen. Virol.* **2018**, *99*, 328–343. [[CrossRef](#)] [[PubMed](#)]
93. Somberg, M.; Schwartz, S. 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.* **2010**, *84*, 8219–8230. [[CrossRef](#)] [[PubMed](#)]
94. Li, X.; Johansson, C.; Glahder, J.; Mossberg, A.K.; Schwartz, S. 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.* **2013**, *41*, 10488–10508. [[CrossRef](#)] [[PubMed](#)]
95. Kajitani, N.; Glahder, J.; Wu, C.; Yu, H.; Nilsson, K.; Schwartz, S. hnRNP L controls HPV16 RNA polyadenylation and splicing in an Akt kinase-dependent manner. *Nucleic Acids Res.* **2017**, *45*, 9654–9678. [[CrossRef](#)] [[PubMed](#)]
96. Dhanjal, S.; Kajitani, N.; Glahder, J.; Mossberg, A.K.; Johansson, C.; Schwartz, S. 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.* **2015**, *290*, 13354–13371. [[CrossRef](#)] [[PubMed](#)]
97. Somberg, M.; Zhao, X.; Fröhlich, M.; Evander, M.; Schwartz, S. Polypyrimidine tract binding protein induces human papillomavirus type 16 late gene expression by interfering with splicing inhibitory elements at the major late 5' splice site, SD3632. *J. Virol.* **2008**, *82*, 3665–3678. [[CrossRef](#)]

98. Zhao, X.; Fay, J.; Lambkin, H.; Schwartz, S. Identification of a 17-nucleotide splicing enhancer in HPV-16 L1 that counteracts the effect of multiple hnRNP A1-binding splicing silencers. *Virology* **2007**, *369*, 351–363. [[CrossRef](#)] [[PubMed](#)]
99. Gilmartin, G.M. Eukaryotic mRNA 3' processing: A common means to different ends. *Genes Dev.* **2005**, *19*, 2517–2521. [[CrossRef](#)]
100. Jurado, A.R.; Tan, D.; Jiao, X.; Kiledjian, M.; Tong, L. Structure and function of pre-mRNA 5'-end capping quality control and 3'-end processing. *Biochemistry* **2014**, *53*, 1882–1898. [[CrossRef](#)]
101. Chekanova, J.A.; Belostotsky, D.A. Evidence that poly(A) binding protein has an evolutionarily conserved function in facilitating mRNA biogenesis and export. *RNA* **2003**, *9*, 1476–1490. [[CrossRef](#)]
102. Mandel, C.R.; Kaneko, S.; Zhang, H.; Gebauer, D.; Vethantham, V.; Manley, J.L.; Tong, L. Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. *Nature* **2006**, *444*, 953–956. [[CrossRef](#)] [[PubMed](#)]
103. Xiang, K.; Tong, L.; Manley, J.L. Delineating the structural blueprint of the pre-mRNA 3'-end processing machinery. *Mol. Cell. Biol.* **2014**, *34*, 1894–1910. [[CrossRef](#)] [[PubMed](#)]
104. Gruber, A.J.; Zavolan, M. Alternative cleavage and polyadenylation in health and disease. *Nat. Rev. Genet.* **2019**, *20*, 599–614. [[CrossRef](#)] [[PubMed](#)]
105. Terhune, S.S.; Hubert, W.G.; Thomas, J.T.; Laimins, L.A. Early polyadenylation signals of human papillomavirus type 31 negatively regulate capsid gene expression. *J. Virol.* **2001**, *75*, 8147–8157. [[CrossRef](#)]
106. Oberg, D.; Collier, B.; Zhao, X.; Schwartz, S. 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.* **2003**, *77*, 11674–11684. [[CrossRef](#)]
107. Oberg, D.; Fay, J.; Lambkin, H.; Schwartz, S. A downstream polyadenylation element in human papillomavirus type 16 L2 encodes multiple GGG motifs and interacts with hnRNP H. *J. Virol.* **2005**, *79*, 9254–9269. [[CrossRef](#)]
108. Johansson, C.; Somberg, M.; Li, X.; Backström Winquist, E.; Fay, J.; Ryan, F.; Pim, D.; Banks, L.; Schwartz, S. HPV-16 E2 contributes to induction of HPV-16 late gene expression by inhibiting early polyadenylation. *EMBO J.* **2012**, *31*, 3212–3227. [[CrossRef](#)]
109. Yang, Q.; Gilmartin, G.M.; Doublé, S. Structural basis of UGUA recognition by the Nudix protein CFI(m)25 and implications for a regulatory role in mRNA 3' processing. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 10062–10067. [[CrossRef](#)]
110. Yang, Q.; Coseno, M.; Gilmartin, G.M.; Doublé, S. Crystal structure of a human cleavage factor CFI(m)25/CFI(m)68/RNA complex provides an insight into poly(A) site recognition and RNA looping. *Structure* **2011**, *19*, 368–377. [[CrossRef](#)]
111. Furth, P.A.; Choe, W.T.; Rex, J.H.; Byrne, J.C.; Baker, C.C. Sequences homologous to 5' splice sites are required for the inhibitory activity of papillomavirus late 3' untranslated regions. *Mol. Cell. Biol.* **1994**, *14*, 5278–5289. [[CrossRef](#)]
112. Gunderson, S.I.; Polycarpou-Schwarz, M.; Mattaj, I.W. U1 snRNP inhibits pre-mRNA polyadenylation through a direct interaction between U1 70K and poly(A) polymerase. *Mol. Cell* **1998**, *1*, 255–264. [[CrossRef](#)]
113. Kennedy, I.M.; Haddow, J.K.; Clements, J.B. Analysis of human papillomavirus type 16 late mRNA 3' processing signals in vitro and in vivo. *J. Virol.* **1990**, *64*, 1825–1829. [[CrossRef](#)] [[PubMed](#)]
