LUND UNIVERSITY

Regulation of human papillomavirus 16 gene expression at the level of mRNA processing

Hao, Chengyu

2022

Document Version:
Publisher's PDF, also known as Version of record

Regulation of human papillomavirus 16 genes expression at the level of mRNA processing

CHENGYU HAO
DEPARTMENT OF LABORATORY MEDICINE | FACULTY OF MEDICINE | LUND UNIVERSITY

Regulation of human papillomavirus 16 genes expression at the level of mRNA processing

Regulation of human papillomavirus 16 genes expression at the level of mRNA processing

Chengyu Hao

LUND
UNIVERSITY

DOCTORAL DISSERTATION

by due permission of the Faculty of Medicine, Lund University, Sweden.
To be defended on September 1, 2022, at 9:00 in Segerfalksalen, BMC, Lund.

Faculty opponent
Prof. Magnus Johansson

Örebro University

Organization LUND UNIVERSITY	**Document name** Doctoral dissertation
	Date of issue September 1 2022
Author(s): Chengyu Hao	Sponsoring organization

Title and subtitle Regulation of human papillomavirus 16 genes expression at the level of mRNA processing

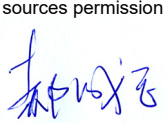
Abstract

Human papillomavirus (HPV) are small double-stranded DNA viruses. To date, over 400 genotypes have been identified. HPV type 16 and HPV type 18 are associated with approximately 70% of all cervical cancers in which HPV16 accounts for about 50%. HPV16 genome contains eight genes which are broadly divided into early genes (E1, E2, E4, E5, E6, and E7) and late genes (L1 and L2). Generally, early genes encode proteins E1 to E7 that promote cell proliferation and enable viral DNA replication while late genes L1 and L2 encode structural proteins that package virus DNA genome to form progeny virions. Alternative splicing plays a major role in generation of the range of mRNAs required to express these proteins. Regulation of alternative splicing negatively or positively could change many functional aspects of mRNAs and their encoded proteins. The RNA-binding proteins, such as serine/arginine (SR)-rich protein family and heterogeneous nuclear ribonucleoprotein (hnRNP) family, are the most prominent mediators of splice site recognition and contribute to alternative splicing regulation. So, it is of great interest to investigate how the expression of the HPV16 early and late genes is regulated by alternative splicing or which splicing regulatory factors.

Our results demonstrate regulation of HPV16 early and late genes expression or mRNA alternative splicing by cis-acting RNA elements and trans-acting regulatory factors. We found that the majority of all hnRNPs had the potential to control HPV16 early and late genes expression and that most SR proteins contained the potential to regulate HPV16 late genes expression. We also found a cis-acting RNA element downstream of HPV16 E2 splice site SA2709 required for efficient production of spliced HPV16 E2 mRNA. We further found that hnRNP G to be the corresponding regulator of this enhancer element of HPV16 E2 spliced mRNA. We also found that hnRNP G has different roles in regulating of HPV16 early E6/E7 and E1/E2 genes expression as well as mRNA processing. When HPV16 early promoter is active, hnRNP G has a splicing inhibitory effect that inhibits production of E7 mRNAs, thereby reducing expression of E7 protein. When transcription of HPV16 genome switches from the early promoter to the late promoter, the activation of hnRNP G by cell differentiation and DNA damage response recruit the splicing factor U2AF65 to exert a positive effect on E2 mRNA splicing through interactions with the splicing enhancer downstream of SA2709, thereby promote expression of E2 protein. Moreover, the enhancement of HPV16 E2 mRNA splicing and inhibition of HPV16 E6/E7 mRNA splicing are exerted by non-overlapping domains of hnRNP G. Furthermore, the enhancement of HPV16 E2 mRNA splicing and the interaction of hnRNP G with HPV16 E2 mRNA and with U2AF65 are regulated by sumoylation of hnRNP G upon activation of the DNA damage response. In the separate project, we demonstrated hnRNP D can regulate HPV16 early gene E1 and E6 mRNAs expression by promoting intron retention and export E1- and E6-encoding mRNAs to the cytoplasm. Finally, we deviced two subgenomic HPV16 reporter plasmids named pE2sLuc and pE1NIuc that may be used to establish bioassays that monitor either expression of E1 encoded by the intron retained, unspliced E1 mRNAs or E2 protein produced by the spliced E2 mRNAs that utilize SA2709 or the small E1C protein encoded by the alternatively spliced E1C mRNA utilizing SA2582.

In summary, in this thesis we demonstrate the vital role of one cis-acting enhancer element downstream of HPV16 E2 mRNA splice site SA2709 and the multifunctional roles of the two trans-acting regulatory proteins hnRNP G and hnRNP D in HPV16 gene expression and mRNAs processing. Both trans-acting proteins hnRNP G and hnRNP D show critical and irreplaceable functions in HPV16 oncogenic E6/E7 and regulatory E1/E2 gene expression. These results lay the foundation and pave ways for development of biomarkers or targets for treatment of HPV16 infection or cancer. Furthermore, the bioassays could be used to identify antiviral or antitumor drugs to HPV16- infections and HPV-driven cancer.

Key words		
Classification system and/or index terms (if any)		
Supplementary bibliographical information		**Language** English
ISSN 1652-8220 Lund University, Faculty of Medicine Doctoral Dissertation Series 2022:102		**ISBN** 978-91-8021-263-2
Recipient's notes	**Number of pages** 87	Price
	Security classification	

Signature 

Date 2022-06-17

Regulation of human papillomavirus 16 genes expression at the level of mRNA processing

Chengyu Hao

LUND
UNIVERSITY

Coverphoto by Chengyu Hao

Faculty of Medicine
Department of Laboratory Medicine

ISBN 978-91-8021-263-2
ISSN 1652-8220

Printed in Sweden by Media-Tryck, Lund University
Lund 2022

Dedicated to my beloved parents, husband and child

If there is a will, there is a way

Table of Contents

List of publications

1. **Hao, C**., Gong, L., Cui, X., Jönsson, J., Zheng, Y., Wu, C., Kajitani, N., Schwartz, S. 2021. Identification of heterogenous nuclear ribonucleoproteins (hnRNPs) and serine and arginine rich (SR) proteins that induce human papillomavirus type 16 late gene expression and alter L1 mRNA splicing. **Archives of Virol**., 167:563-70.

2. **Hao, C**., Zheng, Y., Jönsson, J., Cui, X., Yu, H., Wu, C., Kajitani, N., Schwartz, S. 2022. hnRNP G/RBMX enhances HPV16 E2 mRNA splicing through a novel splicing enhancer and inhibits production of spliced E7 oncogene mRNAs. **Nucleic Acids Res**., Doi:10.1093/nar/gkac213.

3. Cui, X., **Hao, C**., Gong, L., Kajitani, N., Schwartz, S. 2022. hnRNP D activates production HPV 16 E1 and E6 mRNAs by promoting intron retention. **Nucleic Acids Res**., Mar 2:gkac132. Doi:10.1093/nar/gkac132.

4. **Hao, C**., Zheng, Y., Cui, X., Jönsson, J., Wu, C., Kajitani, N., Schwartz, S. 2022. Generation of reporter plasmids for HPV16 E1 or E2 mRNA splicing using nanoluciferase and secreted luciferase. **Manuscript**.

Articles not included in the thesis

1. Zheng, Y., Jönsson, J., **Hao, C**., 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." **Journal of Virology** 94, no. 20 (2020): e01509-20. DOI:10.1128/JVI.01509-20

Abstract

Human papillomavirus (HPV) are small double-stranded DNA viruses. To date, over 400 genotypes have been identified. HPV type 16 and HPV type 18 are associated with approximately 70% of all cervical cancers in which HPV16 accounts for about 50%. HPV16 genome contains eight genes which are broadly divided into early genes (E1, E2, E4, E5, E6, and E7) and late genes (L1 and L2). Generally, early genes encode proteins E1 to E7 that promote cell proliferation and enable viral DNA replication while late genes L1 and L2 encode structural proteins that package virus DNA genome to form progeny virions. Alternative splicing plays a major role in generation of the range of mRNAs required to express these proteins. Regulation of alternative splicing negatively or positively could change many functional aspects of mRNAs and their encoded proteins. The RNA-binding proteins, such as serine/arginine (SR)-rich protein family and heterogeneous nuclear ribonucleoprotein (hnRNP) family, are the most prominent mediators of splice site recognition and contribute to alternative splicing regulation. So, it is of great interest to investigate how the expression of the HPV16 early and late genes is regulated by alternative splicing or which splicing regulatory factors.

Our results demonstrate regulation of HPV16 early and late genes expression or mRNA alternative splicing by cis-acting RNA elements and trans-acting regulatory factors. We found that the majority of all hnRNPs had the potential to control HPV16 early and late genes expression and that most SR proteins contained the potential to regulate HPV16 late genes expression. We also found a cis-acting RNA element downstream of HPV16 E2 splice site SA2709 required for efficient production of spliced HPV16 E2 mRNA. We further found that hnRNP G to be the corresponding regulator of this enhancer element of HPV16 E2 spliced mRNA. We also found that hnRNP G has different roles in regulating of HPV16 early E6/E7 and E1/E2 genes expression as well as mRNA processing. When HPV16 early promoter is active, hnRNP G has a splicing inhibitory effect that inhibits production of E7 mRNAs, thereby reducing expression of E7 protein. When transcription of HPV16 genome switches from the early promoter to the late promoter, the activation of hnRNP G by cell differentiation and DNA damage response recruit the splicing factor U2AF65 to exert a positive effect on E2 mRNA splicing through interactions with the splicing enhancer downstream of SA2709, thereby promote expression of E2 protein. Moreover, the enhancement of HPV16 E2 mRNA splicing and inhibition of HPV16 E6/E7 mRNA splicing are exerted by non-overlapping

domains of hnRNP G. Furthermore, the enhancement of HPV16 E2 mRNA splicing and the interaction of hnRNP G with HPV16 E2 mRNA and with U2AF65 are regulated by sumoylation of hnRNP G upon activation of the DNA damage response. In the separate project, we demonstrated hnRNP D can regulate HPV16 early gene E1 and E6 mRNAs expression by promoting intron retention and export E1- and E6-encoding mRNAs to the cytoplasm. Finally, we devised two subgenomic HPV16 reporter plasmids named pE2sLuc and pE1Nluc that may be used to establish bioassays that monitor either expression of E1 encoded by the intron retained, unspliced E1 mRNAs or E2 protein produced by the spliced E2 mRNAs that utilize SA2709 or the small E1C protein encoded by the alternatively spliced E1C mRNA utilizing SA2582.

In summary, in this thesis we demonstrate the vital role of one cis-acting enhancer element downstream of HPV16 E2 mRNA splice site SA2709 and the multifunctional roles of the two trans-acting regulatory proteins hnRNP G and hnRNP D in HPV16 gene expression and mRNAs processing. Both trans-acting proteins hnRNP G and hnRNP D show critical and irreplaceable functions in HPV16 oncogenic E6/E7 and regulatory E1/E2 gene expression. These results lay the foundation and pave ways for development of biomarkers or targets for treatment of HPV16 infection or cancer. Furthermore, the bioassays could be used to identify antiviral or antitumor drugs to HPV16- infections and HPV-driven cancer.

Populärvetenskaplig sammanfattning

Humant papillomvirus (HPV) är små dubbelsträngade DNA-virus. Hittills har över 400 genotyper identifierats. All livmoderhalscancer orsakas av högrisk HPV-typer. HPV16 är den mest framträdande. HPV16-genomet innehåller åtta gener som brett kan delas in i tidiga gener (E1, E2, E4, E5, E6 och E7) och sena gener (L1 och L2). I allmänhet kodar tidiga gener för proteiner El till E7 som främjar cellförökning och möjliggör viral DNA-replikation medan sena gener L1 och L2 kodar för strukturella proteiner som packar virusets DNA-genom för att bilda nya viruspartiklar. Alternativ splitsning kan påverka HPV16 virala proteiner genom att negativt eller positivt förändra HPV16 mRNA-splitsning. RNA-bindande proteiner, såsom serin/arginin (SR)-rik proteinfamilj och heterogen nukleär ribonukleoprotein (hnRNP) familj, är de mest framträdande aktörerna för igenkänning av splitsningsställen och bidrar till alternativ splitsningsreglering. Det finns ett stort intresse i att undersöka hur uttrycket av HPV16 tidiga och sena gener regleras av alternativ splitsning och vilka splitsningsreglerande faktorer som är inblandade.

Våra resultat visar regleringen av uttrycket av HPV16 tidiga och sena gener som huvudsakligen sker på mRNA-splitsningsnivå. Vi fann att majoriteten av alla hnRNPs kunde kontrollera HPV16 tidiga och sena geners expression och att de flesta SR-proteiner kunde reglera HPV16 sena geners expression. Vi hittade också ett cis-verkande RNA-element nedströms om HPV16 E2s splitsningsställe SA2709, som krävs för effektiv produktion av splitsat HPV16 E2-mRNA. Vi fann också att hnRNP G var den motsvarande regulatorn för detta förstärkande element av HPV16 E2-splitsat mRNA. Vi såg att hnRNP G har olika roller för att kontrollera HPV16 tidig E6/E7 och E1/E2 mRNA-bearbetning och genexpression beroende på infekterad celldifferentiering. Den huvudsakliga splitsningsfaktorn U2AF65 och DNA reparation hjälper hnRNP G att öka splitsning av HPV16 E2-mRNA efter celldifferentiering. Dessutom bidrar olika delar av hnRNP G protein till förstärkande och hämmande funktioner. I dessa ingår förstärkning av HPV16 E2-mRNA-splitsning och inhibering av HPV16 E6/E7-mRNA-splitsning av icke-överlappande hnRNP G-domäner. Förstärkningen av HPV16 E2-mRNA-splitsning och interaktionen av hnRNP G med HPV16 E2-mRNA och U2AF65 regleras därtill av en posttranslationell modifiering av hnRNP G kallad sumoylering. Detta sker vid aktivering av respons på DNA-skada. I ett separat projekt visade vi att hnRNP D kan reglera HPV16 tidiga geners E1- och E6-mRNA-uttryck genom att främja osplitsat uttryck och exportera E1- och E6-kodande mRNA till cytoplasman.

Slutligen producerade vi två subgenomiska HPV16 reporterplasmider kallade pE2sLuc och pE1Nluc som kan användas för att etablera bioanalyser som antingen övervakar uttryck av de osplitsade E1-mRNA, de splitsade E2-mRNA eller det alternativt splitsade E1C mRNA.

Sammanfattningsvis lägger våra resultat grunden och banar väg för utveckling av biomarkörer eller mål för behandling av HPV16-infektion eller cancer. Dessutom kan bioanalyserna användas för att identifiera anti-virala eller anti-tumörläkemedel mot HPV16-infektioner och HPV-driven cancer.

Abbreviations

ATM	Ataxia telangiectasia mutated
ATR	Ataxia telangiectasia and Rad3-related
AUF1	AU-rich element RNA-binding protein
BPS	Branch point sequence
CMV	Cytomegalovirus
DC	Dendritic cell
DDR	DNA damage response
ESE	Exonic splicing enhancer
ESS	Exonic splicing silencer
E6AP	E6 associated protein
EGFR	Epidermal growth factor receptor
eUTR	Early 3' untranslated region
HPV	Human papillomavirus
HR-HPVs	High risk human papillomaviruses
hnRNP	Heterogeneous nuclear ribonucleoprotein
ISS	Intronic splicing silencer
ISE	Intronic splicing enhancer
ICTY	International committee on Taxonomy of Viruses
lUTR	Late 3' untranslated region
IL-1	Interleukin-1
LR-HPVs	Low risk human papillomaviruses
LC	Langerhans cells
LCR	Long control region
MHC-I	Major histocompatibility complex

NTR	Nascent transcription region
pRb	Retinoblastoma protein
pre-mRNA	Primary transcript
PKR	Protein kinase R
PPT	Polypyrimidine tract
RRM	RNA recognition motif
RS	Arginine and serine repeats
RBD	RNA binding domain
snRNPs	Small nuclear ribonucleoproteins
SD	Splice donor (5'-splice site)
SA	Splice acceptor (3'-splice site)
SRSF	Serine and arginine-rich splicing factor
SR	Serine and arginine-rich protein
TNF- α	Tumor necrosis factor-α
STAT-1	Signal transducer and activator of transcription
U2AF	U2 auxiliary factor
VLPs	Virus like particles
TYK-2	Tyrosine kinase 2

1. Introduction

1.1 Human papillomavirus and cancer

Human papillomavirus (HPV) is a small circular double-stranded DNA virus belongs to the papillomavirus family [1]. At present, there are more than 400 HPV genotypes (types) that identified of which about 200 HPV types have been investigated and authorized by the International Committee on Taxonomy of Viruses (ICTV) [2-6]. HPVs tend to infect mucosa or skin [5-9]. Skin infected HPV types are likely transmitted by direct contact with infected individuals or surfaces containing virions. The mucosal infected HPV types are sexually transmitted [8, 10]. Moreover, the genital tract epithelia in either men or women serve as the virus reservoir [8]. Approximately 90% of individuals with HPV infections clear the infections within 1 to 2 years [6, 10-12]. Those infections are often asymptomatic and generally considered as part of normal skin microbial flora [13]. However, about 10% infections by some HPVs could induce benign lesions.

One group of HPVs exhibit a great tropism for cutaneous epithelia, such as HPV2, HPV4 and HPV7, and can cause common warts on the hands, while HPV1 causes deep plantar warts on the soles of feet [8, 14, 15]. One group of HPVs have tropism for mucosal epithelia. There are over 40 HPV types in this group infect the genital tract and induce genital warts, or precancerous lesions [16-18]. Among, such HPV types HPV6 and 11 cause benign or low-grade genital warts and are classified as ¨low risk¨ (LR)-HPVs. LR-HPV type 6 account for about 90% of genital warts [18]. Infection by LR-HPVs rarely develops into cancer and can be recognized and cleared by the host immune system [19]. In contrast, approximately fifteen HPV types including HPV16, HPV18 which can cause precancerous lesions are referred to as ¨high risk¨ (HR)-HPV types [3, 9, 20]. Infection by HR-HPVs can evade the elimination of immune response resulting in progression of high-grade cervical preneoplastic lesions after persistent infection (at least 3 years post transfection) [10, 21-24]. HR-HPVs are associated with a range of cancers and account for about 5% of all cancer worldwide **(Figure 1.1A)**, of which virtually all cervical cancers are caused by HR-HPVs [1, 25, 26].

Cervical cancer is the second most common cancer threatening the public health of middle-aged women worldwide particularly in developing countries [26-28]. Virtually all cervical cancer cases result from persistent infections with at least one

of HR-HPVs [6, 29]. Worldwide, the HR-HPV type 16 is the most prominent in patients with cervical cancer **(Figure 1.1B)** [1, 9, 24, 26, 30, 31].

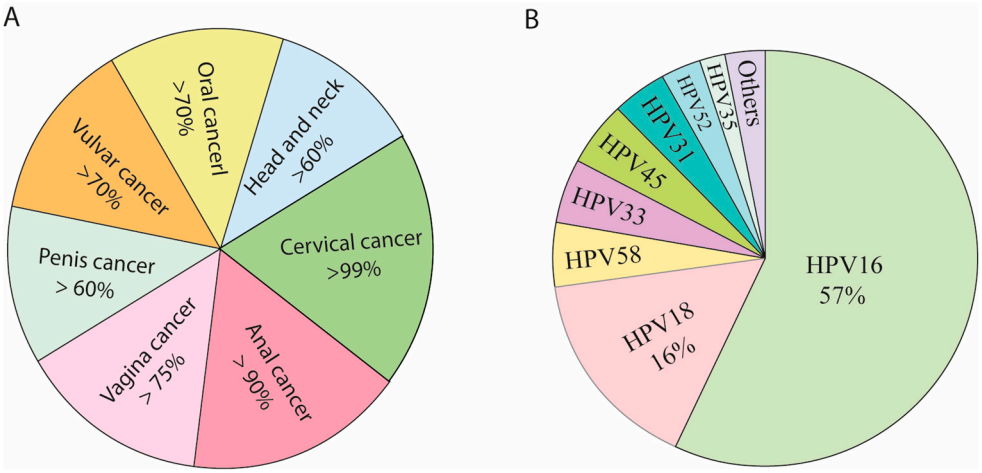


Figure 1.1 (**A**) Cancers attributed to human papillomavirus infection. (**B**) The most common HR-HPV types in cervical cancer.