114. Cumming, S.A.; McPhillips, M.G.; Veerapraditsin, T.; Milligan, S.G.; Graham, S.V. 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.* **2003**, *77*, 5167–5177. [[CrossRef](#)] [[PubMed](#)]
115. Dietrich-Goetz, W.; Kennedy, I.M.; Levins, B.; Stanley, M.A.; Clements, J.B. A cellular 65-kDa protein recognizes the negative regulatory element of human papillomavirus late mRNA. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 163–168. [[CrossRef](#)] [[PubMed](#)]
116. Koffa, M.D.; Graham, S.V.; Takagaki, Y.; Manley, J.L.; Clements, J.B. The human papillomavirus type 16 negative regulatory RNA element interacts with three proteins that act at different posttranscriptional levels. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 4677–4682. [[CrossRef](#)] [[PubMed](#)]
117. McPhillips, M.G.; Veerapraditsin, T.; Cumming, S.A.; Karali, D.; Milligan, S.G.; Boner, W.; Morgan, I.M.; Graham, S.V. SF2/ASF binds the human papillomavirus type 16 late RNA control element and is regulated during differentiation of virus-infected epithelial cells. *J. Virol.* **2004**, *78*, 10598–10605. [[CrossRef](#)]
118. Zhao, X.; Rush, M.; Carlsson, A.; Schwartz, S. The presence of inhibitory RNA elements in the late 3'-untranslated region is a conserved property of human papillomaviruses. *Virus Res.* **2007**, *125*, 135–144. [[CrossRef](#)]
119. Goracznik, R.; Gunderson, S.I. The regulatory element in the 3'-untranslated region of human papillomavirus 16 inhibits expression by binding CUG-binding protein 1. *J. Biol. Chem.* **2008**, *283*, 2286–2296. [[CrossRef](#)]
120. Kaida, D.; Berg, M.G.; Younis, I.; Kasim, M.; Singh, L.N.; Wan, L.; Dreyfuss, G. U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature* **2010**, *468*, 664–668. [[CrossRef](#)]
121. Cai, Z.; So, B.R.; Dreyfuss, G. Comprehensive RNP profiling in cells identifies U1 snRNP complexes with cleavage and polyadenylation factors active in telescripting. *Methods Enzymol.* **2021**, *655*, 325–347. [[CrossRef](#)]
122. Gerstberger, S.; Hafner, M.; Tuschl, T. Learning the language of post-transcriptional gene regulation. *Genome Biol.* **2013**, *14*, 130. [[CrossRef](#)] [[PubMed](#)]
123. Castello, A.; Fischer, B.; Hentze, M.W.; Preiss, T. RNA-binding proteins in Mendelian disease. *Trends Genet.* **2013**, *29*, 318–327. [[CrossRef](#)] [[PubMed](#)]
124. Castello, A.; Fischer, B.; Eichelbaum, K.; Horos, R.; Beckmann, B.M.; Strein, C.; Davey, N.E.; Humphreys, D.T.; Preiss, T.; Steinmetz, L.M.; et al. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* **2012**, *149*, 1393–1406. [[CrossRef](#)] [[PubMed](#)]

125. Krecic, A.M.; Swanson, M.S. hnRNP complexes: Composition, structure, and function. *Curr. Opin. Cell Biol.* **1999**, *11*, 363–371. [[CrossRef](#)]
126. Geuens, T.; Bouhy, D.; Timmerman, V. The hnRNP family: Insights into their role in health and disease. *Hum. Genet.* **2016**, *135*, 851–867. [[CrossRef](#)]
127. Fay, J.; Kelehan, P.; Lambkin, H.; Schwartz, S. Increased expression of cellular RNA-binding proteins in HPV-induced neoplasia and cervical cancer. *J. Med. Virol.* **2009**, *81*, 897–907. [[CrossRef](#)]
128. Kim, Y.J.; Kim, B.R.; Ryu, J.S.; Lee, G.O.; Kim, H.R.; Choi, K.H.; Ryu, J.W.; Na, K.S.; Park, M.C.; So, H.S.; et al. HNRNPA1, a Splicing Regulator, Is an Effective Target Protein for Cervical Cancer Detection: Comparison With Conventional Tumor Markers. *Int. J. Gynecol. Cancer* **2017**, *27*, 326–331. [[CrossRef](#)]
129. Qing, S.; Tulake, W.; Ru, M.; Li, X.; Yuemaier, R.; Lidifu, D.; Rouzibilali, A.; Hasimu, A.; Yang, Y.; Rouziahong, R.; et al. Proteomic identification of potential biomarkers for cervical squamous cell carcinoma and human papillomavirus infection. *Tumour Biol.* **2017**, *39*, 1010428317697547. [[CrossRef](#)]
130. Xu, J.; Liu, H.; Yang, Y.; Wang, X.; Liu, P.; Li, Y.; Meyers, C.; Banerjee, N.S.; Wang, H.K.; Cam, M.; et al. Genome-Wide Profiling of Cervical RNA-Binding Proteins Identifies Human Papillomavirus Regulation of RNASEH2A Expression by Viral E7 and E2F1. *mBio* **2019**, *10*, e02687-18. [[CrossRef](#)]
131. Cheunim, T.; Zhang, J.; Milligan, S.G.; McPhillips, M.G.; Graham, S.V. The alternative splicing factor hnRNP A1 is up-regulated during virus-infected epithelial cell differentiation and binds the human papillomavirus type 16 late regulatory element. *Virus Res.* **2008**, *131*, 189–198. [[CrossRef](#)]