1.2 HPV16 genome

The HPV virion is 55 nm in diameter and encapsulates an about 8 kb circular double-stranded DNA genome [16, 32]. All HPVs genome consists of long control region (LCR), early region and late region **(Figure 1.3)** [33, 34]. LCR, also known as upstream regulatory region (URR), is a non-protein coding region, approximately 1kb in length. It contains the viral early promoter, the replication origin, the late polyadenylation site, RNA element sequences and many transcriptional regulatory factor binding sites [16, 35, 36]. The genes in the early region encode E6, E7, E1, E2, E4 and E5 non-structural proteins which have multifunctional roles in the HPV life cycle [37]. The late genes in the late region encode L1 and L2 structural proteins that necessary for virion formation, transmission and spread [34, 38].

There are two promoters and polyadenylation signal sites in the HPV16 genome can control viral gene transcription **(Figure 1.2)**. It is the early promoter p97, late promoter P670, early polyadenylation signal site pAE and late polyadenylation signal site pAL. The early promoter p97 is located in the LCR and is activated to produce early mRNAs and polyadenylated at pAE once virus genome has been transported into the host cell nucleus. The p97 appears to be constitutively activated throughout major part of the virus replication cycle but is downregulated by E2

protein [17, 38]. The late promoter P670 is differentiated-dependent and is located in the E7 coding region. Activation of P670 resulting in high level of E1, E2, E4 and E5 mRNAs and abundant production of E1, E2, E4 and E5 proteins followed by viral capsid proteins L1 and L2. The early and late mRNAs are produced and polyadenylated at pAE and pAL, respectively. Those early proteins are important for viral genomic vegetive replication in the upper layer of epithelium **(Table 1)** [10]. The pAE is located downstream of E5 coding region and is preceded by the early 3' untranslated region (eUTR) while pAL is located downstream of L1gene is preceded by the late 3' untranslated region (lUTR). The 3' untranslated regions upstream of polyadenylation signal sites contain some RNA processing proteins binding sites and can regulate the efficiency of polyadenylation as well as the half-life of the transcripts [39].

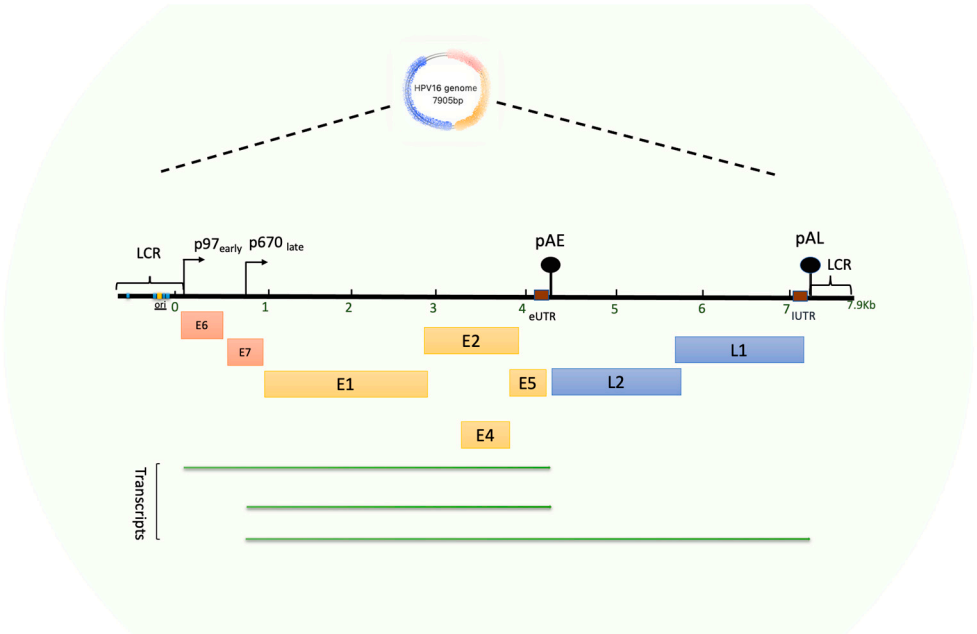


Figure 1.2. Schematic presentation of the HPV16 genome and transcripts. Ori: The replication origin. LCR: Long control region. eUTR: The early 3' untranslated region. lUTR: The late 3' untranslated region. Yellow box in the LCR represents HPV16 E1 protein binding site. Blue boxes in the LCR represent the HPV16 E2 protein binding sites.

1.3 HPV16 life cycle

The HPV life cycle is strongly linked to the differentiation process of the epithelial cell and confined to keratinocytes [14]. It is believed that the virus particles get access to bind to the basement membrane through abrasions or micro-wounds in the stratified epithelium **(Figure 1.3)** [38]. The cellular receptor of virus is now still unclear, but heparin sulphate is required and believed to facilitate the attachment of the virus [34, 40, 41]. Upon entry, virion capsid structure changes allow the transport of the viral genome into the host nucleus during the process of wound healing which can activate cell division [10]. Shortly, the virus genome is amplified in the host cell nucleus and then is maintained at a low copy number, about 20-100 copies per infected basal cell, after establishing their genome as extrachromosomal episomal DNA with the help of viral E1 and E2 proteins [2, 14, 19, 42]. As basal cells divided, the viral genomes are split and evenly distributed into two daughter cells. One set remains in the basal layer and continue to proliferate. The other migrates to upper layers [19]. Upon basal cells migrating into the upper layer of epithelium, basal cells exit the cell cycle and undergo terminal differentiation resulting in maturity and senescence [19]. However, HPVs rely on host cell DNA polymerase for viral genome amplification. To survive, HPVs produces the relatively abundant viral E6 and E7 proteins [38, 42]. In contrast, with the help of HPVs antiapoptotic E6 protein and promitotic E7 protein, differentiated HPV infected cells can survive and re-enter S-phase of the cell cycle [19, 42]. Along with the help of other viral early proteins E1, E2, E4 and E5, viral genomes in the highly differentiated infected cells in the upper stratified epithelium have a suitable cellular environment to be vegetatively amplified [14, 19, 38, 42]. Upon completion of the vegetative viral genome amplification in the upper layers, the viral two structural capsid proteins L1 and L2 are expressed and package the viral genome [19]. Thus, infectious virions are generated and shed with the dead squamous epithelium for further transmission **(Figure 1.3)**.

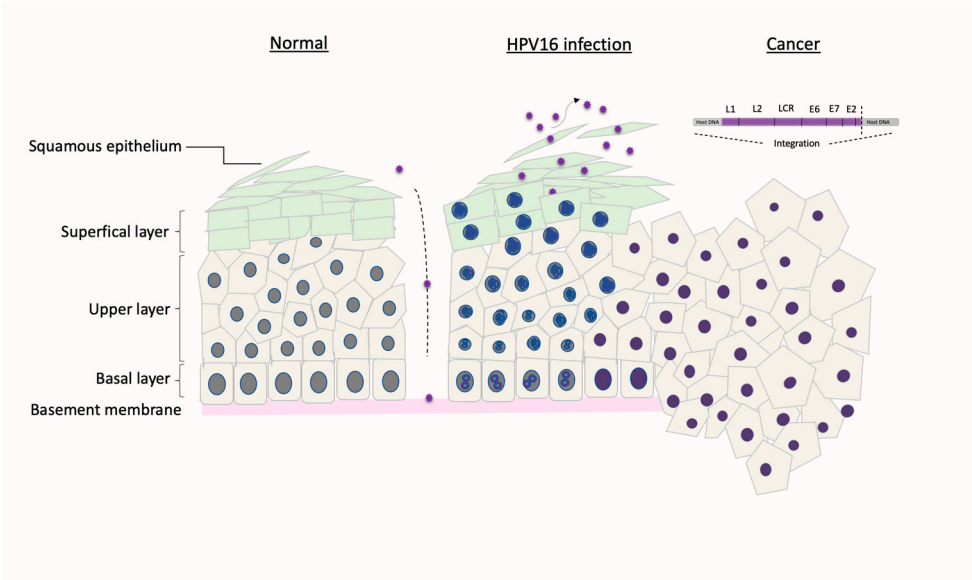


Figure 1.3. Schematic presentation of human papillomavirus life cycle.

Most HPV infections cannot induce cancer and can be cleared by immune system within 1 to 2 years after viral infection. Only 10% of infected individuals show a persistent infection and merely 1% have an increasing risk to cancer [11]. Furthermore, cancer progression is associated with long-term infection of HR-HPVs that remain untreated. In the productive viral life cycle of HR-HPVs, viral gene expression is restrained, especially the expression of oncogenes E6 and E7 [34]. In contrast, in the high-grade preneoplastic lesions or invasive cancer, the restrained expression of oncogenic E6 and E7 proteins is abolished resulting in the HR-HPV DNA integration into the host genome [43]. In the cases of HPV16 related cervical cancer, the extrachromosomal viral genome has often been found integrated into the host genome in approved 85% of the cancer cases **(Figure 1.3)** [28]. Moreover, this integration event disrupts the expression of E2 protein which can repress the virus early promoter p97 thereby alleviating the restriction of viral E6 and E7 oncogene expression. Lack of viral E2 protein results in a boost of antiapoptotic E6 protein and promitotic E7 protein and a gross accumulation of E6 and E7 oncoproteins [28, 43]. Thus, proliferation of HPV infected cells is excessively promoted, cell cycle critical checkpoints are inactivated, and genetic instability are accelerated resulting in carcinogenesis [10, 43].

1.4. HPV16 proteins

1.4.1 E6 and E7

HPV16 E6 and E7 proteins are expressed under the control of the viral early promoter p97. They are major transforming proteins and well-known as oncoproteins since an abundant expression level of both proteins are found in the cancer progressive lesions [44, 45]. E6 and E7 are multifunctional proteins and play crucial and irreplaceable roles in the early stage of HPV infected cells. E6 is initially expressed at a low level in the basal or lower epithelial layer. E6 protein is comprised of approximately 150 amino acids and is well-known for its ability to regulate the destruction of p53. p53 is the most well investigated tumour suppressor protein till date which arrests cell growth and induces apoptosis [46]. The E7 protein is smaller, contains about 100 amino acids and is also initially expressed at low level in the basal layer of epithelium. The most significant feature of E7 is the ability to target and destabilize the tumor suppressor retinoblastoma (Rb) protein. pRb is a negative regulator of the cell cycle that binds to the cellular transcriptional factor E2F family and represses the function of E2F [8, 19].

As the infected cell exits cell cycle after moving to upper layer and undergoes the terminal differentiation, the HPV dependent cellular DNA replicative machinery from host cell is downregulated. To survive, the virus expresses E6 and E7 proteins resulting in abolishment of the restrained cell cycle and postponement of the cell terminal differentiation [19, 38, 42]. E6 and E7 oncoproteins conduct their effects by regulating the activity of p53 and pRb [38, 45].

The association of E7 with pRb is well characterized [8, 19, 47]. E7 targets pRb and induces ubiquitinylated degradation of pRb **(Figure 1.4)** resulting in the release of E2F family of transcription factor [8, 19]. The activation of E2F transcriptional factor family induces a large expression of necessary replicatory proteins to restore the DNA synthesis and then forces the transition of cell cycle progression from G1 to S phase. Thus, HPV genome can continue replication [8, 19, 38, 46]. HR-HPVs E7 protein binds to pRb stronger than LR-HPVs [8, 10]. As the consequence of dysregulated DNA synthesis, the activation of p53 is highly upregulated to clear abnormal cells by inducing cell apoptosis. To counteract this apoptosis induced by p53 and keep HPV infected cells alive, E6 protein expression is needed [38]. Viral E6 protein binds to cellular E6 associated protein (E6AP) to form E6-E6AP complex that targets p53 for ubiquitination and degradation of p53 in proteasome **(Figure 1.4)**, thereby rescuing the cell from p53-mediated growth suppression and apoptosis [19]. In a comparison, the LR-HPVs E6 protein can bind to E6AP but does not mediate the degradation of p53 [8, 10, 48]. In a summary, the viral E6 and E7 proteins work together to promote cell proliferation and prolong cell cycle

progression, thus creating a suitable intracellular environment for continuous replication of the viral DNA [38].

In addition to p53 degradation, HPV16 E6 protein has been shown to significantly induce telomerase activity in primary epithelial cells [45, 49-53]. Enhanced telomerase activity is related to prevent the shortening of telomeres during DNA replication and associated to chromosomal stability [54, 55]. This activation is suggested to be mediated by upregulation of human telomerase reverse transcriptase (hTERT) transcription by E6 protein or by the targeting of a cellular protein such as a telomerase inhibitor, for ubiquitin-mediated degradation by E6/E6-AP complex [45, 50, 51].

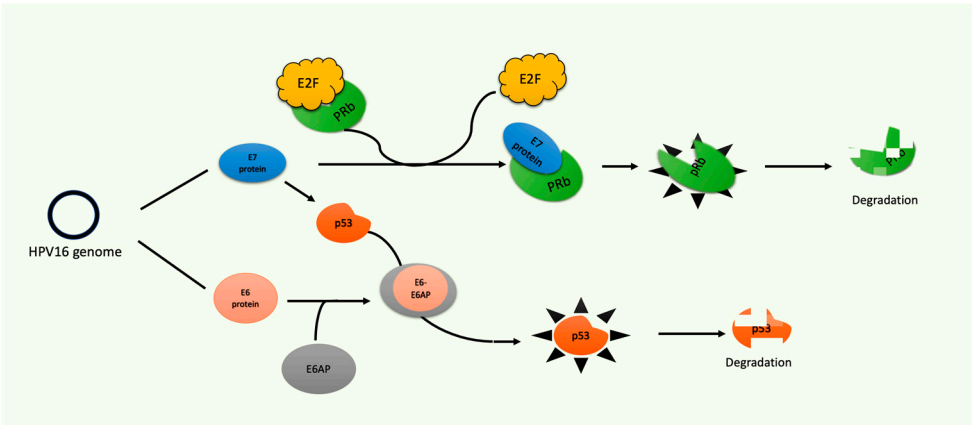


Figure 1.4 Schematical presentation of HPV16 E6 protein mediated p53 and HPV16 E7 protein mediated pRb

1.4.2 E1 and E2

The HPV E1 and E2 proteins are essential regulatory proteins required for viral DNA replication in the HPV16 life cycle, either in the basal layer or upper layer of epithelium. Both proteins are expressed in the early infected stage and distributed in both nuclear and cytoplasmic fractions [56, 57]. E1 is a unique protein in HPVs. It is an ATP-dependent DNA helicase, the only viral protein with enzymatic activity [58]. It is encoded from the largest and most conserved gene in papillomaviruses, reflecting the essential role in the viral life cycle. E1 contains three functional domains, the N-terminal regulatory domain, a DNA-binding domain which recognizes the origin in the LCR, and a C-terminal ATP-dependent helicase domain **(Figure 1.5)**. The main function of E1 protein is as the primary replication protein contributing to initial replication of viral genome. E1 recognizes and binds specifically to the viral DNA replication origin located at 3' end of the viral genome

LCR with the aid of E2 protein which has three binding sites flanking the E1 binding site **(Figure 1.5)** [57]. E1 and E2 bind to these sites as a heterodimer. E2 loads E1 helicase onto the origin, then E2 is displaced as E1 turns to a double hexamer in the presence of ATP [57]. Then, the viral DNA is synthesized after recruitment of the host DNA replication machinery [59]. Both E1 and E2 proteins are essential for initial replication of the viral genome and for the viral genomic maintenance and vegetative amplification.

The viral E2 protein is a specific DNA-binding protein, consisting of an N-terminal domain generally called transactivation domain, a flexible serine-arginine rich region often called hinge domain and a C-terminal DNA-binding/dimerization domain **(Figure 1.5)** [57, 60]. The E2 DNA binding domain and transactivation domain are indirectly contributing to tethering viral genome to the host mitotic chromosome to facilitate accurate segregation of viral genomes in dividing cells [57]. E2 is a multifunctional protein that regulates many virus activities. In addition to replication function, the main function of E2 protein is related to viral transcription. As the main transcriptional factor, E2 protein can activate or repress the viral transcription depending on its binding sequences or associated cellular factors [57]. E2 can repress the viral early promoter and play a vital role in HPV pathogenesis by strictly controlling the expression of E6 and E7 oncogenes. Moreover, mounting evidence show that E2 can induce cell growth arrest and apoptosis associated with p53 [57]. However, this E2-mediated repression of E6 and E7 expression will be disrupted upon viral genome integration into host chromosomes [43, 57].

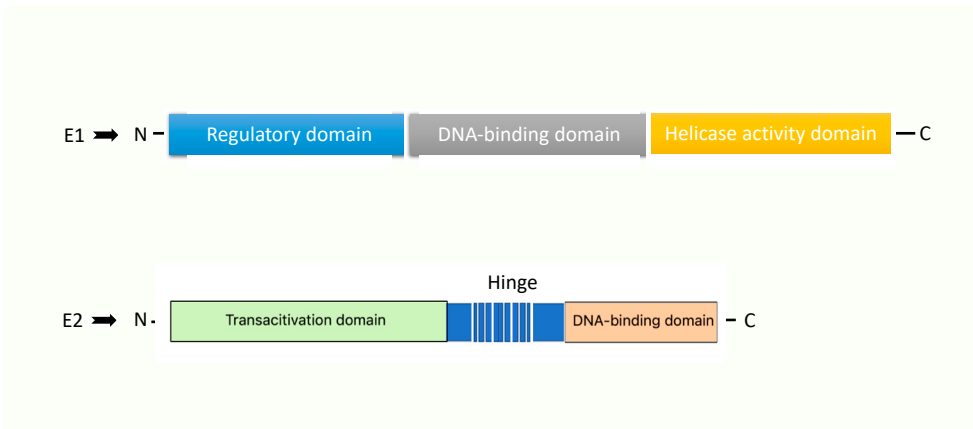


Figure 1.5 Schematical representation of domain structures of HPV16 E1 and E2 proteins.

1.4.3 E4 and E5

Viral proteins E4 and E5 are contributing to the viral genome replication as well [10]. E4 is one of the most abundantly expressed protein in the upper layer of epithelium, also known as E1^E4 late protein. The N-terminal part contains five amino acids from E1 open reading frame which provide the start codon for E4 translation [14]. E4 protein can arrest cell cycle in G2 phase and affect the efficiency of viral genome replication by destroying the cyclin dependent kinase 1(CDK1)-cyclin B which associated with the initiation of mitosis in the cytoplasm [61]. Except the role in viral genome replication, the abundant E4 protein has been suggested to enhance virus synthesis, affect virus infectivity, as well as to promote virus release from the upper layer of epithelium by interfering with the integrity of keratin and the normal assembly of cornified envelope [10, 19].

E5 protein is a small membrane-associated hydrophobic protein and expressed in both basal and suprabasal layers [62]. The expression levels of E5 during virus infection is very low which is an obstacle to elucidation of E5 various roles during virus life cycle [38, 63]. E5 protein can interfere with the presence of peptides on the major histocompatibility complex (MHC). It is also believed that E5 protein plays an important role in stimulating cell growth by augmenting the epidermal growth factor receptor (EGFR) signalling [10, 14] a pathway to inhibit keratinocytes differentiation [44]. The activation of EFFR by E5 protein leads to stimulate cell proliferation [64].

1.4.4 L1 and L2

Viral L1 and L2 proteins are coded by L1 and L2 genes located in the late region of viral genome [65]. Structural L1 and L2 proteins are only expressed in the terminally differentiated superficial layer of epithelium [66]. Transcription of late L1 and L2 genes is initiated from the HPV16 late promoter and polyadenylate at late polyadenylation site pAL Under the control of late promoter, L2 protein expresses prior to L1 protein. Viral L1 and L2 are two capsid proteins which package the circular double-stranded DNA viral genome to form virus particles. The viral L2 protein is the minor capsid protein while the L1 protein is the major capsid protein since one capsid of the virus particle contains 360 copies of L1 protein but only 12-72 copies of L2 protein [19, 67-69].

The L1 protein largely contributes to virus entry into host cell by binding heparin sulphate proteoglycans in the epithelial basement membrane and retains in the endosome following degradation by lysosomes [10, 68]. The viral L1 protein can spontaneously self-assemble into capsid structures without L2 and viral genome to generates virus like particles (VLPs) which are used as HPV vaccines [68, 70-72].

The L2 minor capsid protein contributes to inducing capsid structure changes when the virion binds to cells and disrupt the endosomal membrane to ensure the released

viral genome correctly enter nucleus [10, 19, 67, 69]. Furthermore, the L2 protein is suggested to enhance the assembly of virus particles and infectivity although it is dispensable for the capsid formation [19].

1.5 Immune evasion

The ability to evade immune system attack is one remarkable features of HPVs [73-76]. This hallmark evasion capacity results in a low-profile and no inflammation-associated life cycle of HPVs [77,78]. The HPV life cycle is non-lytic and exclusively intraepithelial [29, 77, 78]. Consequently, there is no viraemia, no cell damage or cell death caused by viral-induced cytolysis [10, 48]. During the virus replication in the early stage, the non-secreted early proteins are expressed in low levels resulting in insufficient triggering of host immune responses [77]. Moreover, the immunogenic proteins L1 and L2 are only produced in the superficial epithelial layer of epithelium compromising detection by immune system. Furthermore, the infected cornified cells containing progeny virions are shed from the epithelial surface thereby avoiding triggering immune cells [10, 19, 48, 77].

HPVs have evolved numerous ways to suppress the immune system [73, 76, 77]. HPV positive keratinocytes show significantly downregulated innate immune system signalling pathways resulting in no or little release of interferons, particularly type-1 interferon (IFN-1) and reduce levels of inflammatory cytokines including interleukin-1 (IL-1), IL-6 and tumour necrosis factor-α (TNF- α) [10, 48, 77]. This reduced production of cytokines alleviates the ability of immune cells to infiltrate infected epithelium and prevent migration of inflammatory cells such as Langerhans cells (LC) the major dendritic cell (DC) into the stratified squamous epithelium [10, 48, 77]. Some HPV16 proteins produced within the infected cells are suggested to contribute to repression of the innate immune response [79]. Viral E6 and E7 proteins can interfere with and repress many interferons induced transcript factors. E6 and E7 proteins repress the activation of the signal transducer and activator of transcription 1 (STAT-1) which activate the transcription of interferon-induced genes of keratinocytes. HPV16 E6 protein binds to and represses the activity of interleukin response factor (IRF) -3, tyrosine kinase 2 (TYK-2) as well as protein kinase R (PKR), while E7 protein is associated with IRF-1. The host cell-mediated immune response is downregulated by HPV16, resulting in inadequate presentation on MHC-I [10, 19, 48, 77]. Viral E5 protein can interfere with the process of MHC peptide present on cell surface. Viral E6 and E7 proteins may also affect this process [10, 48]. For example, viral E7 interferes with the MHC-I pathway and blocks the presentation of antigens via inactivation of the transporter associated with antigen processing (TAP) protein [48, 80]. Moreover, the HPV16 capsid L2 protein has the capacity to affect the ability of LC to induce T-cell responses by repressing maturation, migration and secretion of cytokines [10, 48].

Despite such strategies of HPVs to evade elimination by the immune system, 90% of HPV infections are cleared within one or two years [10]. It is important to understand HPV immune evasion strategies to improve immunotherapeutic approaches. Unfortunately, so far, the strategies of HPV immune evasion still not less clear [76].

Table 1. Functions of HPV16 proteins

Protein	Functions
E1	Binds to replication origin, initiates DNA replication, DNA helicase.
E2	Control of DNA replication, transcription, HPV genome segregation. Enhances binding of E1 to the replication origin. Transcription factor of E6 and E7gene expression. Control of early region viral gene expression. Cell cycle and apoptosis regulation.
E4	Most abundantly expressed protein. Disruption of cytokeratin network. Cell cycle arrest. Virion assembly and release.
E5	Small, membrane-associated protein. Enhances growth factor responses in infected cells, stimulates cell proliferation. Contributes to immune evasion.
E6	Oncoprotein. Directs degradation of p53, inhibits apoptosis and differentiation. Inhibits cell cycle blocks. Involved in immune evasion.
E7	Oncoprotein. Control of Cell cycle. Binds and degrades pRb, permitting cell progression to S phase of cell cycle. Contributes to immune evasion.
L1	Major capsid protein. Required for the attachment to cell surface receptors. Highly immunogenic. Self assembles into capsid structures.
L2	Minor capsid protein. Contributes to binding of virion to the cell receptor, favoring uptake, transport to the nucleus and delivery of viral DNA to replication centers. Virus assembly. Contributes to immune evasion.

1.6 Regulation of HPV16 gene expression

The viral proteins of HPV16 play critical roles in the process of infection **(Table 1)**. To some extent, viral proteins control the virus destiny, to be cleared by immune system or to persistent and to develop into intraepithelial neoplasia that nay progress to cancer. Understanding the way how those proteins gene expressions is regulated

is one key approach to uncover targets for drug development for HPV16-induced cancer.

Gene expression consists of many steps and virtually each step can be regulated, from transcription initiation to posttranscriptional modifications or RNA processing and posttranslational modification of regulatory factors [60, 81]. RNA processing is a set of biological processes by which the newly made primary transcript is altered through the addition of a 5' cap, the addition of a 3' poly A tail and RNA splicing to form a mature, functional mRNA which can leave the nucleus to perform its various roles in the cells [60, 82, 83]. These processes are critical for the correct eukaryotic mRNA translation. The add a of 5' cap and a 3' poly A tail prevent RNA degradation and facilitate the transport of mature mRNA to the cytoplasmic ribosome, while RNA splicing gathers protein-coding exons and removes introns from primary transcripts [84-86]. When it comes to the control of HPV gene expression, RNA processing, particularly RNA splicing, is one of the most common and well-studied controlled steps in the HPV gene regulation in the past decade [23, 87-90].

1.6.1 RNA splicing

mRNA splicing is a basic but crucial cellular RNA processing by which newly formed primary transcripts (pre-mRNAs) are transformed into a mature mRNA [91]. Splicing is carried out by the spliceosome, a macromolecular ribonucleoprotein complex, including small nuclear ribonucleoproteins (snRNPs) U1, U2, U4, U5 and U6 **(Figure 1.6A)** [92-94]. The spliceosome recognizes and binds to splice sites to remove non-coding introns and to gather exons [94-96]. The boundaries of introns are known as 5' splice sites, also called splice donors (SD), and 3'splice sites, also called splice acceptors (SA). The 5' splice site contains a GU dinucleotide where the intron starts. The 5' splice site can be specifically recognized and bound to the U1 snRNP as the result of the sequence complementarity between the splice site and the U1 snRNA [97-99]. The 3' splice site can be further divided into three parts, the branch point sequence (BPS) containing a conserved adenosine residue, the polypyrimidine tract (PPT), and the AG dinucleotide that mark the end of the intron [97]. The 3' splice site can be recognized and bound to the U2 snRNP along with its auxiliary factor U2AF65 as well as U2AF35, the heterodimer partner of U2AF65 **(Figure 1.6A)** [97, 100]. U2AF35 recognizes the AG dinucleotide, while U2AF65 recognizes the PPT and modulates the binding proteins to the BPS such as splicing factor-1 (SF-1) during the initial stage of spliceosome assembly, and U2 snRNP during later steps of the spliceosome assembly [97, 100].

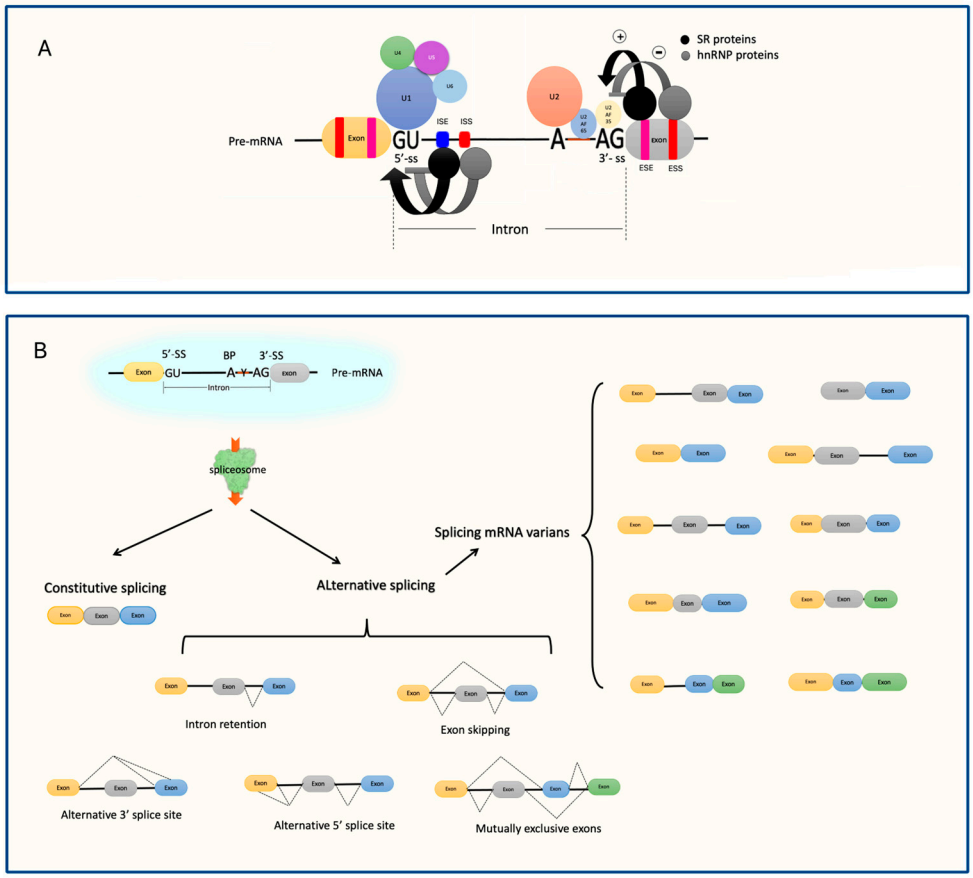


Figure 1.6 (**A**) Schematical representation of the recognition and regulation of RNA splicing site by spliceosome and regulatory proteins. (**B**) Schematical representation of pattern of mRNA splicing.

mRNA splicing can be classified as either constitutive splicing or alternative splicing **(Figure 1.6B)**. Constitutive splicing is when all introns are removed from the pre-mRNA and all protein-coding exons are brought together to form the mature mRNA [101]. Alternative splicing is when intron retention, exon skipping or exon size alteration (selection of splice sites) can occur and result in production of a range of different mRNA variants derived from one pre-mRNA **(Figure 1.6B)** [17, 23, 102, 103]. Alternative splicing plays a key role in biological diversity of proteins by greatly contributing to maximizing protein diversity [23, 91, 103-105]. All papillomaviruses use alternative mRNA polyadenylation and splicing extensively to control gene expression [38, 85, 97, 105]. HPV16 proteins are produced from a myriad of alternatively spliced and polyadenylated mRNAs **(Figure 1.7)** [38, 42, 87, 105].

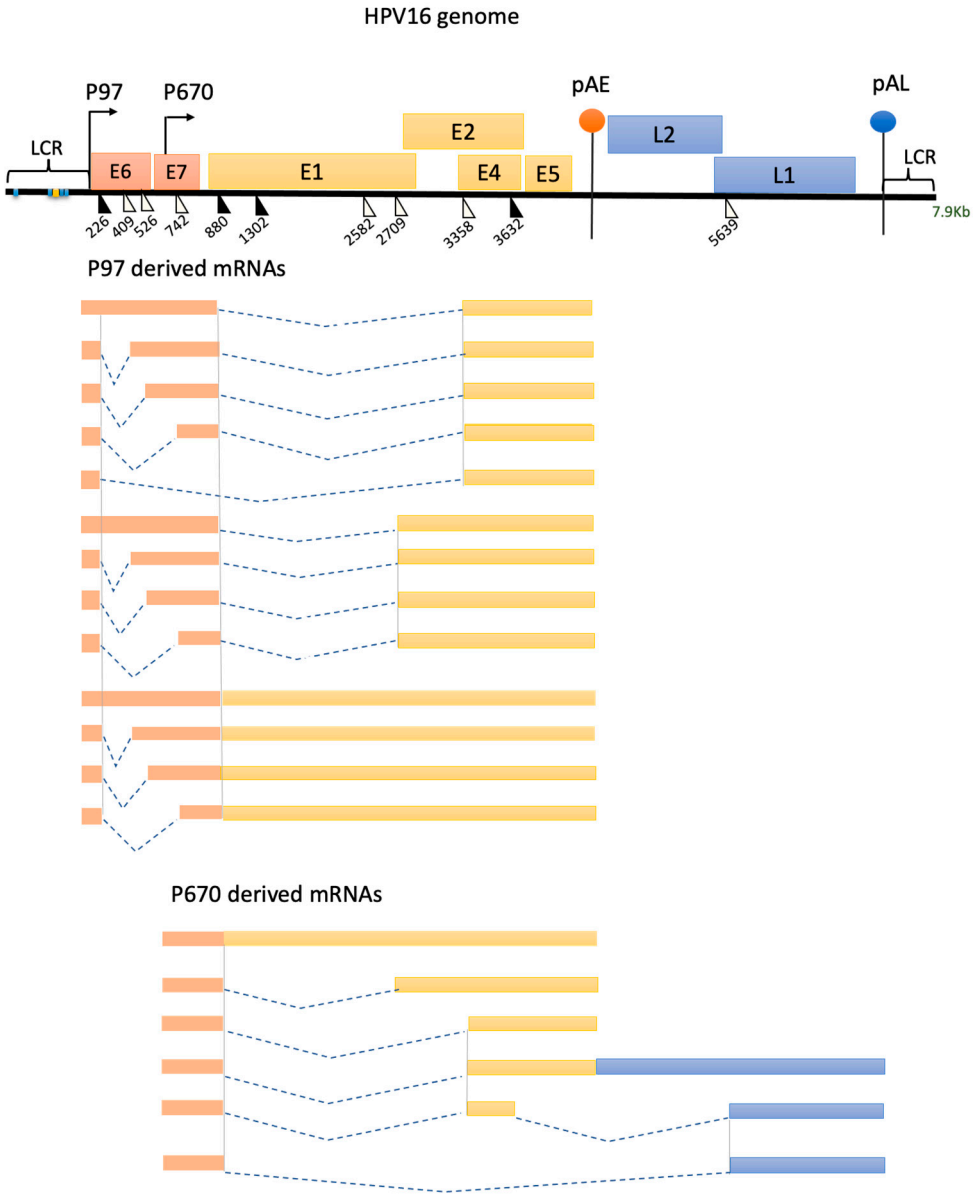


Figure 1.7 Schematical map of linearized HPV16 genome and representative HPV16 mRNAs generated by alternative splicing.

1.6.2 Control of mRNA alternative splicing

mRNA splicing is performed and catalysed by components of the spliceosome. The spliceosome is necessary for splicing but is not sufficient [23]. Importantly, not all the splice sites are well recognized by components of the spliceosome. Accurate and efficient mRNA splicing is also modulated by various splicing regulatory factors, such as cis-acting regulatory elements and trans-acting factors [106]. The trans-acting factors interact with cis-acting regulatory elements control splicing [107]. The effect of trans-acting proteins on mRNA splicing, however, cannot be roughly classified as up- or down- regulation, owing to their dependence of specific cis-acting regulatory sequences in the pre-mRNA. One splicing regulatory protein that bind to one specific RNA sequence may serve as an enhancing factor but show repressive ability when it binds to a similar sequence in another position of the pre-mRNA, and vice versa [108].

1.6.2.1 Cis-acting element

Alternative splicing is exerted by the recognition of the pair of 5' and 3' splice sites which are under the competition with at least one other 5' splice site or 3' splice site [109]. The strength of 5' and 3' splice sites therefore determine the efficiency of recognition of the splice site and then governs which kind of mRNA that is produced [82, 104]. Cis-acting elements are RNA sequences that are located close to the splice sites both in pre-mRNA exons and introns and have the capacity to alter the strength of 5' and 3'splice site resulting in enhanced splicing or reduced splicing [106, 110]. Cis-acting elements are therefore conventionally classified as splicing enhancers or splicing silencers [108]. Splicing enhancer is an RNA sequence that splicing factors can bind and increase the possibility of recognition of a nearby splice site. Splicing enhancers are separated into intronic splicing enhancers (ISE) or exon splicing enhancers (ESE) due to the location **(Figure 1.6A)** [111, 112]. In contrast, a splicing silencer is an RNA sequence to which splicing negative regulatory proteins can bind. Based on its position on the pre-mRNA, a splicing silencer can also be divided into two groups: intronic splicing silencers (ISS) and exon splicing silencers (ESS) **(Figure 1.6A)** [111].

1.6.2.2 Trans-acting factor

The protein or complex that binds to the cis-acting RNA sequence on the pre-mRNA and regulates mRNA splicing is termed trans-acting factor [107]. A trans-acting factor binding to its corresponding cis-acting element facilitates the recruitment of specific components of the spliceosome to a nearby splice site, therefore regulating splicing **(Figure 1.6A)** [113]. A large number of cellular RNA-binding protein serve as trans-acting regulatory factors can stimulate or repress mRNA splicing, termed activators or repressors, respectively [114-118]. The most well characterized and predominate splicing regulatory RNA-binding proteins are serine-arginine rich (SR) proteins and heterogeneous nuclear ribonucleoproteins (hnRNPs) [119-124].

1.6.3 SR protein

SR proteins are conserved in eukaryotes and are involved in facilitating splice site recognition by components of the spliceosome [110, 122, 125, 126]. In humans, there are about 12 SR proteins identified based on the N-terminal RNA recognition motifs (RRMs) and the C-terminal RS domain which contains more than 40% uninterrupted RS or SR repeats **(Figure 1.8)** [82]. The proteins are also called serine and arginine-rich splicing factors (SRSF) [82]. SR proteins primarily localize to the nucleus, but some have the ability to shuttle from nucleus to cytoplasm [23, 111]. In general, SR proteins act through their C-terminal RS domain to control the formation of early splice complex assembly by protein-protein interactions and through their N-terminal RRM thereby binding to cis-acting RNA elements in the pre-mRNA [23, 111]. SR proteins can interact with U2AF35 allowing U2 snRNP to stably bind to the intronic 3' splice site, greatly increasing the possibility of utilization of the nearby splice site [23, 126]. This function of SR proteins is particularly important for weak or poorly conserved splice sites [23]. The majority of SR proteins are activators of mRNA splicing by binding to ESEs, but some have been shown to have the ability to repress splicing, such as SRSF-9. SRSF-9, also called SRp30c, binds to an ISS to repress the recognition of a 3' splice site [128].

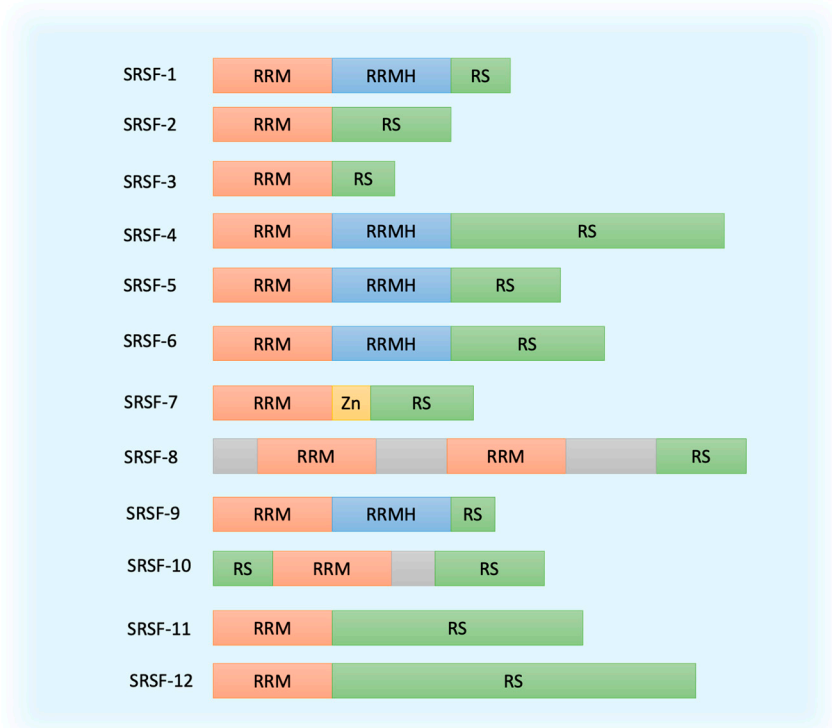


Figure 1.8 Schematic representation of human SR proteins.