132. Patry, C.; Bouchard, L.; Labrecque, P.; Gendron, D.; Lemieux, B.; Toutant, J.; Lapointe, E.; Wellinger, R.; Chabot, B. Small interfering RNA-mediated reduction in heterogeneous nuclear ribonucleoproteins A1/A2 proteins induces apoptosis in human cancer cells but not in normal mortal cell lines. *Cancer Res.* **2003**, *63*, 7679–7688. [[PubMed](#)]
133. Wang, C.; Norton, J.T.; Ghosh, S.; Kim, J.; Fushimi, K.; Wu, J.Y.; Stack, M.S.; Huang, S. Polypyrimidine tract-binding protein (PTB) differentially affects malignancy in a cell line-dependent manner. *J. Biol. Chem.* **2008**, *283*, 20277–20287. [[CrossRef](#)] [[PubMed](#)]
134. Pillai, M.R.; Chacko, P.; Kesari, L.A.; Jayaprakash, P.G.; Jayaram, H.N.; Antony, A.C. Expression of folate receptors and heterogeneous nuclear ribonucleoprotein E1 in women with human papillomavirus mediated transformation of cervical tissue to cancer. *J. Clin. Pathol.* **2003**, *56*, 569–574. [[CrossRef](#)] [[PubMed](#)]
135. Liu, B.; Hu, J.; Wang, J.; Kong, D. Direct Visualization of RNA-DNA Primer Removal from Okazaki Fragments Provides Support for Flap Cleavage and Exonucleolytic Pathways in Eukaryotic Cells. *J. Biol. Chem.* **2017**, *292*, 4777–4788. [[CrossRef](#)]
136. Ule, J.; Stefani, G.; Mele, A.; Ruggiu, M.; Wang, X.; Taneri, B.; Gaasterland, T.; Blencowe, B.J.; Darnell, R.B. An RNA map predicting Nova-dependent splicing regulation. *Nature* **2006**, *444*, 580–586. [[CrossRef](#)]
137. Mole, S.; McFarlane, M.; Chuen-Im, T.; Milligan, S.G.; Millan, D.; Graham, S.V. RNA splicing factors regulated by HPV16 during cervical tumour progression. *J. Pathol.* **2009**, *219*, 383–391. [[CrossRef](#)]
138. Jia, R.; Li, C.; McCoy, J.P.; Deng, C.X.; Zheng, Z.M. SRp20 is a proto-oncogene critical for cell proliferation and tumor induction and maintenance. *Int. J. Biol. Sci.* **2010**, *6*, 806–826. [[CrossRef](#)]
139. Liu, F.; Dai, M.; Xu, Q.; Zhu, X.; Zhou, Y.; Jiang, S.; Wang, Y.; Ai, Z.; Ma, L.; Zhang, Y.; et al. SRSF10-mediated IL1RAP alternative splicing regulates cervical cancer oncogenesis via mIL1RAP-NF- $\kappa$ B-CD47 axis. *Oncogene* **2018**, *37*, 2394–2409. [[CrossRef](#)]
140. Graham, S.V. Human Papillomavirus E2 Protein: Linking Replication, Transcription, and RNA Processing. *J. Virol.* **2016**, *90*, 8384–8388. [[CrossRef](#)]
141. Karni, R.; de Stanchina, E.; Lowe, S.W.; Sinha, R.; Mu, D.; Krainer, A.R. The gene encoding the splicing factor SF2/ASF is a proto-oncogene. *Nat. Struct. Mol. Biol.* **2007**, *14*, 185–193. [[CrossRef](#)]
142. Jia, R.; Ajiro, M.; Yu, L.; McCoy, P., Jr.; Zheng, Z.M. Oncogenic splicing factor SRSF3 regulates ILF3 alternative splicing to promote cancer cell proliferation and transformation. *RNA* **2019**, *25*, 630–644. [[CrossRef](#)] [[PubMed](#)]
143. Ajiro, M.; Jia, R.; Yang, Y.; Zhu, J.; Zheng, Z.M. A genome landscape of SRSF3-regulated splicing events and gene expression in human osteosarcoma U2OS cells. *Nucleic Acids Res.* **2016**, *44*, 1854–1870. [[CrossRef](#)] [[PubMed](#)]
144. Mole, S.; Milligan, S.G.; Graham, S.V. Human papillomavirus type 16 E2 protein transcriptionally activates the promoter of a key cellular splicing factor, SF2/ASF. *J. Virol.* **2009**, *83*, 357–367. [[CrossRef](#)] [[PubMed](#)]
145. Mole, S.; Faizo, A.A.A.; Hernandez-Lopez, H.; Griffiths, M.; Stevenson, A.; Roberts, S.; Graham, S.V. Human papillomavirus type 16 infection activates the host serine arginine protein kinase 1 (SRPK1)—Splicing factor axis. *J. Gen. Virol.* **2020**, *101*, 523–532. [[CrossRef](#)] [[PubMed](#)]
146. Klymenko, T.; Hernandez-Lopez, H.; MacDonald, A.I.; Bodily, J.M.; Graham, S.V. Human Papillomavirus E2 Regulates SRSF3 (SRp20) To Promote Capsid Protein Expression in Infected Differentiated Keratinocytes. *J. Virol.* **2016**, *90*, 5047–5058. [[CrossRef](#)]
147. Gauson, E.J.; Windle, B.; Donaldson, M.M.; Caffarel, M.M.; Dornan, E.S.; Coleman, N.; Herzyk, P.; Henderson, S.C.; Wang, X.; Morgan, I.M. Regulation of human genome expression and RNA splicing by human papillomavirus 16 E2 protein. *Virology* **2014**, *468–470*, 10–18. [[CrossRef](#)]
148. Bodaghi, S.; Jia, R.; Zheng, Z.M. Human papillomavirus type 16 E2 and E6 are RNA-binding proteins and inhibit in vitro splicing of pre-mRNAs with suboptimal splice sites. *Virology* **2009**, *386*, 32–43. [[CrossRef](#)]
149. Lai, M.C.; Teh, B.H.; Tarn, W.Y. A human papillomavirus E2 transcriptional activator. The interactions with cellular splicing factors and potential function in pre-mRNA processing. *J. Biol. Chem.* **1999**, *274*, 11832–11841. [[CrossRef](#)]



150. Jang, M.K.; Anderson, D.E.; van Doorslaer, K.; McBride, A.A. A proteomic approach to discover and compare interacting partners of papillomavirus E2 proteins from diverse phylogenetic groups. *Proteomics* **2015**, *15*, 2038–2050. [[CrossRef](#)]
151. Gui, J.F.; Lane, W.S.; Fu, X.D. A serine kinase regulates intracellular localization of splicing factors in the cell cycle. *Nature* **1994**, *369*, 678–682. [[CrossRef](#)]
152. Wang, H.Y.; Lin, W.; Dyck, J.A.; Yeakley, J.M.; Songyang, Z.; Cantley, L.C.; Fu, X.D. SRPK2: A differentially expressed SR protein-specific kinase involved in mediating the interaction and localization of pre-mRNA splicing factors in mammalian cells. *J. Cell Biol.* **1998**, *140*, 737–750. [[CrossRef](#)] [[PubMed](#)]
153. Prescott, E.L.; Brimacombe, C.L.; Hartley, M.; Bell, I.; Graham, S.; Roberts, S. Human papillomavirus type 1 E1<sup>E4</sup> 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.* **2014**, *88*, 12599–12611. [[CrossRef](#)] [[PubMed](#)]
154. Liu, X.; Mayeda, A.; Tao, M.; Zheng, Z.M. Exonic splicing enhancer-dependent selection of the bovine papillomavirus type 1 nucleotide 3225 3' splice site can be rescued in a cell lacking splicing factor ASF/SF2 through activation of the phosphatidylinositol 3-kinase/Akt pathway. *J. Virol.* **2003**, *77*, 2105–2115. [[CrossRef](#)] [[PubMed](#)]
155. Hummel, M.; Lim, H.B.; Laimins, L.A. Human papillomavirus type 31b late gene expression is regulated through protein kinase C-mediated changes in RNA processing. *J. Virol.* **1995**, *69*, 3381–3388. [[CrossRef](#)]
156. Dor, Y.; Cedar, H. Principles of DNA methylation and their implications for biology and medicine. *Lancet* **2018**, *392*, 777–786. [[CrossRef](#)]
157. Jones, P.A. Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* **2012**, *13*, 484–492. [[CrossRef](#)]
158. Lin, R.K.; Wu, C.Y.; Chang, J.W.; Juan, L.J.; Hsu, H.S.; Chen, C.Y.; Lu, Y.Y.; Tang, Y.A.; Yang, Y.C.; Yang, P.C.; et al. Dysregulation of p53/Sp1 control leads to DNA methyltransferase-1 overexpression in lung cancer. *Cancer Res.* **2010**, *70*, 5807–5817. [[CrossRef](#)]
159. Au Yeung, C.L.; Tsang, W.P.; Tsang, T.Y.; Co, N.N.; Yau, P.L.; Kwok, T.T. HPV-16 E6 upregulation of DNMT1 through repression of tumor suppressor p53. *Oncol. Rep.* **2010**, *24*, 1599–1604. [[CrossRef](#)]
160. McCabe, M.T.; Davis, J.N.; Day, M.L. Regulation of DNA methyltransferase 1 by the pRb/E2F1 pathway. *Cancer Res.* **2005**, *65*, 3624–3632. [[CrossRef](#)]
161. Leonard, S.M.; Wei, W.; Collins, S.I.; Pereira, M.; Diyaf, A.; Constandinou-Williams, C.; Young, L.S.; Roberts, S.; Woodman, C.B. Oncogenic human papillomavirus imposes an instructive pattern of DNA methylation changes which parallel the natural history of cervical HPV infection in young women. *Carcinogenesis* **2012**, *33*, 1286–1293. [[CrossRef](#)]
162. Burgers, W.A.; Blanchon, L.; Pradhan, S.; de Launoit, Y.; Kouzarides, T.; Fuks, F. Viral oncoproteins target the DNA methyltransferases. *Oncogene* **2007**, *26*, 1650–1655. [[CrossRef](#)] [[PubMed](#)]
163. Clarke, M.A.; Gradissimo, A.; Schiffman, M.; Lam, J.; Sollecito, C.C.; Fetterman, B.; Lorey, T.; Poitras, N.; Raine-Bennett, T.R.; Castle, P.E.; et al. Human Papillomavirus DNA Methylation as a Biomarker for Cervical Precancer: Consistency across 12 Genotypes and Potential Impact on Management of HPV-Positive Women. *Clin. Cancer Res.* **2018**, *24*, 2194–2202. [[CrossRef](#)] [[PubMed](#)]
164. Lorincz, A.T.; Brentnall, A.R.; Scibior-Bentkowska, D.; Reuter, C.; Banwait, R.; Cadman, L.; Austin, J.; Cuzick, J.; Vasiljević, N. Validation of a DNA methylation HPV triage classifier in a screening sample. *Int. J. Cancer* **2016**, *138*, 2745–2751. [[CrossRef](#)] [[PubMed](#)]
165. Ekanayake Weeramange, C.; Tang, K.D.; Vasani, S.; Langton-Lockton, J.; Kenny, L.; Punyadeera, C. DNA Methylation Changes in Human Papillomavirus-Driven Head and Neck Cancers. *Cells* **2020**, *9*, 1359. [[CrossRef](#)] [[PubMed](#)]
166. Vasiljević, N.; Scibior-Bentkowska, D.; Brentnall, A.R.; Cuzick, J.; Lorincz, A.T. Credentialing of DNA methylation assays for human genes as diagnostic biomarkers of cervical intraepithelial neoplasia in high-risk HPV positive women. *Gynecol. Oncol.* **2014**, *132*, 709–714. [[CrossRef](#)] [[PubMed](#)]
167. Kan, Y.Y.; Liou, Y.L.; Wang, H.J.; Chen, C.Y.; Sung, L.C.; Chang, C.F.; Liao, C.I. PAX1 methylation as a potential biomarker for cervical cancer screening. *Int. J. Gynecol. Cancer* **2014**, *24*, 928–934. [[CrossRef](#)]