1.6.4 Heterogeneous nuclear ribonucleoproteins (hnRNPs)

Heterogeneous nuclear ribonucleoproteins (hnRNPs) are a set of abundant nuclear RNA-binding proteins with distinct functions in eukaryotes [129-131]. In humans, there are over 20 types of hnRNPs identified and named alphabetically from A to U [124, 132]. The molecular weight of hnRNPs varies from 34 to 120 kDa [129]. hnRNPs are known to be modular proteins and primarily affect RNA metabolism [97, 129, 133]. However, the effect of hnRNPs on RNA metabolism is different and depends on the variety of hnRNPs domain structures **(Figure 1.9)** [97, 129, 133].

hnRNPs have distinct functions at both transcriptional and posttranscriptional levels, ranging from packaging nascent transcripts to preventing degradation, to transcriptional modulation, to regulation of alternative splicing, to transcript stability, to control of nucleocytoplasmic mRNA transport [124, 129, 134, 135]. Some hnRNPs control alternative splicing through ISS or ESS to regulate competition between splice sites. Some of hnRNPs regulate the recruitment of specific components of the spliceosome [97]. Some hnRNPs are subjected to the post translational modifications, such as phosphorylation, methylation, acetylation and sumoylation [97, 129]. It is believed that the function of posttranslational modifications is likely to modulate the interacts between hnRNPs and RNA or other proteins which could be the partners or controllers of hnRNPs in cells. Posttranslational modifications are also suggested to regulate the ability of hnRNPs shuttle from nucleus to cytoplasm or mediate other subcellular localization [129].

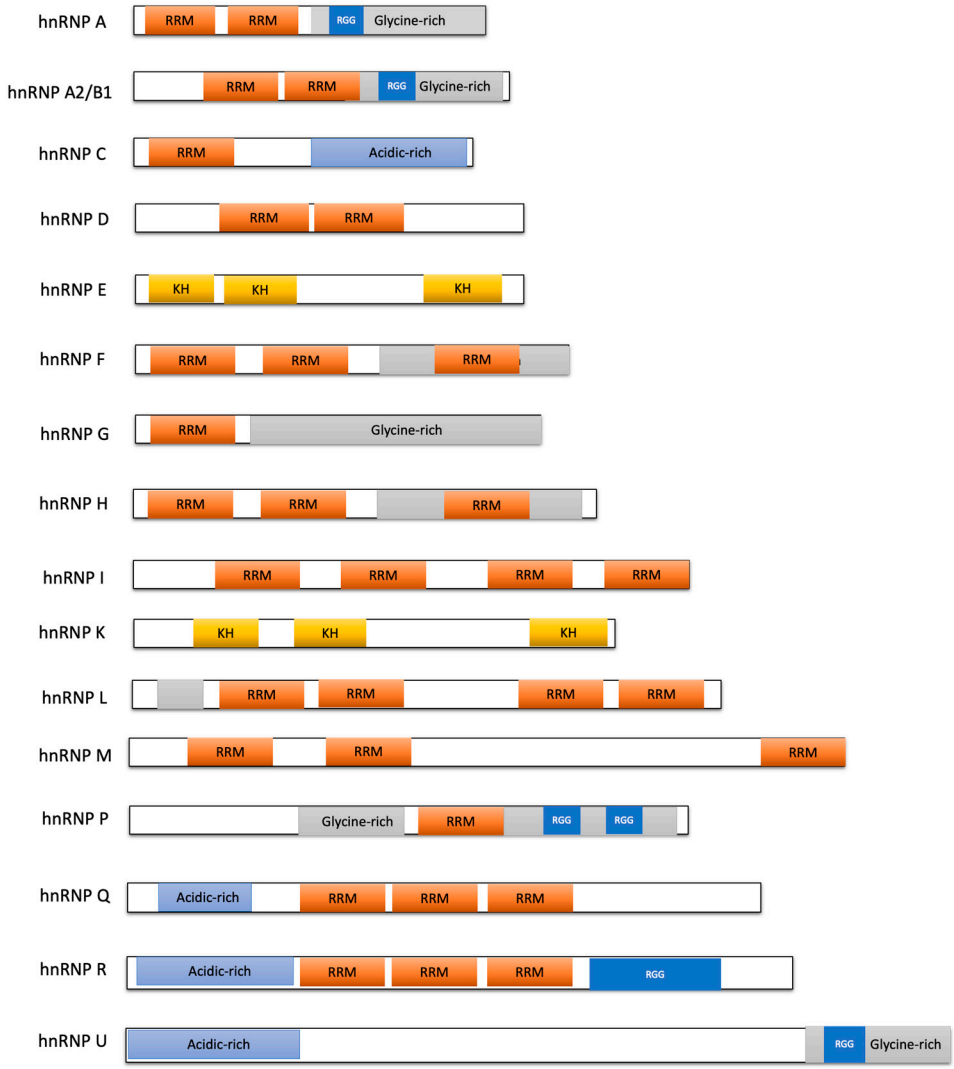


Figure 1.9 Schematic representation of human hnRNPs.

1.6.4.1 hnRNP D

hnRNP D, also known as AU-rich element RNA-binding protein (AUF1), has the capacity to shuttle between the nucleus and cytoplasm [136]. Cloning and characterization of hnRNP D cDNAs indicated that hnRNP D gene is transcribed into a pre-mRNA that undergoes alternative pre-mRNA splicing to give rise to four different protein isoforms with apparent molecular weights of 37, 40, 42, and 45 kDa [137]. The inclusion or exclusion of exon 2 and/or exon 7, near the N- and C-termini of hnRNP D respectively, is responsible for the differences between the

isoforms **(Figure 1.10)** [138]. The four proteins of hnRNP D, are named according to their apparent molecular weights. The hnRNP D45 isoform contains sequences encoded by both exon 2 and exon 7. hnRNP D42 retains the exon 7 encoded domain, hnRNP D40 retains the exon 2 encoded domain, while hnRNP D37 lacks both exons. All four protein isoforms contain two RRM domains, an alanine (A)-rich domain located in the N-terminus, and an 8-amino acid glutamine (Q)-rich motif located in the C-terminus followed by Arg-Gly-Gly (RGG) motifs [138, 139]. The RRM domains are required but not sufficient for high-affinity RNA binding [140]. hnRNP D is involved in multiple biological processes. It can work as a transcription factor or regulatory protein to regulate mRNA decay, or mRNA stabilization and translation [141,142].

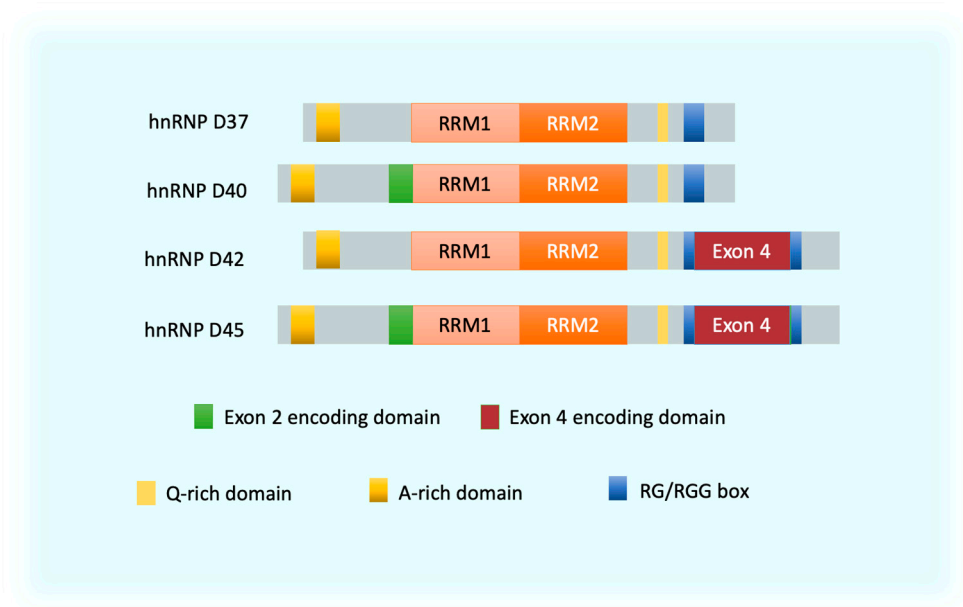


Figure 1.10 Schematic representation of the four AUF1/hnRNP-D isoforms.

1.6.4.2 hnRNP G

hnRNP G belongs to the hnRNP family and is also known as RNA-binding motif protein X-linked (RBMX). It is the first hnRNP protein that was shown to be glycosylated which was found to be necessary for protein-protein interactions [143]. hnRNP G is a ubiquitously expressed but tissue-specific nuclear RNA-binding protein that is encoded by the RBMX gene located on the X chromosome [144]. hnRNP G is composed of 391 amino acids and has a molecular weight of 43 kDa [145]. It is composed of an N-terminal RRM followed by a proline-rich motif, an RGG box with three Arg–Gly–Gly repeats, a nascent RNA targeting domain (NTD),

an SRGY box, Arg–Ser (RS) repeats and a C-terminal, second RNA-binding domain (C-RBD) **(Figure 1.11)** [146]. hnRNP G is a component of the supraspliceosome but can modulate multiple cellular processes [147,148]. It regulates selection of alternative splice sites and facilitates splicing of pre-mRNAs [149, 150]. In addition, it also functions as a transcription factor which may inhibit or activate gene transcription in the different cellular environment [151, 152]. Moreover, hnRNP G has been identified as a protein involved in DNA damage response [148, 153, 154].

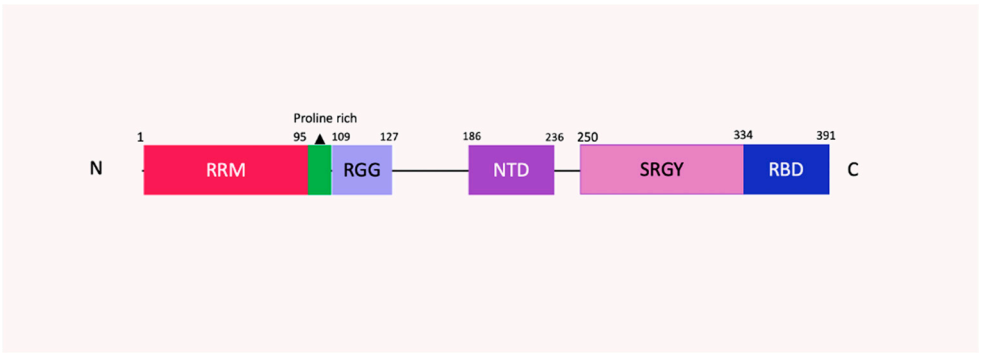


Figure 1.11 Schematic representation of hnRNP G.

1.7 Control of HPV16 mRNA alternative splicing

Upon HPV genome entry into the nucleus, the transcription initiates from the early promoter P97 and produces polycistronic mRNAs [38, 42, 87, 155]. These polycistronic mRNAs undergo alternative splicing to generate a range of different mRNAs. The alternative splicing is of paramount importance for HPV16 gene expression since it can generate at least one mRNA that ensures efficient expression of each protein. Alternative splicing of HPV16 early polycistronic mRNAs differentially utilizes a number of HPV16 splice sites including SD226, SA409, SA526, SA742, SD880, SA2582, SA2709 and SA3358 to generate a range of mRNA variants, all polyadenylated at the early polyadenylation signal pAE. These mRNAs will be translated into the early proteins E6, E7, E1, E2, E4 and E5 **(Figure 1.12)** [38, 42, 87, 155-157].

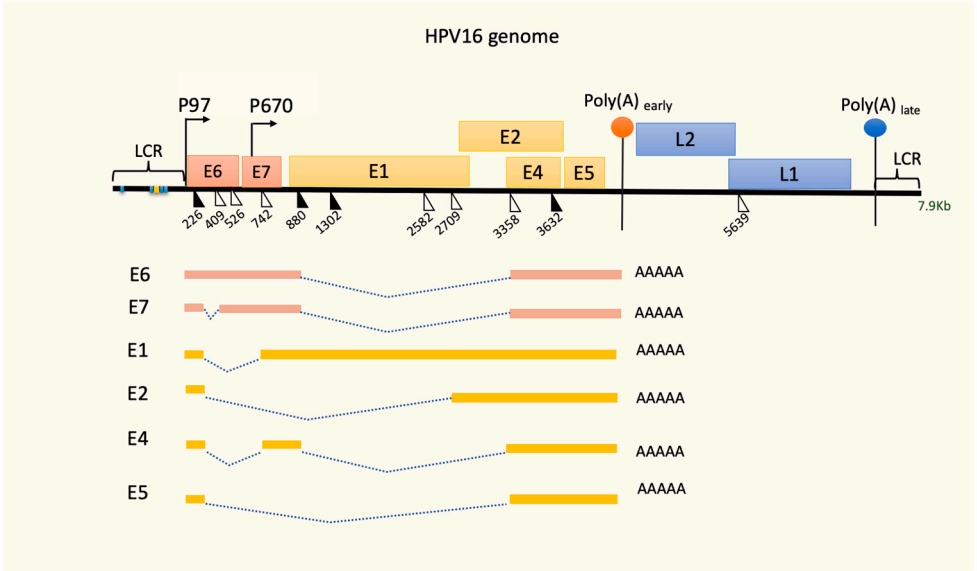


Figure 1.12 Schematic representation of the HPV16 genome and representative early mRNAs.

The E6 and E7 mRNAs are only transcribed from P97 and share the same pre-mRNA [38, 42, 155]. This pre-mRNA is exclusively alternatively spliced by utilizing the 5' splice site SD226 and either of the 3' splice site SA409, SA526 or SA742 to produce at least four spliced mRNA variants **(Figure 1.13)** [155-157]. Intron retention between SD226 and SA409 generates unspliced E6 mRNA or retained E6 mRNA (E6-IR), encoding full-length E6 protein [42]. Excision of the intron between SD226 and SA409 produces the spliced 226^409 mRNA. This transcript can be translated into the E6*1 protein, which is a splicing maimed E6 protein, but may also produce full-length E7 protein [42]. Selection of SA526 by spliceosome produces the spliced 226^526 mRNA encoding E6*II or E7 protein and spliced 226^742 mRNA encoding E6-E7 fusion protein named E6^E7 **(Figure 1.13)** [42, 158-161]. Among these spliced mRNA variants, the 226^409 mRNA is the most abundant due to the utilization efficiency of the E6/E7 splice sites [42]. Consequently, E6/E7 mRNA splicing is tightly controlled by cis-acting RNA elements and trans-acting regulators [38, 42]. For example, cellular RNA-binding protein hnRNPA1 can bind directly to an ESS near SA409 to repress production of E6*1 encoding mRNA 226^409 thereby promoting production of unspliced intron retained E6 mRNA, while hnRNP A2 binds the same ESS and redirects splicing to SA742 instead of SA409 or SA526 thereby allowing production of spliced 226^742 mRNA [159].

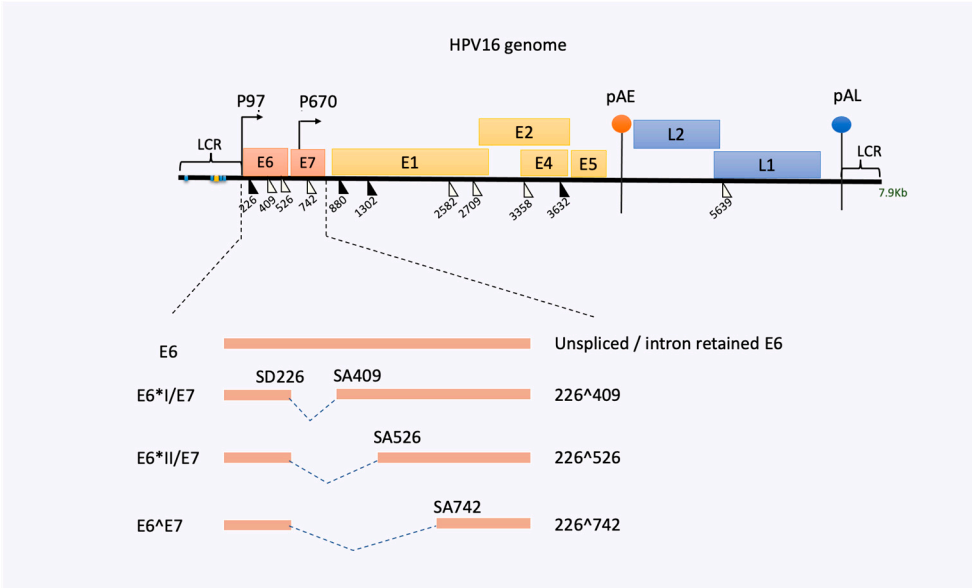


Figure 1.13 Schematic representation of the HPV16 E6 and E7 alternatively spliced mRNAs

Upon cell differentiation, the late promoter P670 is activated. Thus, the promoter is switched from early promoter P97 to late promoter P670 [38, 42, 87, 155]. Transcription from P670 produces abundant mRNAs encoding E1, E2 and E4 using primarily SD880, SA2709 and SA3358 and polyadenylated at pAE **(Figure 1.14)** [155]. There are three splice donors SD880, SD1302 and SD3632 and three splice acceptors SA2582, SA2709 and SA3358 in the E1/E2 coding region **(Figure 1.14)**. Among those splice sites, intron retention between SD880 and SA2709 produce unspliced E1 mRNA or intron retained E1 mRNA, while efficiently utilized SD880 and either SA2709 or SA3358 produce spliced E2 mRNA or E4/E1^E4 mRNAs, respectively [38, 42, 87, 155]. These mRNAs are polyadenylated at pAE. Moreover, SA3358 is utilized within the entire HPV16 life cycle to produce both early and late mRNAs polyadenylated at pAE and late mRNAs encoding L1 and L2 proteins which are polyadenylated at pAL **(Figure 1.12 and 1.14)** [42, 87, 155].

In fact, splicing acceptors SA2582 and SA2709 which are located upstream of the E2 ATG within the E1 open reading frame may both contribute to E2 mRNA production. However, SA2709 is verified to be more efficient than SA2582 and to generate spliced E2 mRNA that is proved efficiently producing E2 protein [162]. The major splice acceptor SA2709 is also competing against SA3358 that is are utilized in a mutually exclusive manner [42, 155]. Despite the important roles of SD880 and SA2709 in HPV16 mRNA production, regulatory factors that control

these spice sites have not been reported [42]. In contrast, as one of most efficiently or frequently used splice acceptor, SA3358 has been well studied and numerous reports show that SA3358 is regulated by downstream cis-acting RNA elements along with RNA-binding proteins including SRSF1, SRSF3, and SRSF9 [42, 155].