168. Ramírez, A.T.; Sánchez, G.I.; Nedjai, B.; Agudelo, M.C.; Brentnall, A.R.; Cuschieri, K.; Castañeda, K.M.; Cuzick, J.; Lorincz, A.T. Effective methylation triage of HPV positive women with abnormal cytology in a middle-income country. *Int. J. Cancer* **2021**, *148*, 1383–1393. [[CrossRef](#)]
169. Banila, C.; Lorincz, A.T.; Scibior-Bentkowska, D.; Clifford, G.M.; Kumbi, B.; Beyene, D.; Wheeler, C.M.; Cuschieri, K.; Cuzick, J.; Nedjai, B. Clinical performance of methylation as a biomarker for cervical carcinoma in situ and cancer diagnosis: A worldwide study. *Int. J. Cancer* **2022**, *150*, 290–302. [[CrossRef](#)]
170. Chang, C.L.; Ho, S.C.; Su, Y.F.; Juan, Y.C.; Huang, C.Y.; Chao, A.S.; Hsu, Z.S.; Chang, C.F.; Fwu, C.W.; Chang, T.C. DNA methylation marker for the triage of hrHPV positive women in cervical cancer screening: Real-world evidence in Taiwan. *Gynecol. Oncol.* **2021**, *161*, 429–435. [[CrossRef](#)]
171. Burley, M.; Roberts, S.; Parish, J.L. Epigenetic regulation of human papillomavirus transcription in the productive virus life cycle. *Semin. Immunopathol.* **2020**, *42*, 159–171. [[CrossRef](#)]
172. Thain, A.; Jenkins, O.; Clarke, A.R.; Gaston, K. CpG methylation directly inhibits binding of the human papillomavirus type 16 E2 protein to specific DNA sequences. *J. Virol.* **1996**, *70*, 7233–7235. [[CrossRef](#)] [[PubMed](#)]
173. Kim, K.; Garner-Hamrick, P.A.; Fisher, C.; Lee, D.; Lambert, P.F. Methylation patterns of papillomavirus DNA, its influence on E2 function, and implications in viral infection. *J. Virol.* **2003**, *77*, 12450–12459. [[CrossRef](#)] [[PubMed](#)]

174. Kalantari, M.; Calleja-Macias, I.E.; Tewari, D.; Hagmar, B.; Lie, K.; Barrera-Saldana, H.A.; Wiley, D.J.; Bernard, H.U. Conserved methylation patterns of human papillomavirus type 16 DNA in asymptomatic infection and cervical neoplasia. *J. Virol.* **2004**, *78*, 12762–12772. [[CrossRef](#)] [[PubMed](#)]
175. Mirabello, L.; Schiffman, M.; Ghosh, A.; Rodriguez, A.C.; Vasiljevic, N.; Wentzensen, N.; Herrero, R.; Hildesheim, A.; Wacholder, S.; Scibior-Bentkowska, D.; et al. Elevated methylation of HPV16 DNA is associated with the development of high grade cervical intraepithelial neoplasia. *Int. J. Cancer* **2013**, *132*, 1412–1422. [[CrossRef](#)]
176. Mirabello, L.; Sun, C.; Ghosh, A.; Rodriguez, A.C.; Schiffman, M.; Wentzensen, N.; Hildesheim, A.; Herrero, R.; Wacholder, S.; Lorincz, A.; et al. Methylation of human papillomavirus type 16 genome and risk of cervical precancer in a Costa Rican population. *J. Natl. Cancer Inst.* **2012**, *104*, 556–565. [[CrossRef](#)]
177. Sun, C.; Reimers, L.L.; Burk, R.D. Methylation of HPV16 genome CpG sites is associated with cervix precancer and cancer. *Gynecol. Oncol.* **2011**, *121*, 59–63. [[CrossRef](#)]
178. Kalantari, M.; Osann, K.; Calleja-Macias, I.E.; Kim, S.; Yan, B.; Jordan, S.; Chase, D.M.; Tewari, K.S.; Bernard, H.U. Methylation of human papillomavirus 16, 18, 31, and 45 L2 and L1 genes and the cellular DAPK gene: Considerations for use as biomarkers of the progression of cervical neoplasia. *Virology* **2014**, *448*, 314–321. [[CrossRef](#)]
179. Nishikura, K. Functions and regulation of RNA editing by ADAR deaminases. *Annu. Rev. Biochem.* **2010**, *79*, 321–349. [[CrossRef](#)]
180. Tan, M.H.; Li, Q.; Shanmugam, R.; Piskol, R.; Kohler, J.; Young, A.N.; Liu, K.I.; Zhang, R.; Ramaswami, G.; Ariyoshi, K.; et al. Dynamic landscape and regulation of RNA editing in mammals. *Nature* **2017**, *550*, 249–254. [[CrossRef](#)]
181. Chen, Y.; Wang, H.; Lin, W.; Shuai, P. ADAR1 overexpression is associated with cervical cancer progression and angiogenesis. *Diagn. Pathol.* **2017**, *12*, 12. [[CrossRef](#)]
182. Snoek, B.C.; Babion, I.; Koppers-Lalic, D.; Pegtel, D.M.; Steenbergen, R.D. Altered microRNA processing proteins in HPV-induced cancers. *Curr. Opin. Virol.* **2019**, *39*, 23–32. [[CrossRef](#)] [[PubMed](#)]
183. Jarmuz, A.; Chester, A.; Bayliss, J.; Gisbourne, J.; Dunham, I.; Scott, J.; Navaratnam, N. An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. *Genomics* **2002**, *79*, 285–296. [[CrossRef](#)] [[PubMed](#)]
184. Delviks-Frankenberry, K.A.; Desimmie, B.A.; Pathak, V.K. Structural Insights into APOBEC3-Mediated Lentiviral Restriction. *Viruses* **2020**, *12*, 587. [[CrossRef](#)] [[PubMed](#)]
185. Vartanian, J.P.; Guétard, D.; Henry, M.; Wain-Hobson, S. Evidence for editing of human papillomavirus DNA by APOBEC3 in benign and precancerous lesions. *Science* **2008**, *320*, 230–233. [[CrossRef](#)]
186. Henderson, S.; Chakravarthy, A.; Su, X.; Boshoff, C.; Fenton, T.R. APOBEC-mediated cytosine deamination links PIK3CA helical domain mutations to human papillomavirus-driven tumor development. *Cell Rep.* **2014**, *7*, 1833–1841. [[CrossRef](#)]
187. Vieira, V.C.; Leonard, B.; White, E.A.; Starrett, G.J.; Temiz, N.A.; Lorenz, L.D.; Lee, D.; Soares, M.A.; Lambert, P.F.; Howley, P.M.; et al. Human papillomavirus E6 triggers upregulation of the antiviral and cancer genomic DNA deaminase APOBEC3B. *mBio* **2014**, *5*, e02234-14. [[CrossRef](#)]
188. Warren, C.J.; Xu, T.; Guo, K.; Griffin, L.M.; Westrich, J.A.; Lee, D.; Lambert, P.F.; Santiago, M.L.; Pyeon, D. APOBEC3A functions as a restriction factor of human papillomavirus. *J. Virol.* **2015**, *89*, 688–702. [[CrossRef](#)]
189. Periyasamy, M.; Singh, A.K.; Gemma, C.; Kranjec, C.; Farzan, R.; Leach, D.A.; Navaratnam, N.; Pálincás, H.L.; Vértessy, B.G.; Fenton, T.R.; et al. p53 controls expression of the DNA deaminase APOBEC3B to limit its potential mutagenic activity in cancer cells. *Nucleic Acids Res.* **2017**, *45*, 11056–11069. [[CrossRef](#)]
190. Westrich, J.A.; Warren, C.J.; Klausner, M.J.; Guo, K.; Liu, C.W.; Santiago, M.L.; Pyeon, D. Human Papillomavirus 16 E7 Stabilizes APOBEC3A Protein by Inhibiting Cullin 2-Dependent Protein Degradation. *J. Virol.* **2018**, *92*, e01318-17. [[CrossRef](#)]
191. Hirose, Y.; Onuki, M.; Tenjimbayashi, Y.; Mori, S.; Ishii, Y.; Takeuchi, T.; Tasaka, N.; Satoh, T.; Morisada, T.; Iwata, T.; et al. Within-Host Variations of Human Papillomavirus Reveal APOBEC Signature Mutagenesis in the Viral Genome. *J. Virol.* **2018**, *92*, e00017-18. [[CrossRef](#)]
192. Mirabello, L.; Yeager, M.; Yu, K.; Clifford, G.M.; Xiao, Y.; Zhu, B.; Cullen, M.; Boland, J.F.; Wentzensen, N.; Nelson, C.W.; et al. HPV16 E7 Genetic Conservation Is Critical to Carcinogenesis. *Cell* **2017**, *170*, 1164–1174.e6. [[CrossRef](#)] [[PubMed](#)]
193. Wang, Z.; Wakae, K.; Kitamura, K.; Aoyama, S.; Liu, G.; Koura, M.; Monjurul, A.M.; Kukimoto, I.; Muramatsu, M. APOBEC3 deaminases induce hypermutation in human papillomavirus 16 DNA upon beta interferon stimulation. *J. Virol.* **2014**, *88*, 1308–1317. [[CrossRef](#)]
194. Kukimoto, I.; Mori, S.; Aoyama, S.; Wakae, K.; Muramatsu, M.; Kondo, K. Hypermutation in the E2 gene of human papillomavirus type 16 in cervical intraepithelial neoplasia. *J. Med. Virol.* **2015**, *87*, 1754–1760. [[CrossRef](#)] [[PubMed](#)]
195. Faden, D.L.; Kuhs, K.A.L.; Lin, M.; Langenbucher, A.; Pinheiro, M.; Yeager, M.; Cullen, M.; Boland, J.F.; Steinberg, M.; Bass, S.; et al. APOBEC Mutagenesis Is Concordant between Tumor and Viral Genomes in HPV-Positive Head and Neck Squamous Cell Carcinoma. *Viruses* **2021**, *13*, 1666. [[CrossRef](#)] [[PubMed](#)]
196. Wakae, K.; Aoyama, S.; Wang, Z.; Kitamura, K.; Liu, G.; Monjurul, A.M.; Koura, M.; Imayasu, M.; Sakamoto, N.; Nakamura, M.; et al. Detection of hypermutated human papillomavirus type 16 genome by Next-Generation Sequencing. *Virology* **2015**, *485*, 460–466. [[CrossRef](#)] [[PubMed](#)]
197. TCGA Research Network. Integrated genomic and molecular characterization of cervical cancer. *Nature* **2017**, *543*, 378–384. [[CrossRef](#)]
198. Gillison, M.L.; Akagi, K.; Xiao, W.; Jiang, B.; Pickard, R.K.L.; Li, J.; Swanson, B.J.; Agrawal, A.D.; Zucker, M.; Stache-Crain, B.; et al. Human papillomavirus and the landscape of secondary genetic alterations in oral cancers. *Genome Res.* **2019**, *29*, 1–17. [[CrossRef](#)]



199. Buisson, R.; Langenbucher, A.; Bowen, D.; Kwan, E.E.; Benes, C.H.; Zou, L.; Lawrence, M.S. Passenger hotspot mutations in cancer driven by APOBEC3A and mesoscale genomic features. *Science* **2019**, *364*, eaaw2872. [[CrossRef](#)]
200. Swanton, C.; McGranahan, N.; Starrett, G.J.; Harris, R.S. APOBEC Enzymes: Mutagenic Fuel for Cancer Evolution and Heterogeneity. *Cancer Discov.* **2015**, *5*, 704–712. [[CrossRef](#)]
201. Ahasan, M.M.; Wakae, K.; Wang, Z.; Kitamura, K.; Liu, G.; Koura, M.; Imayasu, M.; Sakamoto, N.; Hanaoka, K.; Nakamura, M.; et al. APOBEC3A and 3C decrease human papillomavirus 16 pseudovirion infectivity. *Biochem. Biophys. Res. Commun.* **2015**, *457*, 295–299. [[CrossRef](#)]
202. Warren, C.J.; Westrich, J.A.; Doorslaer, K.V.; Pyeon, D. Roles of APOBEC3A and APOBEC3B in Human Papillomavirus Infection and Disease Progression. *Viruses* **2017**, *9*, 233. [[CrossRef](#)] [[PubMed](#)]
203. Zhu, B.; Xiao, Y.; Yeager, M.; Clifford, G.; Wentzensen, N.; Cullen, M.; Boland, J.F.; Bass, S.; Steinberg, M.K.; Raine-Bennett, T.; et al. Mutations in the HPV16 genome induced by APOBEC3 are associated with viral clearance. *Nat. Commun.* **2020**, *11*, 886. [[CrossRef](#)] [[PubMed](#)]