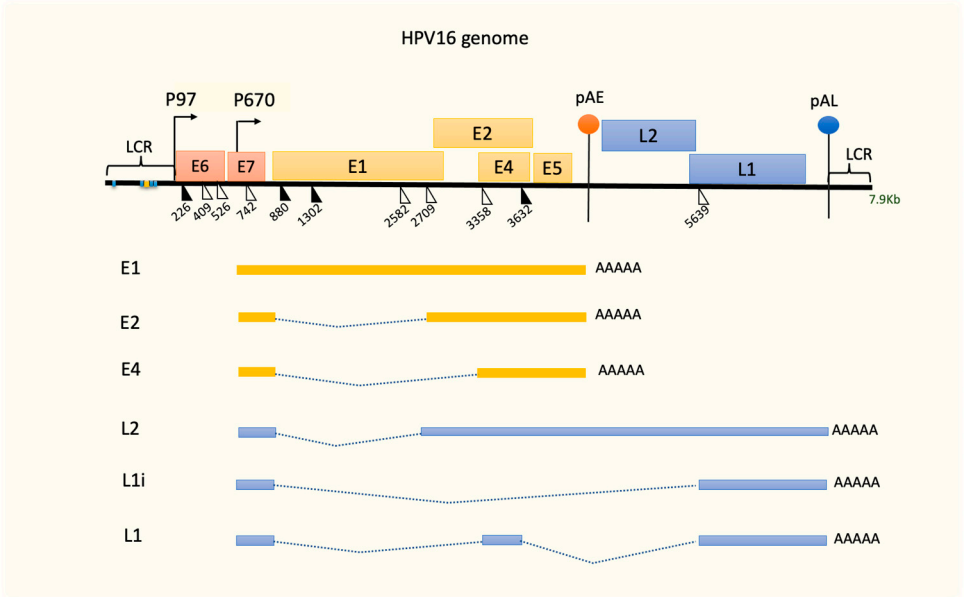


Figure 1.14 Schematic representation of the HPV16 genome and representative early and late mRNAs.

1.8 DNA damage response

The DNA damage response (DDR) or DNA repair is a vital ability of cells to protect its genomic integrity to ensure a normal function of cells and organs in eukaryotes [163]. The DDR is a complex network of proteins by which DNA damages are signalled, transduced, detected and repaired [164, 165]. DDR is constantly activated as it responds to DNA damage followed by either cell senescence or apoptosis or alternatively, a risk of carcinogenesis if DDR fail [2, 164, 166-168]. Ataxia telangiectasis mutated (ATM) and Ataxia telangiectasia and Rad3-related (ATR) proteins are two key signal transducers of the DDR and that are also known as two major regulators downstream of the DDR pathways [169, 170]. Upon DNA strand breaks, ATM and ATR kinases are activated. ATM responds to double-strand DNA breaks while ATR is activated and responds to DNA single-strand breaks **(Figure 1.15)** [169, 171]. Importantly, mountainous reports show that HPVs require and

manipulate the DDR for replication and particularly utilize the ATR and ATM pathways [2, 59, 169, 172-174]. It has been suggested that HPV16 viral proteins E1, E2, E6 and E7 can independently activate ATM and ATR pathways, resulting in the constitutive accumulation of numerous recombination/repair factors at viral replication sites [2, 169, 172, 173, 175]. These factors contribute to efficient viral genomic replication, particularly at the productive phase of the viral life cycle, although the molecular mechanism is still unclear [171-175]. Furthermore, efficient viral amplification of HPV16 may also partly attribute to the ability of the ATM pathway to arrest the cell cycle in G2 phase [169, 176].

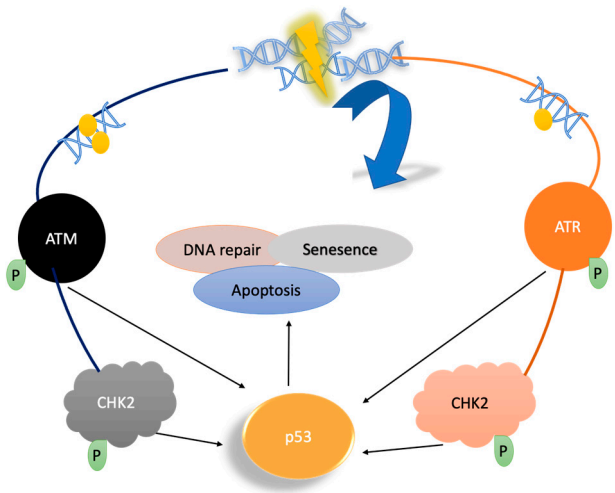


Figure 1.15. Simplified representation of the ATM and ATR DNA damage response pathway. Simply, once phosphorylation of ATM or ATR has occured, the cell cycle effectors checkpoint kinase 2 (CHK2) or CHK1 are activated. Upon phosphorylation of CHK2 or CHK1, the cell cycle is paused allowing repair of DNA damages before continued cell division. Finally, the fate of cell with DNA damages depends on the level of damage or the efficiency of the DDR which induces the critical sensor of DNA damage the p53 protein.

In response to DNA breaks, the DDR has been demonstrated to involve mRNA metabolism. The DDR can influence mRNA stability, alter pre-mRNA splicing pattern and modulate the distribution of splicing factors. Additionally, the DDR affects alternative splicing [177]. Lately our laboratory has reported that DDR contributes to the induction of HPV16 late gene expression at the level of pre-mRNA splicing [178]. However, it is unknown if the DDR contributes to HPV16 early gene expression by influencing the alternative splicing of those mRNAs. Interestingly, recently interactions between DDR and RNA-binding proteins have been identified [179]. Large-scale proteomics and genetic screenings have identified roles of RNA-binding proteins as ATM and ATR substrates although the exact

contribution of each RNA-binding protein is still unclear [154]. Surprisingly, the RNA-binding protein hnRNP G/RBMX which is associated with regulation of alternative splicing has been identified as a component of the DDR [148, 153, 1544]. hnRNP G has been demonstrated to be a positive regulator of homologous recombination associated with the ATM pathway [153, 154, 169]. During the DDR, hnRNP G accumulates at DNA lesions and can facilitate DNA repair [66]. Since hnRNP G is a DDR factor and papillomaviruses utilize the DDR for genome replication as well as for control of viral gene expression, it is of great interest to investigate if activation of the DDR affects HPV16 gene expression via hnRNP G.

2. Aim of the Thesis

Overall aim: To investigate the regulation of HPV16 early and late gene expression at the level of mRNA processing.

Specific aims for each project:

Project I: Identification of heterogenous nuclear ribonucleoproteins (hnRNPs) and serine and arginine rich (SR) proteins that induce human papillomavirus type 16 late gene expression and alter L1 mRNA splicing.

Specific aims: To investigate the role of RNA-binding proteins (hnRNPs and SR proteins) in the regulation of HPV16 late gene mRNA expression and late mRNA splicing.

Project II: hnRNP G/RBMX enhances HPV16 E2 mRNA splicing through a novel splicing enhancer and inhibits production of spliced E7 oncogene mRNAs.

Specific aim: 1. To determine the role of the cellular hnRNP G protein in the control of HPV16 E1-, E2-, E6- and E7-mRNAs.

2. To determine the role of a splicing enhancer RNA element on the HPV16 E2 mRNAs in the control of HPV16 E2 mRNA splicing.

3.To determine if hnRNP G regulates HPV16 gene expression upon activation of the DDR.

Project III: hnRNP D activates production of human papillomavirus type 16 E1 and E6 mRNAs by promoting intron retention.

Specific aims: To investigate the role of cellular factor hnRNP D in the production of the intron retained E1 and E6 mRNAs.

Project IV: Generation of reporter plasmids for HPV16 E1 or E2 mRNA splicing using nanoluciferase and secreted luciferase.

Specific aims: 1. To develop a bioassay for HPV16 E1 and E2 mRNA splicing.

2. To investigate small molecules or novel proteins that affect HPV16 E1/E2 mRNA splicing.

3. Materials and Methods

Plasmids

The following plasmids have been described previously: pBEL [180], pBELsL[181, 182], pC97ELsLuc [181], pHPV16AN [181], hnRNP G [183], phnRNP A1 [159], phnRNP A2 [159], phnRNP D [181] , pX856F [159], p556F [159] and pCL0806 [184]. The other plasmids were described in detail respectively in Papers.

Cells

HeLa cells, C33A2 cells and SiHa cells were cultured in Dulbecco's modified Eagle medium (DMEM) (HyClone) supplemented with 10% bovine calf serum and 1% penicillin-streptomycin. The C33A2 cell line has been described previously [79]. Briefly, C33A2 is derived from the HPV-negative cervical cancer cell line C33A which stably integrated the subgenomic HPV16 plasmid pBELsL contains a gene segment encoding poliovirus 2A internal ribosome entry site together with the Metridia longa secreted luciferase (sLuc) gene in the L1 coding region [79]. Induction of HPV16 late gene expression results in the appearance of sLuc in the cell culture medium. HPV16-infected tonsillar cancer cell line HN26 has been described previously [82]. Briefly, the HN26 cells are derived from a tumuor of a 48-year-old nonsmoking man with non-keratinizing, HPV16-positive tonsil oral squamous cell carcinoma, stage T2N0M0. The HN26 cells contain episomal HPV16 DNA and have an intact p53 gene. HN26 cells were cultured in RPMI 1640 medium (HyClone) with 10% iron-supplemented bovine calf serum (HyClone), 5%MEM Nonessential Amino Acid Solution (Sigma Aldrich) and 5% sodium pyruvate (Sigma Aldrich). The HPV16-immortalized keratinocyte cell line 3310 cell line wascultured in EpiLife medium (Gibco) supplemented with 1% human keratinocyte growth supplement (HKGS, Gibco) and 0.2% Gentamicin/Amphotericin (Gibco). Differentiation of 3310 cells was induced by addition of CaCl2 at a final concentration of 2.4 mM in the keratinocyte culture medium for 24 hrs. 3310 has been described previously and was generated by stable transfection of normal neonatal human foreskin keratinocytes (nHFK) with HPV16 genome plasmid pHPV16ANE2fs [83]. For treatment of C33A2 cells or HN26 cells with melphalan, cell culture medium was replaced with medium containing indicated concentrations of melphalan for indicated time periods. Melphalan hydrochloride (Y0001158, Sigma) was dissolved in DMSO and DMSO without melphalan was used as a control in all experiments.

Transfecton

Transfection of cells was carried out using Turbofect according to the manufacturer's instructions (Fermentas). Turbofect was mixed with plasmid DNA and incubated at room temperature for 15 min prior to drop-wise addition to 60-mm plates with subconfluent cells. Cells were harvested at indicated time points post transfection. Each plasmid was transfected in triplicate, in a minimum of two independent experiments.

RNA extraction, RT-PCR and Real-time quantitative PCR

Total RNA was extracted using TRI Reagent and Directzol RNA MiniPrep kit (ZYMO Research) according to the manufacturer's protocol. For RT-PCR, 1µg of total RNA was reverse transcribed in a 20µl reaction at 37 ℃ using M-MLV Reverse Transcriptase (Invitrogen) and random primers (Thermo Scientific). 1µl of cDNA was subjected to PCR amplification. qPCR was performed on 1µl of cDNA prepared as described above in a MiniOpticon (Bio-Rad) using the SsoAdvanced SYBR Green Supermix (Bio-Rad) according to the manufacturer's instructions. Specific primers for RT-PCR and RT-qPCR are listed respectively in supplemental files in Papers I, II, III and IV.

Nuclear and Cytoplasmic Extraction

Nuclear and cytoplasmic extracts were prepared from HeLa cells grown in 10 cm dishes harvested at 24 h post-transfection. Cells were harvested by scraping and spun at 1500 rpm for 5 min in an Eppendorf centrifuge. Cell pellets were resuspended in 250 ul ice-cold Buffer l (10 mM HEPES pH 7.9, 10 mM KCl, 0.1 mM EGTA, 1 mM DTT and protease inhibitor) and allowed to swell for 15 min on ice. Then, 250 ul of ice-cold NP40-Buffer I (10 mM HEPES pH 7.9, 10 mM KCl, 0.1 mM EGTA, 1 mM DTT, 0.8% NP-40 and protease inhibitor) was added. When lysis was observed under the microscope, samples were spun at 2000 rpm for 10min. Supernatants were collected and 5 M NaCl was added to a final concentration of 137 mM and stored on ice until use (cytoplasmic extracts). Pellets were resuspended in ice-cold RIP buffer and incubated with rotation for 30 min, followed by freezing and thawing and passage twice through a needle.

Finally, samples were centrifuged at 14000 rpm for 20 min and the supernatants collected as nuclear extracts.

siRNAs and siRNA transfections

siRNA knock downs were carried out using Dharma-FECT™ transfection reagent according to the manufacturer's instructions. Briefly, the siRNA was diluted to 40nM final concentration in 250 ul serum free medium, and the mixture was added

to 250 ul of serum free medium with 5 ul transfection reagent. The mixture was incubated at room temperature for 20 min prior to the addition of the mixture to a 60mm plate with subconfluent C33A2, HeLa or SiHa cells. siRNA to hnRNP G was ON-TARGET plus SMART pool Human RBMX (L-011691-01-0005 Dharmacon™). The scrambled control (scr) was siGENOME Control pool non-targeting #2 (D-001206-14-20, Dharmacon™).

RNA-mediated protein pull-down assay

RNA-mediated protein pull-down assay nuclear extracts were prepared according to the procedure described previously [84]. Briefly, the cells were lysed by using lysis buffer A (10 mM HEPES,10 mM KCl, 0.1 mM EGTA, 1 mM DTT, 0.4% NP40 pH 7.9 and protease inhibitors) for cytoplasmic proteins, then using buffer B (10 mM HEPES, 400 mM NaCl, 1 mM EGTA, 1 mM EDTA, 1 mM DTT pH7,9 and protease inhibitor) for nuclear proteins. The nuclear extracts were mixed with streptavidin-coated magnetic beads (Dynabeads M-280 Streptavidin, Invitrogen) carrying biotin-labelled single stranded RNA oligonucleotides in binding buffer (10 mM HEPES, 130 mM NaCl, 2.5 mM $MgCl_2$, 1 mM EGTA, 1mMEDTA, 1mM DTT, 10% glycerol pH 7.9 and protease inhibitor). RNA oligos are listed in Paper II and III supplementary tables. The mixtures were incubated at room temperature with rotation for 1 h, followed by washing 10 times with 1 ml wash buffer (10 mM HEPES, 200 mM NaCl, 2.5 mM $MgCl_2$, 0.5% Triton X-100, 1 mM DTT pH 7.9). Proteins were eluted by boiling of the beads in SDS-PAGE loading buffer and subjected to SDS-PAGE followed by western blot analysis with indicated antibodies.

Preparation of proteins for mass spectrometry

Protein extracts were subjected to SDS-polyacrylamide gel electrophoresis followed by either Western blotting that was performed as described previously, or by staining with silver statin (Silver Quest^{TM} staining Kit, Invitrogen). Bands were excised and subjected to liquid chromatography– mass spectrometry (LC–MS) analysis at the SCIBLU Proteomics Resource Centre at Lund university. Primary and secondary antibodies used for Western blotting are listed in Paper II and III supplementary table. Filters were stained with the Clarity Western ECL Substrate (BioRad) or the Super Signal West Femto chemiluminescence substrate (Pierce).

Immunoprecipitation and Western blotting

Proteins for Western blotting were extracted from cells using the radioimmunoprecipitation assay (RIPA) buffer (50 mM Tris pH 7.6, 150 mM NaCl, 0.5% Na-DOC, 1% NP-40, 1 mM DTT) with 30 min incubation on ice and occasional vortexing followed by centrifugation to remove cell debris.

Immunoprecipitation was performed by over-night incubation of 400 ul of cell extract with Dynabeads and 1–2 ug of antibody followed by washing three times in RIPA buffer. Protein extracts were subjected to SDS-polyacrylamide gel electrophoresis and Western blotting. Primary and secondary antibodies used for Western blotting are listed in Paper II and III supplementary tables. Filters were stained with the Clarity Western ECL Substrate (Bio Rad) or the Super Signal West Femto chemiluminescence substrate (Pierce).

RNA-protein immunoprecipitation assay (RIP)

Cells were lysed by resuspension in 1ml RIP lysis buffer (200 mM Tris, pH 8, 137 mM NaCl, 1% NP-40, 2 mM EDTA, protease inhibitor and RNase inhibitor) followed by incubation on a rotator at 4 ℃ for 30 min. Cell debris was removed by centrifugation. For immunoprecipitations, 1 ug of the indicated antibody or normal mouse IgG was incubated overnight at 4◦C in 0.5 ml of lysate. Antibodies are listed in Paper II and III supplementary tables. 50 ul (1.5 mg) of Dynabeads Protein A (#10002D, Invitrogen) were washed three times with lysis buffer and added to the cell extracts with antibody. The beads were washed six times with wash buffer (200 mM Tris, pH 8, 137 mM NaCl, 1% NP-40, 2 mM EDTA) and RNA was extracted from the immunoprecipitations using Trizol-chloroform, dissolved in water and analysed by RT-PCR directly. RT-PCR primers are listed in Supplementary Table in Paper II and III. The RNA was ethanol precipitated and dissolved in 20 ul of water. 10 ul of immunoprecipitated RNA was reverse transcribed using MMLV reverse transcriptase (Invitrogen) and random hexamer primers (Thermo Scientific) according to the protocol of the manufacturer. 1 ul of cDNA were subjected to PCR amplification using primers indicated in each figure.

Application of small molecule substances to cells

C33A2 or HN26 cells were grown in 10-cm dishes. Cell culture medium was replaced with medium containing indicated concentrations of inhibitors for 21hrs or indicated time points. All inhibitors were dissolved in dimethyl sulfoxide (DMSO) (Sigma) and DMSO in the absence of inhibitor was used as a control in all experiments. Inhibitors were: Melphalan Hydrochloride (#Y0001158, European pharmacopoeia reference standard), 2-D08/S8696 (Selleckchem) (sumoylation-inhibitor) and TAK-981/S8829 (Selleckchem) (sumoylation-inhibitor).

sLuc assay

The Metridia longa secreted luciferase activity in the cultured medium of the C33A2 cells was monitored with the help of the Ready To Glow secreted luciferase reporter assay (Clontech) according to the instructions of the manufacturer.

Nanoluciferase assay

The nanoluciferase activity of Hela cells was measured using the Nano-Glo Luciferase Assay Kit (Promega) according to the manufacturer's instruction. Briefly, Hela cells were washed twice in ice cold PBS then lysed in 100uL Passive Lysis Buffer (Promega) at room temperature for 15min by pipetting up and down. 25 uL of cell lysate were mixed with 25uL of NanoGlo reagent consisting of 1ul nanoluciferase substrate and 24 ul nanoluciferase reaction buffer. After incubated in 3 min, the luminescence was measured in a Tristar LB941 luminometer (Berthold Technologies).

Cell viability assay

HeLa cells were seeded in 96-well plates and incubated at 37◦C in CO2 incubator for 24 h. Transfections were performed with a total amount of 100 ng plasmid DNA per well using 0.3 ul Fugene6 (Promega). The transfected cells were incubated for 24 h followed by addition of 10 ul of WST-1 cell proliferation reagent (Roche) to each well and incubation for 2 h at 37◦C in 5%CO2. The absorbance was monitored at 450 nm in a Tristar LB941.

Quantitations

The software used to determine band intensity in Western blots and RT-PCR gels is 'Image Lab 6.0.1' and quantitations were performed with the software 'Prism GraphPad 8.4.0'.

4. Results

4.1 Identification of heterogenous nuclear ribonucleoproteins (hnRNPs) and serine- and arginine-rich (SR) proteins that induce human papillomavirus type 16 late gene expression and alter L1 mRNA splicing

HPV16 viral late capsid L1 and L2 proteins are immunogenetic proteins expressed exclusively in the late stage of viral life cycle which contribute to immune evasion. HPV16 late L1 and L2 proteins are translated from mRNAs generated by alternative splicing which utilize multiple splice sites including SD880, SA3358, SD3632 and SA5639. The major HPV16 L1 mRNAs are produced by splicing between SD3632 and SA5639. To determine the effect of HPV16 late gene mRNA splicing by trans-acting regulatory proteins including SR proteins and hnRNPs, we compared the effect of various SR proteins and hnRNPs family on HPV16 late gene expression especially L1 alternative splicing in the same experimental system. We concluded that the majority of SR proteins and hnRNPs have the potential to regulate HPV16 L1 mRNA splicing.

4.1.1 Three SR proteins induce HPV16 late gene expression and five SR proteins affect HPV16 late L1 mRNA splicing

We transfected plasmids expressing either of seven SR proteins into HeLa cells with an HPV16-derived, subgenomic reporter plasmid pC97ELsL. This plasmid has a secreted luciferase (sLuc) reporter gene inserted in the L1 coding region that serves as a marker for HPV16 late gene expression **(Figure 4.1A)**. The sLuc activity results revealed that SRSF1, SRFS3 and SRSF9 induced HPV16 late gene expression **(Figure 4.1B)**. The RT-PCR results revealed that SRSF1 and SRSF3 promoted inclusion of exon SA3358-SD3632 while SRSF9 promoted exclusion of this exon compared the empty pUC plasmid **(Figure 4.1C)**. Surprisingly, two other SR proteins SRSF5 and SRSF 9 could not induce L1 gene expression but promoted either inclusion (SRSF5) or exclusion (SRSF9) of exon between SA3358 and SD3632 **(Figure 4.1C)**. Moreover, deletion of the RS domain of SRSF9

(SRp30DRS) did not alleviate the function of SRSF9, suggesting that the RS domain is not required for production of HPV16 late gene expression **(Figure 4.1C)**.

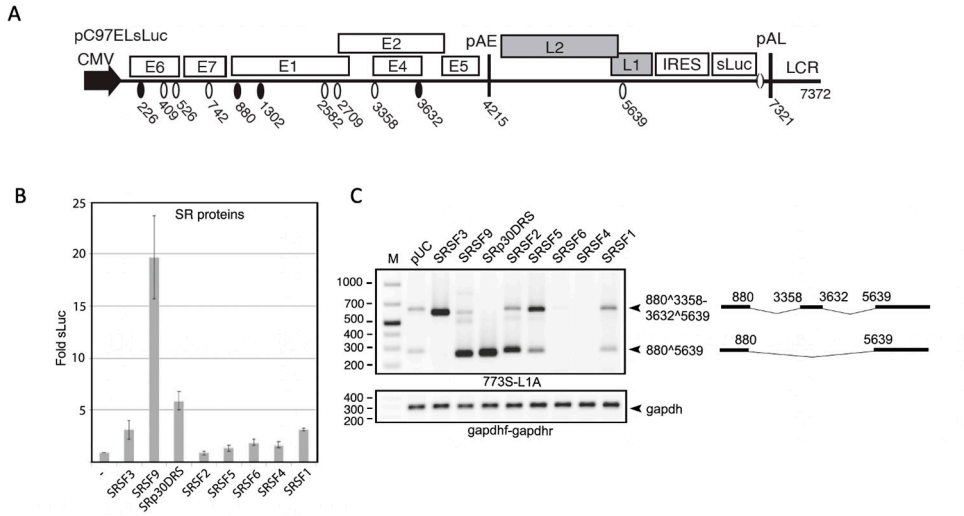


Figure 4.1 (**A**) Schematic representation of the HPV16 subgenomic plasmid pC97ELsLuc. (**B**) Secreted luciferase enzyme activity (sLuc) in the cell culture medium at 24h after transfection of HeLa cells with pC97ELsL and plasmids expressing the indicated SR proteins. SRp30DRS is a deletion mutant of SRSF9 in which the RS domain has been deleted. Mean values and standard deviations of triplicate transfections are displayed. (**C**) RT-PCR with primers 773S and L1A on total RNA extracted from HeLa cells transfected with the indicated plasmids. Schematic representations of the alternatively spliced HPV16 L1 mRNAs detected by the RT-PCR are shown to the right of the gel image. gapdh cDNA was PCR amplified with primers gapdhf and gapdhr. M, molecular weight marker; pUC, pUC vector.

4.1.2 Majority of the hnRNPs affect HPV16 L1 mRNA alternative splicing and HPV16 late gene expression

We investigated the effect of fifteen hnRNPs on HPV16 late gene expression and HPV16 late L1 mRNA splicing. pC97ELsL was individually co-transfected with each of fifteen hnRNP expression plasmids in HeLa cells and sLuc levels were determined **(Figure 4.2A and B)**. The sluc activity results revealed that at least three hnRNPs particularly hnRNP A2, hnRNP F and hnRNP H induced HPV16 late gene expression **(Figure 4.2A)**. The RT-PCR results revealed that hnRNP proteins affected the HPV16 L1 mRNAs in different ways **(Figure 4.2C and D)**. hnRNP A2, hnRNP C, hnRNP E1, hnRNP F, hnRNP H, hnRNP Q and hnRNP R promoted inclusion of the exon between SA3358 and SD3632 on the L1 mRNAs **(Figure 4.2C)**, whereas hnRNP A1, hnRNP AB, hnRNP G, hnRNP K and hnRNP L, caused exclusion of the same exon from the L1 mRNAs **(Figure 4.2C)**.

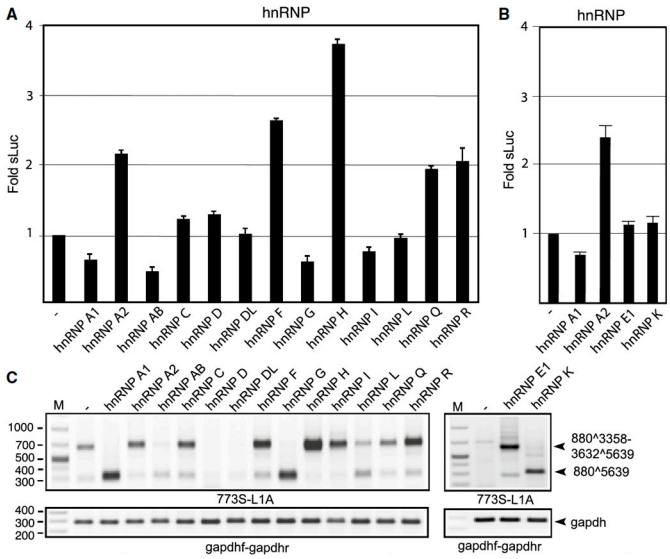


Figure 4.2. (**A and B**) Secreted luciferase enzyme activity (sLuc) in the cell culture medium at 24h after transfection of HeLa cells with pC97ELsL and plasmids expressing the indicated hnRNP proteins. Mean values and standard deviations of triplicate transfections are displayed. (**C**) RT-PCR with primers 773S and L1A on total RNA extracted from HeLa cells transfected with the indicated plasmids. gapdh cDNA was PCR-amplified by primers gapdhf and gapdhr. M, molecular weight marker.

In conclusion, our results suggest that the majority of all hnRNPs and SR proteins have the potential to control HPV16 late gene expression.

4.2 hnRNP G/RBMX enhances HPV16 E2 mRNA splicing through a novel splicing enhancer and inhibits production of spliced E7 oncogene mRNAs

Cellular RNA-binding protein hnRNP G/RBMX is a component of the DNA damage response (DDR) as well as a regulator of RNA processing. hnRNP G/RBMX had been reported by our lab previously to modulate HPV16 late L1 mRNA splicing by promoting exclusion of the central exon between splice sites SA3358 and SD3632 [183]. Thus, it is of great interest to investigate the role of hnRNP G/RBMX in HPV16 early E1-, E2-, E6- and E7- mRNA splicing and to investigate if hnRNP G/RBMX can regulate HPV16 early gene expression upon activation of the DDR. Understanding how HPV activates and utilizes DDR to modulate viral gene regulation at the RNA processing level may identify potential therapy targets that could be exploited limit viral replication and block disease progression.

4.2.1 A splicing enhancer downstream of SA2709 is required to produce HPV16 E2 mRNAs that are spliced to SA2709

Previously, our lab has demonstrated that HPV16 E2 protein was primarily translated by E2 mRNA spliced from SD880 to SA2709 [162]. To further understand how E2 mRNA splicing is regulated, we constructed a set of truncated plasmids based on the HPV16 subgenomic reporter plasmid pBEL **(Figure 4.2.1A and B)**. Transcription of the HPV16 sequences in the pBEL plasmid is driven by the human cytomegalovirus immediate early promoter (CMV) but lack the E6 and E7 parts **(Figure 4.2.1A)**. RT-PCR and RT-qPCR results suggested that a splicing enhancer is located within the sequence between nucleotide position 2731 to 2922. Plasmid pD200 in which a truncation of the sequence between HPV16 nucleotide position 2731 to 2922 produced significantly reduced level of E2 mRNAs spliced from SD880 to SA2708 **(Figure 4.2.1B and C)**. After further mapping of this RNA element, we found that this enhancer element must be located within nucleotide positions 2731 and 2958 **(Figure 4.2.1D)**. The p2758 plasmid which contained a deletion between nucleotide position 2758 to 2922 had the same phenotype as the wildtype control pBEL. Both p2758 and pBEL produced high level of E2 mRNAs in contrast to pD200 in which the entire enhancer was deleted **(Figure 4.2.1D)**. Finally, RT-PCR results from full-length HPV16 genomic reporter plasmids with or without intact enhancer elements (pHPV16AN and pHPV16ANdEN, pC97ELsLuc and pC97ELdEN in **figure 4.2.1F and H**) demonstrated that this splicing enhancer downstream of SA2709 was required to produce HPV16 E2 mRNAs spliced to SA2709 **(Figure 4.2.1G and I)**.

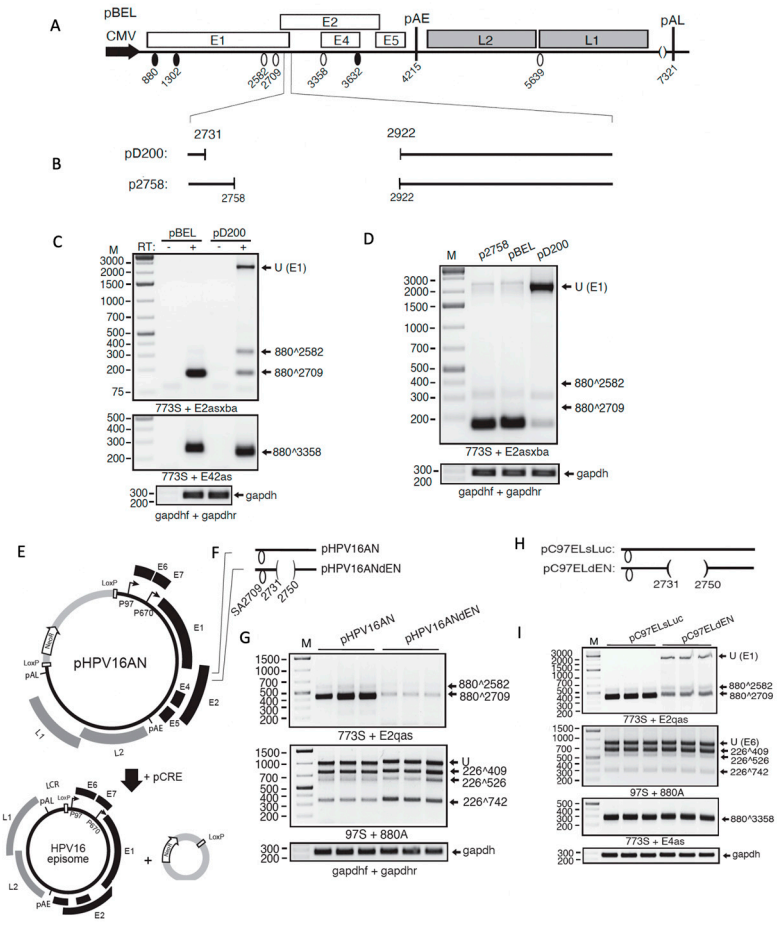


Figure.4.2.1 (**A and B**) Schemetic representation of HPV16 subgenomic plasmid pBEL, pD200 and p2578. (**C and D**) RT-PCR with indicated primers on RNA extracted from HeLa cells transfected with indicated plasmids. RT-PCR was performed in the absence (−) or presence (+) of RT as indicated. (**E**) Schematic representation of HPV16 genomic plasmid pHPV16AN and Cre-loxp transfection system. Co-transfection of pHPV16AN with plasmid pCRE reults in excision and circulation of the viral DNA at the loxp sites to generate episomal HPV16 genome. (**F and H**) Schematic representation of full-length HPV16 genomic reporter plasmids with or without intact enhancer elements (pHPV16AN and pHPV16ANdEN, pC97ELsLuc and pC97ELdEN). (**G and I**) RT-PCR with indicated primers on RNA extracted from HeLa cells transfected with indicated plasmids and spliced mRNAs represented by the amplicons are indicated to the right.

4.2.2 hnRNP G enhances HPV16 E2 mRNA splicing to SA2709 by interacting with splicing enhancer downstream of SA2709

hnRNP G was identified as a cellular factor that interacted with the enhancer element downstream of HPV16 by pull-down assay with various comparable RNA oligos followed by mass spectrometric analysis (data do not show here). The ability of hnRNP G to enhance HPV16 E2 mRNA splicing by interacting with the enhancer element was confirmed by RT-PCR with a set of plasmids with wildtype or mutant enhancer **(Figure 4.2.2A-C)**. The Western blotting were performed with HPV16 subgenomic plasmids shown in **figure 4.2.2D** and the results indicated that hnRNP G promoted HPV16 E2 protein expression only when plasmid contained an intact enhancer element **(Figure 4.2.2E)**. These results indicated that hnRNP G acted on the splicing enhancer downstream of SA2709 to enhance splicing to the HPV16 E2 mRNA 3' splice site SA2709, thereby enhancing production of HPV16 E2 protein.

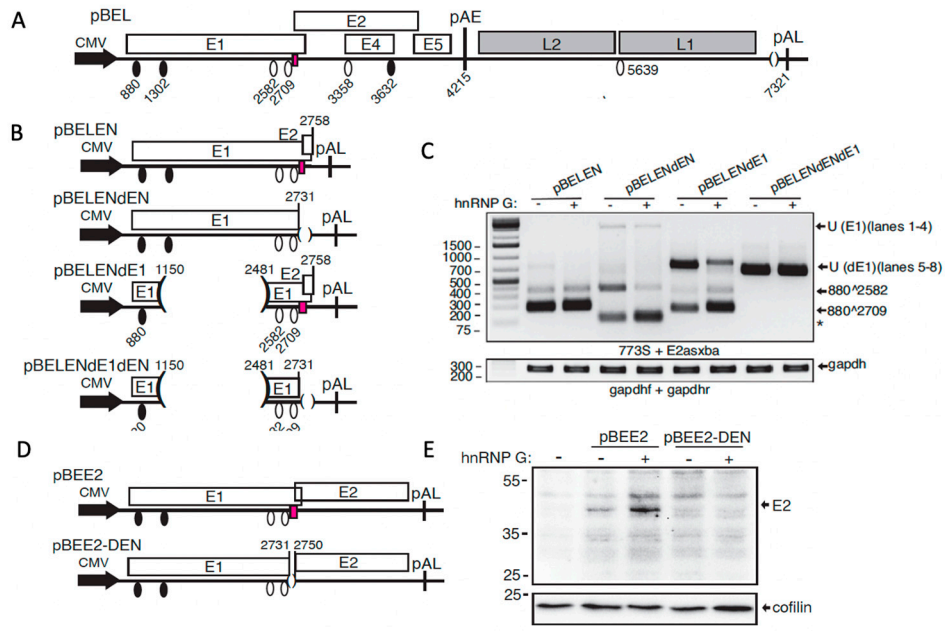


Figure.4.2.2 (**A**) Schematic representation of HPV16 subgenomic pBEL reporter plasmid. (**B and D**) Schematic representation of the HPV16 subgenomic expression plasmids pBELEN, pBELENdEN, pBELENdE1, pBELENdE1dEN, pBEE2 and pBEE2-DEN. Transcription of the HPV16 sequences in these plasmids is driven by the human cytomegalovirus immediate early promoter (CMV). HPV16 splice sites are indicated. (**C**) RT-PCR with indicated primers on RNA extracted from HeLa cells transfected with indicated HPV16 subgenomic plasmids in with or without hnRNP G plasmid. (**E**) Western blotting on cell extracts from HeLa cells transfected with indicated plasmids in the absence or presence of hnRNP G.

4.2.3 hnRNP G affects HPV16 E2, E6 and E7 mRNA splicing in the full-length HPV16 genome

hnRNP G exhibited effects on HPV16 E2 mRNA splicing in a different manner depending on the absence of presence of E6 and E7 regions. When lacking HPV16 E6 and E7 sequence, hnRNP G showed enhancement function on E2 mRNA splicing to SA2709 by interacting with the enhancer element downstream of SA2709 **(Figure 4.2.2A-C)**. In contrast, hnRNP G did not promote E2 mRNA splicing to SA2709 but enhanced production of unspliced E6 mRNA production at the expense of spliced E7 mRNA when the E6 and E7 coding sequences were present on the pre-mRNAs **(Figure 4.2.3A)**. Those results were confirmed by other HPV16 subgenomic reporter plasmids. In addition to effects on HPV16 E6 and E7 mRNA splicing, hnRNP G also affected HPV16 E6 and E7 gene expression, enhancing expression of E6 protein but reducing E7 protein expression **(Figure 4.2.3B and C)**.

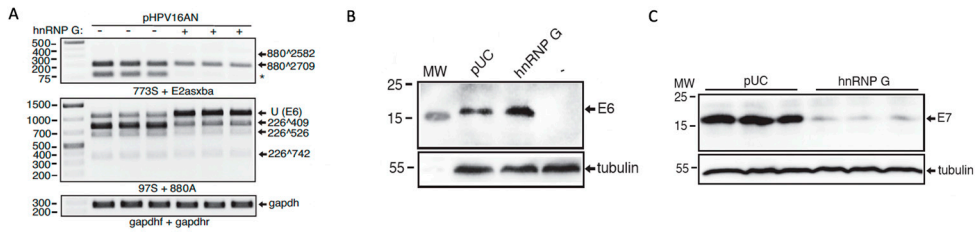


Figure.4.2.3 (**A**) RT-PCR with indicated primers on RNA extracted from HeLa cells co-transfected with pHPV16AN and cre-expressing plasmid in the absence or presence of phnRNP G. (**B and C**) Western blotting on cell extracts from HeLa cells for HPV16 E6 and E7 protein detection. (−), untransfected cells.

4.2.4 Different domains of hnRNP G contribute to HPV16 E2, E6 and E7 mRNA splicing

The enhancement of HPV16 E2 mRNA splicing and inhibition of HPV16 E6/E7 mRNA splicing were exerted by non-overlapping domains of hnRNP G. Primarily, the ability of hnRNP G to enhance HPV16 E2 mRNA splicing to SA2709 was attributed to the central and C-terminal region of hnRNP G. In contrast, hnRNP G mediated inhibition of E6/E7 mRNA splicing required the N-terminal region of hnRNP G. RT-PCR results showed that the deletion mutants separated splicing-enhancing and splicing-inhibitory domains of hnRNP G **(Figure 4.2.4A and B)**. Finally, the splicing-enhancing domain of hnRNP G was mapped to hnRNP G-sequences between amino acids 236 and 286 and the splicing inhibitory region of hnRNP G to hnRNP G-sequences between amino acids 127 and 186. Taken together, the conclusion was that the enhancement of HPV16 E2 mRNA splicing and inhibition of HPV16 E6/E7 mRNA splicing were exerted by non-overlapping domains of hnRNP G.

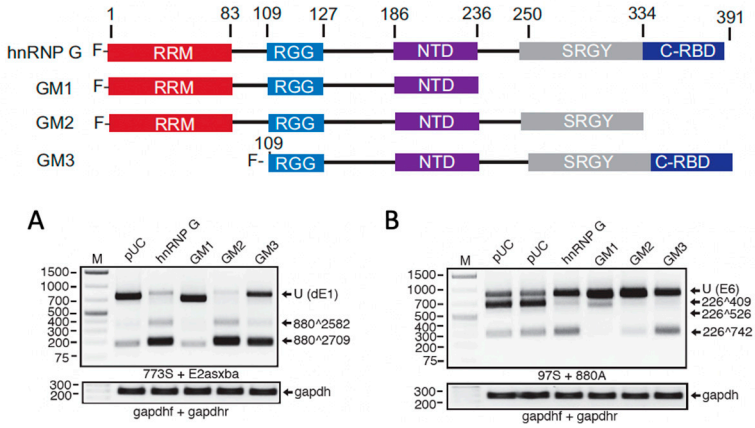


Figure.4.2.4 (**A**) RT-PCR with indicated primers on RNA extracted from HeLa cells co-transfected with pBELENdE1 and empty pUC plasmid or plasmids expressing wild type or mutant hnRNP G. (**B**) RT-PCR with indicated primers on RNA extracted from HeLa cells co-transfected with pC97ELsLuc and empty pUC plasmid or plasmids expressing wild type or mutant hnRNP G.