204. Lorenzo, J.P.; Molla, L.; Ibarra, I.L.; Ruf, S.; Ridani, J.; Subramani, P.G.; Boulais, J.; Harjanto, D.; Vonica, A.; di Noia, J.M.; et al. APOBEC2 is a Transcriptional Repressor required for proper Myoblast Differentiation. *bioRxiv* **2021**. [[CrossRef](#)]
205. Cladel, N.M.; Jiang, P.; Li, J.J.; Peng, X.; Cooper, T.K.; Majerciak, V.; Balogh, K.K.; Meyer, T.J.; Brendle, S.A.; Budgeon, L.R.; et al. Papillomavirus can be transmitted through the blood and produce infections in blood recipients: Evidence from two animal models. *Emerg. Microbes Infect.* **2019**, *8*, 1108–1121. [[CrossRef](#)] [[PubMed](#)]
206. Mangeat, B.; Turelli, P.; Caron, G.; Friedli, M.; Perrin, L.; Trono, D. Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature* **2003**, *424*, 99–103. [[CrossRef](#)]
207. Henderson, S.; Fenton, T. APOBEC3 genes: Retroviral restriction factors to cancer drivers. *Trends Mol. Med.* **2015**, *21*, 274–284. [[CrossRef](#)]
208. Chakravarthy, A.; Henderson, S.; Thirdborough, S.M.; Ottensmeier, C.H.; Su, X.; Lechner, M.; Feber, A.; Thomas, G.J.; Fenton, T.R. Human Papillomavirus Drives Tumor Development throughout the Head and Neck: Improved Prognosis Is Associated with an Immune Response Largely Restricted to the Oropharynx. *J. Clin. Oncol.* **2016**, *34*, 4132–4141. [[CrossRef](#)]
209. Chen, T.W.; Lee, C.C.; Liu, H.; Wu, C.S.; Pickering, C.R.; Huang, P.J.; Wang, J.; Chang, I.Y.; Yeh, Y.M.; Chen, C.D.; et al. APOBEC3A is an oral cancer prognostic biomarker in Taiwanese carriers of an APOBEC deletion polymorphism. *Nat. Commun.* **2017**, *8*, 465. [[CrossRef](#)]
210. Salter, J.D.; Bennett, R.P.; Smith, H.C. The APOBEC Protein Family: United by Structure, Divergent in Function. *Trends Biochem. Sci.* **2016**, *41*, 578–594. [[CrossRef](#)]
211. Smith, N.J.; Fenton, T.R. The APOBEC3 genes and their role in cancer: Insights from human papillomavirus. *J. Mol. Endocrinol.* **2019**, *62*, R269–R287. [[CrossRef](#)]
212. Hoopes, J.I.; Cortez, L.M.; Mertz, T.M.; Malc, E.P.; Mieczkowski, P.A.; Roberts, S.A. APOBEC3A and APOBEC3B Preferentially Deaminate the Lagging Strand Template during DNA Replication. *Cell Rep.* **2016**, *14*, 1273–1282. [[CrossRef](#)] [[PubMed](#)]
213. Seplyarskiy, V.B.; Soldatov, R.A.; Popadin, K.Y.; Antonarakis, S.E.; Bazykin, G.A.; Nikolaev, S.I. APOBEC-induced mutations in human cancers are strongly enriched on the lagging DNA strand during replication. *Genome Res.* **2016**, *26*, 174–182. [[CrossRef](#)] [[PubMed](#)]
214. Chapman, S.; Liu, X.; Meyers, C.; Schlegel, R.; McBride, A.A. Human keratinocytes are efficiently immortalized by a Rho kinase inhibitor. *J. Clin. Investig.* **2010**, *120*, 2619–2626. [[CrossRef](#)] [[PubMed](#)]
215. Kajitani, N.; Schwartz, S. Role of Viral Ribonucleoproteins in Human Papillomavirus Type 16 Gene Expression. *Viruses* **2020**, *12*, 1110. [[CrossRef](#)]
216. Paris, C.; Pentland, I.; Groves, I.; Roberts, D.C.; Powis, S.J.; Coleman, N.; Roberts, S.; Parish, J.L. CCCTC-binding factor recruitment to the early region of the human papillomavirus 18 genome regulates viral oncogene expression. *J. Virol.* **2015**, *89*, 4770–4785. [[CrossRef](#)]
217. Ruiz, F.J.; Inkman, M.; Rashmi, R.; Muhammad, N.; Gabriel, N.; Miller, C.A.; McLellan, M.D.; Goldstein, M.; Markovina, S.; Grigsby, P.W.; et al. HPV transcript expression affects cervical Cancer Response to chemoradiation. *JCI Insight* **2021**, *6*, e138734. [[CrossRef](#)]
218. Wang, H.K.; Duffy, A.A.; Broker, T.R.; Chow, L.T. Robust production and passaging of infectious HPV in squamous epithelium of primary human keratinocytes. *Genes Dev.* **2009**, *23*, 181–194. [[CrossRef](#)]
219. Lee, S.C.; Abdel-Wahab, O. Therapeutic targeting of splicing in cancer. *Nat. Med.* **2016**, *22*, 976–986. [[CrossRef](#)]
220. Havens, M.A.; Duelli, D.M.; Hastings, M.L. Targeting RNA splicing for disease therapy. *Wiley Interdiscip. Rev. RNA* **2013**, *4*, 247–266. [[CrossRef](#)]
221. Bowling, E.A.; Wang, J.H.; Gong, F.; Wu, W.; Neill, N.J.; Kim, I.S.; Tyagi, S.; Orellana, M.; Kurley, S.J.; Dominguez-Vidaña, R.; et al. Spliceosome-targeted therapies trigger an antiviral immune response in triple-negative breast cancer. *Cell* **2021**, *184*, 384–403.e21. [[CrossRef](#)]
222. Lareau, L.F.; Hite, D.H.; Hogan, G.J.; Brown, P.O. Distinct stages of the translation elongation cycle revealed by sequencing ribosome-protected mRNA fragments. *eLife* **2014**, *3*, e01257. [[CrossRef](#)] [[PubMed](#)]
223. Martens, A.T.; Taylor, J.; Hilser, V.J. Ribosome A and P sites revealed by length analysis of ribosome profiling data. *Nucleic Acids Res.* **2015**, *43*, 3680–3687. [[CrossRef](#)] [[PubMed](#)]
224. Ingolia, N.T. Ribosome Footprint Profiling of Translation throughout the Genome. *Cell* **2016**, *165*, 22–33. [[CrossRef](#)] [[PubMed](#)]

225. Wu, C.C.; Zinshteyn, B.; Wehner, K.A.; Green, R. High-Resolution Ribosome Profiling Defines Discrete Ribosome Elongation States and Translational Regulation during Cellular Stress. *Mol. Cell* **2019**, *73*, 959–970.e5. [[CrossRef](#)]
226. Bartel, D.P. Metazoan MicroRNAs. *Cell* **2018**, *173*, 20–51. [[CrossRef](#)]
227. Dong, P.; Xiong, Y.; Yue, J.; Hanley, S.J.B.; Kobayashi, N.; Todo, Y.; Watari, H. Long Non-coding RNA NEAT1: A Novel Target for Diagnosis and Therapy in Human Tumors. *Front. Genet.* **2018**, *9*, 471. [[CrossRef](#)]
228. Statello, L.; Guo, C.J.; Chen, L.L.; Huarte, M. Gene regulation by long non-coding RNAs and its biological functions. *Nat. Rev. Mol. Cell Biol.* **2021**, *22*, 96–118. [[CrossRef](#)]
229. Wang, X.; Wang, H.K.; Li, Y.; Hafner, M.; Banerjee, N.S.; Tang, S.; Briskin, D.; Meyers, C.; Chow, L.T.; Xie, X.; et al. microRNAs are biomarkers of oncogenic human papillomavirus infections. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 4262–4267. [[CrossRef](#)]
230. Wang, X.; Tang, S.; Le, S.Y.; Lu, R.; Rader, J.S.; Meyers, C.; Zheng, Z.M. Aberrant expression of oncogenic and tumor-suppressive microRNAs in cervical cancer is required for cancer cell growth. *PLoS ONE* **2008**, *3*, e2557. [[CrossRef](#)]
231. Wang, X.; Wang, H.K.; McCoy, J.P.; Banerjee, N.S.; Rader, J.S.; Broker, T.R.; Meyers, C.; Chow, L.T.; Zheng, Z.M. Oncogenic HPV infection interrupts the expression of tumor-suppressive miR-34a through viral oncoprotein E6. *RNA* **2009**, *15*, 637–647. [[CrossRef](#)]
232. Melar-New, M.; Laimins, L.A. Human papillomaviruses modulate expression of microRNA 203 upon epithelial differentiation to control levels of p63 proteins. *J. Virol.* **2010**, *84*, 5212–5221. [[CrossRef](#)] [[PubMed](#)]
233. Lajer, C.B.; Garnæs, E.; Friis-Hansen, L.; Norrild, B.; Therkildsen, M.H.; Glud, M.; Rossing, M.; Lajer, H.; Svane, D.; Skotte, L.; et al. The role of miRNAs in human papilloma virus (HPV)-associated cancers: Bridging between HPV-related head and neck cancer and cervical cancer. *Br. J. Cancer* **2012**, *106*, 1526–1534. [[CrossRef](#)] [[PubMed](#)]
234. Snoek, B.C.; Verlaat, W.; Babion, I.; Novianti, P.W.; van de Wiel, M.A.; Wilting, S.M.; van Trommel, N.E.; Bleeker, M.C.G.; Massuger, L.; Melchers, W.J.G.; et al. Genome-wide microRNA analysis of HPV-positive self-samples yields novel triage markers for early detection of cervical cancer. *Int. J. Cancer* **2019**, *144*, 372–379. [[CrossRef](#)] [[PubMed](#)]
235. Tong, F.; Andress, A.; Tang, G.; Liu, P.; Wang, X. Comprehensive profiling of extracellular RNA in HPV-induced cancers using an improved pipeline for small RNA-seq analysis. *Sci. Rep.* **2020**, *10*, 19450. [[CrossRef](#)]
236. Liu, H.; Xu, J.; Yang, Y.; Wang, X.; Wu, E.; Majerciak, V.; Zhang, T.; Steenbergen, R.D.M.; Wang, H.K.; Banerjee, N.S.; et al. Oncogenic HPV promotes the expression of the long noncoding RNA Inc-FANCI-2 through E7 and YY1. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2014195118. [[CrossRef](#)]
237. Martínez-Noël, G.; Szajner, P.; Kramer, R.E.; Boyland, K.A.; Sheikh, A.; Smith, J.A.; Howley, P.M. Identification of MicroRNAs That Stabilize p53 in Human Papillomavirus-Positive Cancer Cells. *J. Virol.* **2022**, *96*, e0186521. [[CrossRef](#)]
238. Bartel, D.P. MicroRNAs: Target recognition and regulatory functions. *Cell* **2009**, *136*, 215–233. [[CrossRef](#)]
239. Yao, R.W.; Wang, Y.; Chen, L.L. Cellular functions of long noncoding RNAs. *Nat. Cell Biol.* **2019**, *21*, 542–551. [[CrossRef](#)]
240. Briata, P.; Gherzi, R. Long Non-Coding RNA-Ribonucleoprotein Networks in the Post-Transcriptional Control of Gene Expression. *Noncoding RNA* **2020**, *6*, 40. [[CrossRef](#)]
241. Pardini, B.; De Maria, D.; Francavilla, A.; Di Gaetano, C.; Ronco, G.; Naccarati, A. MicroRNAs as markers of progression in cervical cancer: A systematic review. *BMC Cancer* **2018**, *18*, 696. [[CrossRef](#)]
242. Mitra, T.; Elangovan, S. Cervical cancer development, chemoresistance, and therapy: A snapshot of involvement of microRNA. *Mol. Cell. Biochem.* **2021**, *476*, 4363–4385. [[CrossRef](#)] [[PubMed](#)]