4.2.5 hnRNP G interacts with splicing factor U2AF65 in a manner that correlates with splicing activation

The mechanism of action of hnRNP G was delineated and hnRNP F was shown to interact with core splicing factor U2AF65 to enhance HPV16 E2 mRNA splicing to SA2709. Co-immunoprecipitation assay results showed that hnRNP G robustly interacted with U2AF65 in either non-transfected or transfected cells **(Figure 4.2.5A)**. hnRNP G mutant 1 GM1 could not immunoprecipitate U2Af65 but GM2 and GM3 could, demonstrating that interactions of hnRNP G mutants with U2AF65 corelated with the splicing enhancing function of hnRNP G **(Figure 4.2.5B and C)**.

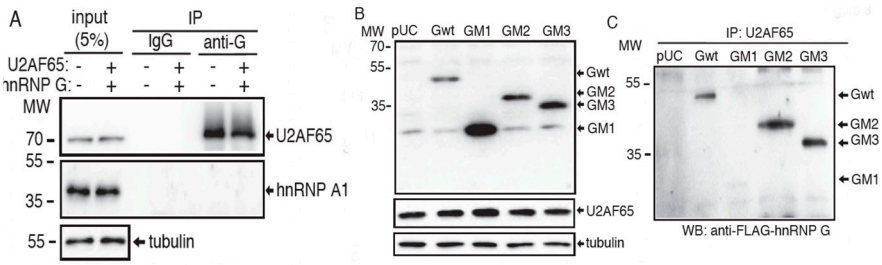


Figure 4.2.5 (**A**) Immunoprecipitation of proteins in cell extracts from untransfected and transfected HeLa cells using either IgG or anti-hnRNPG antibody followed by Western blotting with specific antibodies. (**B**) Western blotting on cell extracts of HeLa cells transfected with plasmid encoding wild type, flag-tagged hnRNP G or the indicated flag-tagged hnRNP G mutants. (**C**) Coimmunoprecipitation experiment on cell extracts from HeLa cells transfected with pUC-plasmid or Flag-tagged hnRNP G or the indicated Flag-tagged hnRNP G mutants using anti-U2AF65 antibody followed by Western blotting with anti-Flag antibody.

4.2.6 Activation of the cellular DNA damage machinery (DDR) enhances HPV16 E2 mRNA splicing and interactions of hnRNP G with E2 mRNAs as well as with U2AF65 in p53-inactivated C33A2 cells

hnRNP G regulated HPV16 E2 mRNA splicing upon activation of the DDR. The activation of DDR by melphalan treatment induced HPV16 E2 mRNA splicing **(Figure 4.2.6A)**. Melphalan is a cancer drug that damage DNA and strongly increases the DDR. hnRNP G protein expression levels increased by activation of the DDR **(Figure 4.2.6B)**, as did the interactions of hnRNP G with HPV16 E2 mRNAs and hnRNP G interactions with U2AF65 as shown by the RIP and Co-IP results **(Figure 4.2.6C and D)**.

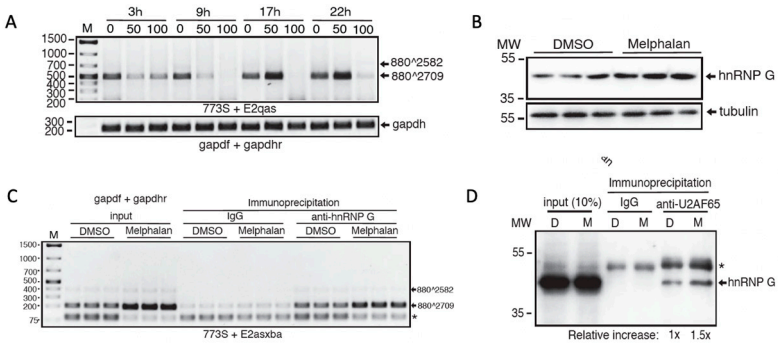


Figure.4.2.6 (**A**) RT-PCR with primers 773S and E2qas on RNA extracted from C33A2 cells treated with the indicated concentrations of melphalan for the indicated time periods. (**B**) Western blotting on cell extracts from C33A2 cells treated with DMSO or 50uM melphalan for 20hrs in triplicates. Blots were stained with antibody specific for hnRNP G or tubulin. (**C**) RIP assay. Immunoprecipitation of C33A2 cells lysis treated with DMSO or melphalan with IgG or anti-hnRNP G antibody followed by RT-PCR. (**D**) Co-IP assay. Immunoprecipitation of proteins in cell extracts from DMSO (D) or melphalan (M)-treated C33A2 cells using either IgG or anti-U2AF65 antibody followed byWestern blotting with antibody to hnRNP G.

4.2.7 Melphalan treatment of apoptosis-prone HN26 cells negatively affects E2 mRNA levels and reduced interactions between hnRNP G and U2AF65 and HPV16 E2 mRNAs

hnRNP G were associated with HPV16 E2 mRNAs in the apoptosis-prone and HPV16- positive tonsillar cancer cells HN26 **(Figure 4.2.7A)**. The interactions of hnRNP G with HPV16 E2 mRNA and with U2AF65 were reduced in HN26 cells when treated with melphalan **(Figure 4.2.7B and C)**. We also found that U2AF65 did not interact with hnRNP A1 in HN26 cells either. Thus, we concluded that in the HPV16-positive tonsillar cancer cell line HN26, induction of the DDR was followed by apoptosis and resulted in reduced levels of HPV16 E2 mRNAs accompanied by reduced association of hnRNP G with HPV16 E2 mRNAs and with U2AF65.

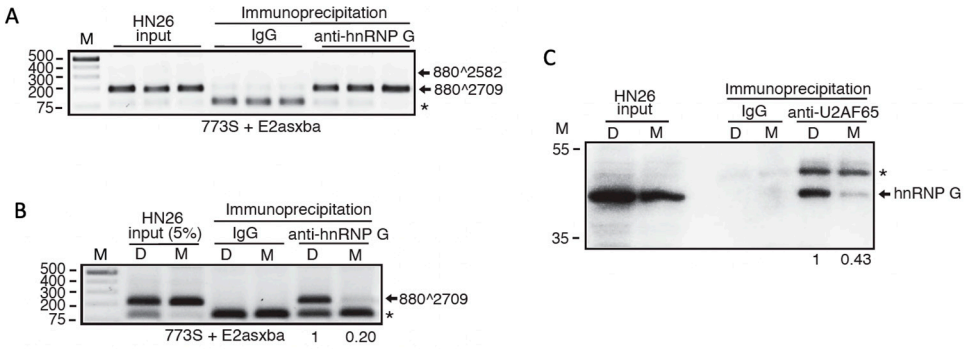


FIgure 4.2.7. (**A**) RIP assay. Immunoprecipitation on the extract from HN26 cell with IgG or anti-hnRNP G antibody followed by extraction of the immunoprecipitated RNA and RT-PCR by HPV16 specific primers 773s and E2asxba. (**B**) RIP assay. Immunoprecipitation on extract from HN26 cell treated with DMSO or Melphalan using IgG or anti-hnRNP G antibody followed by extraction of the immunoprecipitated RNA and RT-PCR by HPV16 specific primers 773s and E2asxba. (**C**) Immunoprecipitation of cell extracts from DMSO or melphalan treated HN26 cells using either indicated antibody followed by Western blotting with antibodies to hnRNP G.

4.2.8 Phosphorylation, parylation and mono-methylation of hnRNP G are enhanced by DDR activation, whereas di-methylation and sumoylation are reduced

hnRNP G was phosphorylated, parylated, methylated and suomylated in C33A2 cells **(Figure 4.2.8A-E)**. Phosphorylation, parylation and mono-methylation of hnRNP G increased in melphalan-treated C33A2 cells **(Figure 4.2.8A, C and D)**, whereas sumoylation and asymmetric di-methylation of hnRNP G decreased **(Figure 4.2.8B and E)**. Taken together, we concluded that increased hnRNP G phosphorylation, parylation and monomethylation and decreased hnRNP G sumoylation and di-methylation correlated with increased interactions of hnRNP G with HPV16 E2 mRNAs.

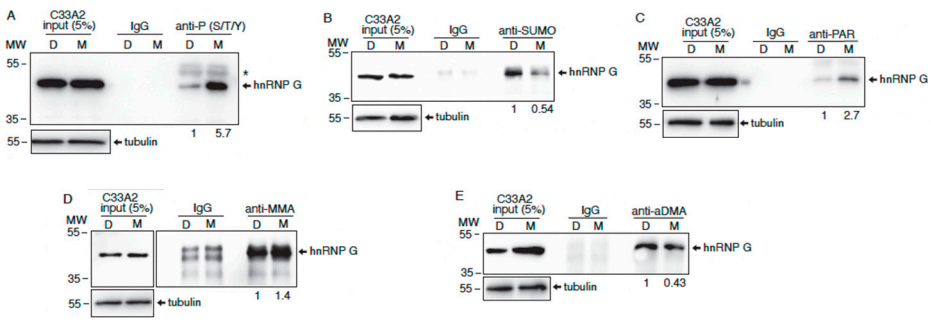


Figure 4.2.8 (**A-E**) Immunoprecipitation of proteins in cell extracts from DMSO (D) or melphalan (M) treated C33A2 cells using either IgG or indicated antibody followed by Western blotting with antibody to hnRNP G. Input samples represent 5% of the sample volume used for immunoprecipitation. MW, molecular weight marker.

4.2.9 Inhibition of sumoylation and protein N-methylation induced HPV16 E2 mRNA splicing and enhanced the association of hnRNP G with HPV16 E2 mRNAs and with U2AF65

HPV16 E2 mRNA levels were significantly induced in C33A2 cells after treatment of these cells by pharmacological inhibitors either of the protein arginine N-methyltransferase 1 (PRMT1) inhibitor MS023 or by sumoylation inhibitor TAK-981 and 2-D08 **(Figure 4.2.9A-C)**. The association of hnRNP G with HPV16 E2 mRNAs was enhanced by either of methyltransferase inhibitor MS023 or sumoylation inhibitors TAK-981 and 2-D08 **(Figure 4.2.9D-F)**. Furthermore, the interactions between hnRNP G and the splicing factor U2AF65 were also enhanced by sumoylation inhibitors TAK-981 or 2-D08 **(Figure 4.2.9G and F)**. Taken together, we concluded that HPV16 E2 mRNA splicing was mediated by interactions between hnRNP G and the HPV16 E2 mRNA splicing enhancer and the splicing factor U2AF65, and that these interactions were regulated by protein methylation and sumoylation.

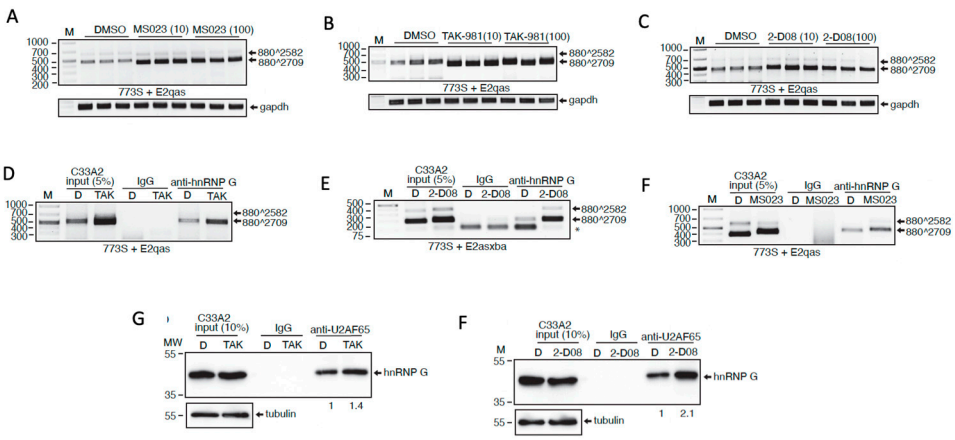


Figure 4.2.9 (**A-C**) RT-PCR. C33A2 cells were treated with DMSO (D) or 10- or 100 uM of methyltransferase inhibitor MS023 or sumoylation inhibitor TAK-981and 2-D08 followed by RNA extraction and RT-PCR with HPV16-specific primers 773s and E2qas. (**D-F**) RIP assay. C33A2 cells treated with DMSO (D) or indicated inhibitors were lysed and subjected to immunoprecipitation with IgG or anti-hnRNP G antibody. RNA was extracted from the Immunoprecipitated RNA-protein complexes and subjected to RT-PCR with HPV16-specific primers 773s and E2qas. (**G-F**) Co-IP assay. Immunoprecipitation of proteins in cell extracts from DMSO (D) or sumoylation inhibitor TAK-981 (TAK) or 2-D08 treated C33A2 cells using either IgG or anti-U2AF65 antibody followed by Western blotting with antibody to hnRNP G. Input samples represent 10% of the sample volume used for immunoprecipitation.

4.2.10 Association between HPV16 E2 mRNAs and hnRNP G in HPV16-immortalized human keratinocytes increases in response to cell differentiation

hnRNP G was associated with HPV16 E2 mRNAs in HPV16-immortalized human keratinocyte 3310 cell line **(Figure 4.2.10A)**. And the association of HPV16 E2 mRNAs with hnRNP G was enhanced by calcium-induced differentiation of the 3310 cells **(Figure 4.2.10B)**. Moreover, RT-PCR on the involucrin differentiation marker confirmed that calcium induced cell differentiation in the 3310 cells **(Figure 4.2.10C)**. We concluded that hnRNP G was associated with HPV16 E2 mRNAs in HPV16-immortalized human keratinocytes and that this interaction increased in response to cell differentiation.

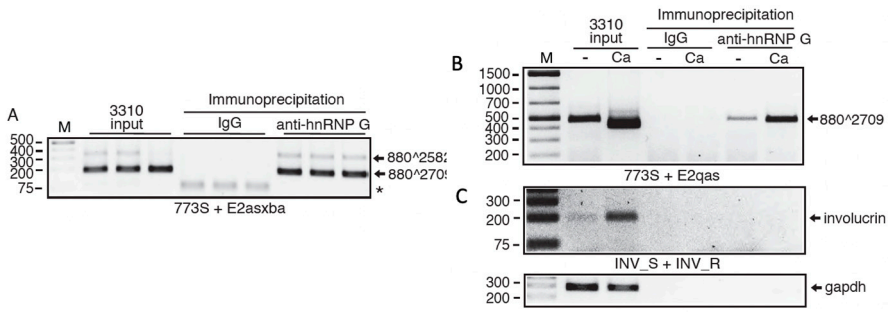


Figure.4.2.10 (**A**) HPV16-immortalized human keratinocyte cell line 3310 cells were lysed in RIPA buffer and subjected to immunoprecipitation with IgG or anti-hnRNP G antibody followed by extraction of the immunoprecipitated RNA and RT-PCR by HPV16 specific primers 773s and E2asxba. (**B**) HPV16-immortalized human keratinocyte cell line 3310 cells were treated with 2.4mM CaCl2 to induce differentiation, lysed in RIPA buffer and subjected to immunoprecipitation with IgG or anti-hnRNP G antibody followed by RT-PCR (**C**) RT-PCR with primers specific for mRNAs encoding the cell differentiation marker involucrin performed on the cDNA samples used in (B).

In conclusion, we determined that:

(1) The cellular RNA-binding protein hnRNP G acts on the E6/E7 coding region to inhibit splicing of the HPV16 E6/E7 mRNAs thereby altering the balance in production of the HPV16 E6 and E7 oncogenes: the levels of the pro-mitotic E7 protein are reduced whereas the anti-apoptosis E6 protein are increased.

(2) hnRNP G recruits the splicing factor U2AF65 and acts on the splicing enhancer downstream of SA2709 to enhance splicing to the HPV16 E2 mRNA 3' splice site SA2709, thereby enhancing production of HPV16 E2 protein when late promoter P670 is activated.

(3) Different domains of the hnRNP G protein contribute to the inhibition and enhancement of HPV16 mRNA splicing.

(4) hnRNP G increases HPV16 E2 mRNA splicing upon the activation of the DDR. The interactions of hnRNP G with HPV16 E2 mRNAs and interactions of hnRNP G with U2AF65 are enhanced by the activation of DDR.

(5) Posttranslational modifications of hnRNP G protein are regulated by the activation of DDR.

(6) hnRNP G sumoylation regulates the HPV16 E2 mRNA splicing and the interactions between hnRNP G and HPV16 E2 mRNAs and the interactions between hnRNP G and the splicing factor U2AF65.

4.3 hnRNP D activates production of HPV16 E1 and E6 mRNA by promoting intron retention

RNA-binding protein hnRNP D has four isoforms generated by alternative splicing of its pre-mRNA. Generally, they are called hnRNP D37, D40, D42 and D45 depending on their molecular weights. In this project, I identified hnRNP D as a HPV16 trans-acting regulatory protein that exhibited inhibitory functions on HPV16 early intron retained mRNA splicing thereby promoting production of intron retained E1 and E6 mRNAs. Furthermore, hnRNP D facilitated export of the intron retained HPV16 E1 mRNAs to the cytoplasm.

4.3.1 All four hnRNP D isoforms promote intron retention of HPV16 E1 and E6 mRNAs

RT-PCR results show that all hnRNP D isoforms had an inhibitory effect on HPV16 E6/E7 mRNA splicing, which resulted in production of intron retained E6-encoding mRNAs at the expense of the spliced E7 mRNAs **(Figure 4.3.1 A)**. RT-PCR results also revealed that hnRNP D proteins promoted production of intron retained mRNAs encoding E1 at the expense of the spliced E2 mRNAs (880^2709) **(Figure 4.3.1B)**. Furthermore, hnRNP D40 has a stronger inhibitory effect than other isoforms **(Figure 4.3.1A-C)**.

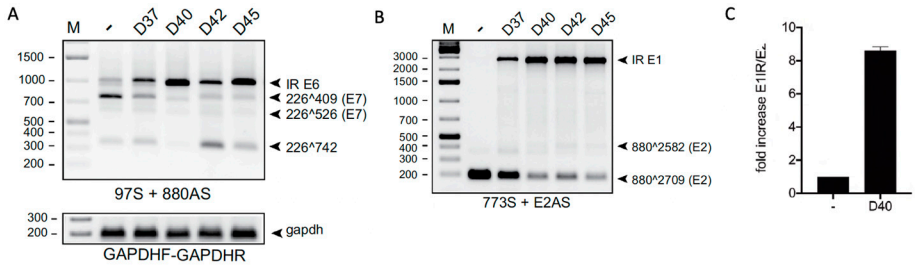


Figure 4.3.1 (**A and B**) RT-PCR with indicated primers on RNA extracted from HeLa cells transfected with HPV16 subgenomic plasmids pC97ELsLuc in with or without hnRNP D isoform plasmids. (**C**) RT-qPCR to quantitate intron-retained E1 mRNAs and spliced E2 mRNAs from pC97ELsLuc in the absence of presence of hnRNP D40 was performed. Fold change of intron retained E1 mRNAs over spliced E2 mRNAs is shown.

4.3.2 hnRNP D40 increase the levels of HPV16 intron retained E1 mRNAs in the cytoplasm

The RT-PCR results indicated that the levels of spliced E2 mRNA were reduced both in nuclear and cytoplasmic fractions in the presence of hnRNP D40, predominantly in the cytoplasm **(Figure 4.3.2 A)**, while the intron retained E1 mRNAs were detected in both the nuclear and cytoplasmic fractions in the presence of hnRNP D40 **(Figure 4.3.2 A)**. These results demonstrated that hnRNP D40 not only inhibited splicing of the HPV16 E1/E2 mRNAs and promoted intron retention to generate intron retained E1 mRNAs, but it also helped to export these mRNAs to the cytoplasm. Actin mRNAs served as controls for cellular fractionation: unspliced actin mRNAs were present primarily in the nuclear fractionation, whereas spliced actin mRNAs were detected in both fractions **(Figure 4.3.2 B)**. The Western blotting results revealed that hnRNP D40 produced from the expression plasmid was present in both nuclear and cytoplasmic fractions, whereas tubulin was detected primarily in the cytoplasmic fraction and histone in the nuclear fractionation **(Figure 4.3.2 C)**. Furthermore, CLIP assay on cytoplasmic extracts revealed that hnRNP D40 was associated with the intron retained HPV16 E1 mRNAs in the cytoplasm since the intron retained HPV16mRNAs could be specifically amplified from the UV-crosslinked RNA–protein complexes immunoprecipitated with anti-FLAG antibody **(Figure 4.3.2 D)**. Taken together, we concluded that hnRNP D40 exported the intron retained HPV16 E1 mRNAs from the nucleus to the cytoplasm.

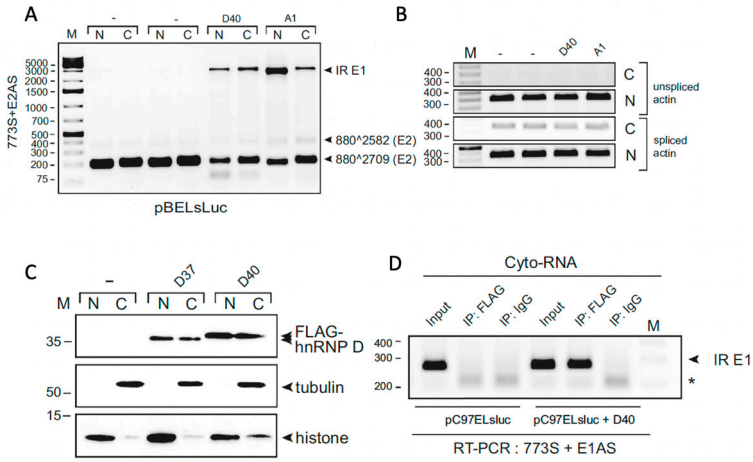


Figure 4.3.2 (**A**) RT-PCR with indicated primers on RNA extracted from Nuclear (N) and cytoplasmic (C) fractions of HeLa cells transfected with HPV16 subgenomic plasmids pBELsluc with or without hnRNP D40 or hnRNP A1 plasmids. (**B**) Cellular fractionation was validated by analysis by RT-PCR of unspliced and spliced actin mRNAs. (**C**) Western blotting of nuclear and cytoplasmic extracts with anti-FLAG antibody or anti-tubulin or anti-histone antibody to control for subcellular fractionation. (**D**) CLIP assay on cytoplasmic extracts from HeLa cells transfected with HPV16 subgenomic plasmid pC97ELsLuc in the absence or presence of FLAG-hnRNP D40 expressing plasmid.

Unexpectedly, the HPV16 E1 and E6 mRNAs that came to the cytoplasm were not efficient translated. These results suggested that the intron retained mRNAs were translationally suppressed by hnRNP D. Other factors may be needed for translation of the intron retained HPV16 E1 and E6 mRNAs.

4.4 Generation of reporter plasmids for HPV16 E1 or E2 mRNA splicing using nanoluciferase and secreted luciferase

HPV16 viral protein E1 and E2 play vital roles in viral genomic replication during virus infection cycle. HPV16 E1 protein is produced from unspliced intron retained mRNA while HPV16 E2 is primarily produced from mRNAs spliced from HPV16 5' splice site SD880 to HPV16 3' splice site SA2709. Alternatively splicing to a competing splice site SA2582 upstream of SA2709 can produce spliced mRNAs encoding a small protein called E1C. We wished to establish bio-assays that can monitor expression of HPV16 spliced E2 mRNAs or intron retained E1 mRNAs.

4.4.1 Generation and characterization of the pE2sLuc reporter plasmid

Subgenomic HPV16 reporter plasmids pE2sLuc was previously generated to monitor E2 mRNA production. In pE2sLuc, the entire E2 open reading frame had been replaced by the secreted luciferase (sLuc) reporter gene (pE2sLuc) **(Figure 4.4.1A)**. sLuc production from pE2sLuc relied on the enhancer element downstream of E2 mRNA splice site SA2709 identified previously and described in **4.2.1**. The sLuc levels were significantly reduced from pE2ENsLuc or from other enhancer mutated plasmids compared with pE2sLuc **(Figure 4.4.1B and C)**. Furthermore, the utility of the pE2sLuc plasmid was verified by several previously identified RNA-binding protein including hnRNP G, hnRNP D40 and hnRNP A1 that affect E2 mRNA splicing **(Figure 4.4.1D)**. We concluded that the pE2sLuc reporter plasmid could be used to investigate regulation of splicing to HPV16 E2 splice site SA2709 in a splicing-enhancer dependent manner.

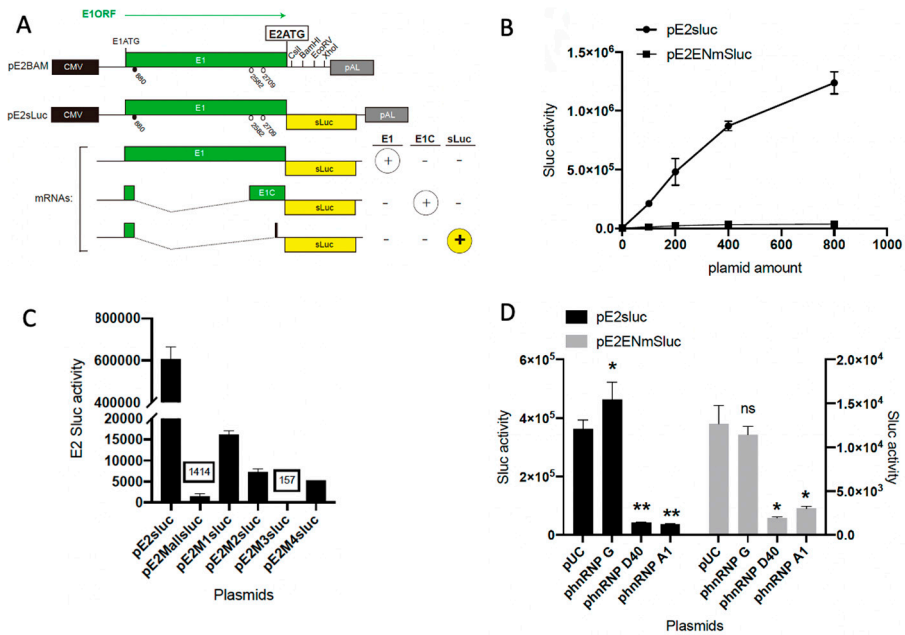


Figure 4.4.1 (**A**) Schematic representation of the HPV16 plasmid pE2BAM and the HPV16 reporter plasmid pE2sLuc. mRNAs produced by pE2sLuc are indicated and the major predicted translation product of each mRNA is indicated to the right. (**B**) HeLa cells were transfected in triplicates with the indicated amounts of serially diluted pE2sLuc or pE2ENmsLuc plasmid DNA. pE2ENmsluc plasmid is generated by deletiion the sequence of enancer element in pE2sLuc plasmid. (**C**) HeLa cells were transfected in triplicates with the pE2sLuc, pE2ENmsluc, pE2M1sluc, pE2M2sluc, pE2M3sluc and pE2M4sluc plasmid DNA. (**D**) HeLa cells were transfected in triplicates with pE2sLuc or pE2ENmsLuc plasmid DNA in the absence or presence of plasmids expressing hnRNP G, hnRNP D40 or hnRNP A1. *, $p < 0.05$; **, $p < 0.01$.

4.4.2 Generation and characterization of the pE1NLuc reporter plasmid

Subgenomic HPV16 reporter plasmids pE1nLuc was generated to monitor product of E1 or E1C mRNA. The E1 or E1C open reading frames were fused in frame with the nanoluciferase reporter gene resulting in plasmid pE1NLuc **(Figure 4.4.2A)**. The pE1Nluc plasmid was characterized by transfection with plasmids expressing RNA-binding proteins of the hnRNP family including hnRNP G, hnRNP D40, hnRNP A1 as well as SR proteins Tra2β, SRp40 and SC35 **(Figure 4.4.2B and C)**. The results revealed that the HPV16 pE1NLuc reporter plasmid could be used to investigate production of unspliced intron retained E1 mRNAs or E1C mRNAs spliced from SD880 to SA2582.

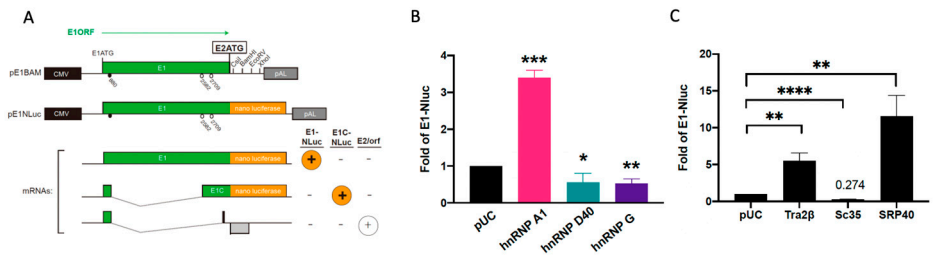


Figure 4.4.2 (**A**) Schematic representation of the HPV16 plasmid pE1BAM and the reporter plasmid pE1NLuc. mRNAs produced by pE1NLuc are indicated and the major predicated translation product of each mRNA is indicated to the right. Numbers represent HPV16 splice sites and refer to the HPV16 reference genome HPV16R. E1C is an ORF in which the 5'-end of the E1 ORF is spliced in frame with the 3'-end of the E1 ORF. The HPV16 E1 open reading frame and the E2 ATG are indicated. The polylinker downstream of the E1 orf is shown. (**B and C**) HeLa cells were transfected in triplicates with the indicated plasmid DNA. *, $p < 0.05$; **, $p < 0.01$.

In conclusion, bio-assays for HPV16 E2 mRNA splicing or E1 intron retention could be generated by establishing stable cell lines containing either pE2sLuc or pE1NLuc. Such bioassays could be used to investigate regulation of HPV16 E1 and E2 expression or to identify antiviral or antitumor drugs to HPV16- infections.

5. Summary, conclusions and further perspectives

In this thesis, we have identified regulation of HPV16 early and late gene expression, primarily mRNA alternative splicing by cis-acting RNA elements and trans-acting regulatory factors. We found that the majority of all hnRNPs had the potential to control HPV16 early and late gene expression **(Paper I and III)** and most SR proteins had the potential to regulate HPV16 late gene expression (Paper I). We also found a cis-acting RNA element downstream of HPV16 E2 splice site SA2709 required for efficient production of spliced HPV16 E2 mRNA **(Paper II)**. We further found that hnRNP G is the corresponding trans-acting regulator of this enhancer element of spliced HPV16 E2 mRNA. Furthermore, hnRNP G has different roles in regulating of HPV16 early E6/E7 and E1/E2 gene expression as well as mRNA processing (Paper II). When HPV16 early promoter is active, hnRNP G has a splicing inhibitory effect that inhibits production of E7 mRNAs, thereby reducing expression of E7 protein. When transcription of HPV16 genome switches from the early promoter to the late promotor, the activation of hnRNP G by cell differentiation and DDR has an entirely positive effect on E2 mRNA splicing through interactions with the splicing enhancer downstream of SA2709 **(Paper II)**. We further demonstrated that hnRNP D can regulate HPV16 early gene E1 and E6 mRNAs expression by promoting intron retention and export of E1- and E6-encoding mRNAs to the cytoplasm **(Paper III)**. Finally, we generated and characterised two subgenomic HPV16 reporter plasmids named pE2sLuc and pE1Nluc that could be used to establish bioassays that can monitor expression unspliced intron retained E1 mRNAs or spliced E2 mRNAs that utilize SA2709, or E1C mRNAs utilizing splice site SA2582 **(Paper IV)**.

The cis-acting RNA element downstream of HPV16 E2 mRNA splice site SA2709 that we identified in **paper II** is definitely required for efficient production of spliced HPV16 E2 mRNAs to SA2709 when the transcription initiated from HPV16 late promoter P670. However, this HPV16 E2 splicing enhancer ACGAGGACGAGGACAAGG is not strictly conserved among other HPV types while all PVs presumably produce E2 protein from spliced E2 mRNAs. In the future, it is of interest, therefore, to evaluate the conserved level of this enhancer in a bigger number of PVs. Furthermore, it is still unclear if the corresponding RNA-binding proteins we identified in **Paper II** bind to this E2 spliced mRNA enhancer

directly or if they acquire specificity for the enhancer via protein-protein interactions. Here, we speculate that hnRNP G may bind to this enhancer directly, since the enhancer at SA2709 is AG-rich, and hnRNP G has affinity for purine-rich sequences, but the exact mode of interaction between hnRNP G and the HPV16 splicing enhancer remains to be determined.

The cellular protein hnRNP G we identified in **Paper II** is multifunctional. And the different functions of hnRNP G in HPV16 E6/E7 and E1/E2 genes expression or mRNA splicing rely on two non-overlapping regions of hnRNP G structure. Since the part of hnRNP G that mapped to its amino acids 236-286 enhanced E2 mRNA splicing is separated from the part of hnRNP G that mapped to amino acids 127-186 inhibited E6/E7 mRNA splicing, the activity of these regions could potentially be independently regulated by posttranslational modifications. We have identified a number of posttranslational modifications of hnRNP G in **4.2.8** including phosphorylation, parylation, methylation and sumoylation that either show an increase (phosphorylation, parylation and mono-methylation) or decrease (sumoylation and dimethylation) in response to the activation of DDR. Inhibition of protein sumoylation or methylation correlated with increased spliced HPV16 E2 mRNAs and an increase in the association of hnRNP G with HPV16 E2 mRNAs and with U2AF65. Intriguingly, lysines that can be sumoylated are present primarily in the N-terminal part of hnRNP G and not between amino acids 236 and 286, whereas multiple serines, tyrosines and arginines that may be either phosphorylated or methylated are present between amino acids 236 and 286 in hnRNP G. Therefore, it is of great interest to further determine which amino acids that are posttranslationally modified and how these modifications affect the HPV16 RNA splicing regulatory function of hnRNP G.

We identified U2AF65 as one protein partner of trans-acting regulatory factors that regulate HPV16 early gene expression, such as hnRNP G **in Paper II** and hnRNP D40 in **Paper III**. Intriguingly, the interaction of U2AF65 and hnRNP G we demonstrated in **Paper II** directly supports previously published results that hnRNP G is part of the supraspliceosome [147]. In the future, one can investigate if posttranslational modifications of hnRNP G may affect the interactions between U2AF65 and hnRNP G. Since U2AF65 interacts with hnRNP G sequences between of amino acids 236 and 286, and these 50 amino acids that are required for the enhancement of splicing to E2 splice site SA2709 include serines, nine tyrosines and eight arginines, suggesting that the interactions of hnRNP G with U2AF65 could potentially be affected by phosphorylation and methylation of hnRNP G. Indeed, hnRNP G appear to be arginine-methylated both by PRMT1[187] and PRMT5 [188]. On the other hand, lysines may be subject to sumoylation and are all located upstream of amino acid position 236 in hnRNP G, suggesting that sumoylation may indirectly affect interactions between hnRNP G and U2AF65. Alternatively, sumoylation in the N-terminal region may interfere with other interactions between hnRNP G and the splicing machinery. Apart from affecting the

interactions between hnRNP G and U2AF65, posttranslational modifications of hnRNP G may affect interactions between hnRNP G and other proteins. Unless hnRNP G binds to the HPV16 enhancer directly, hnRNP G itself may be recruited to the enhancer by other RNA binding proteins. One candidate protein that has been repeatedly described as an hnRNP G partner is the SR-like splicing factor Tra-2β [150, 189, 190]. In the future, it will be of interest to determine if hnRNP G interacts with other proteins to regulate HPV16 mRNA splicing.

The roles of two trans-acting regulatory proteins hnRNP G and hnRNP D we identified here are both multifunctional in HPV16 gene expression and mRNA processing. Both proteins are shown the regulation of HPV16 oncogene E6/E7 gene expression or mRNAs splicing. In the future, it is great interest to further investigate the role of hnRNP G and hnRNP D in the cancer progression and the potential to be biomarker to monitor the progression of HPV16 infections to high-grade lesion or cancer, or to be targets for HPV16-infective lesion treatment.

Our results suggest that posttranslational modification of hnRNP G affect its splicing regulatory activity. These are of particular interest since various inhibitors of the enzymes that modify hnRNP G could potentially be used to alter HPV16 E6/E7 or E1/E2 mRNA splicing. These inhibits could potentially be used to treat HPV16 infects or HPV16 induced cancer.

The two HPV16 subgenomic reporter plasmids pE2sLuc and pE1Nluc described in **Paper IV** may be used to further investigate HPV16 E1 and E2 mRNA splicing and we wish to establish bioassays to monitor HPV16 E1 or E2 mRNAs and protein levels as well as the poorly studied viral small E1C protein which produced by spliced mRNA utilizing splice site SA2582. Regarding splicing of HPV16 mRNAs from SD880 to SA2582, the significance of SA2582 is still unclear. The role of SA2582 may be to express an mRNA that produces the E1C protein. Currently, very little information is available about the E1C protein. It is unclear whether E1C is produced in HPV16-infected cells, although there is no doubt that mRNA spliced from SD880 to SA2582 is present in HPV16-infected cells. In the future, one can utilize these plasmids that we describe here to study expression of the HPV16 E1C protein and the significance and regulation of the alternative 3' splice site SA2582. Moreover, bioassays could be generated by establishing stable cell lines containing either the pE2sLuc or the pE1NLuc. Such bioassays could be used to identify molecules that may interfere with HPV16 E1 and E2 mRNA splicing. Furthermore, the bioassays could be used to identify the antiviral or antitumor drugs to HPV16-infections and HPV-driven cancer.

6. Acknowledgement

I would like to express my heartful appreciation to all the people who helped and supported my PhD study and life in Sweden.

First and foremost, I am extremely grateful to my supervisor, Professor **Stefan Schwartz**, for your patient guidance, continuous support, encouragement and invaluable advice you provided to me through my PhD study. I am always thinking how lucky I am for being your student. You are such a considerate and encouraged supervisor. You instruct my work so much and always respond to my questions so promptly. Without your teaching and help, I would not be able to finish my study.

I would also like to express my sincere thanks to my co-supervisors **Docent Joakim Esbjörnsson** and **Dr. Naoko kajitani**. Joakim, thanks for the encouragement and discussion you offered to me. You are always so nice to me. Naoko, thanks for providing so much support and guidance at my late half stage of PhD study. You show me how a good researcher should be and I have learnt so much from you. I really admire your diligent, persistent, rigorous attitude towards the work. Hope all of you have a sweet life.

In addition, I also give my gratitude to my colleagues **Johanna Jönsson**, **Xiaoxu Cui** and **Ye Zhu**. **Johanna**, thanks so much for your support and attention to help me solve problems both in project and life. I would remember the time we spent together. You are going through the most stressful stage of PhD and wish you good luck! **Xiaoxu**, thanks for being a mental model for me. You showed me how potential a person can be! Your persistence on the work encouraged me during the toughest time. Hope you have a bright future. **Ye**, you are such a nice and warm girl. You always finish your work seriously and responsibly. I really like your sweet smile. I would also like to say thanks to my previous colleagues in the lab **Yunji Zheng**, **Shirin Shoja Chaghervand** and **Chengjun Wu**. Thanks so much for all the help and support in my PhD.

Never forget those friends in Lund and Linköping: **Ping Li**, **Nannan Yao, Hannah Äbacka**, **Yusheng Yuan**, **Tao Sun**, **Hunan**, **Meiting Wang**, **Guanqun Du**, **Meng Lu** and **Ying Guo**, **Liang Zhao** and **Jiaming Hou**, **Jie Cheng**, **Yuwen Xie**. Thanks for the wonderful moment you spent with me. With special thanks to **Ping Li**, thanks for the enjoyable time you spent together with my family. Particularly, I am so appreciated the love you show to Mingche. And you are the talent for cooking I have never met in Sweden. **Nannan Yao**, thanks for your nearly five-year company. The

time we spent in Linköping is so precious to me. Wish you would meet the true love soon. **Hannah,** thanks for your help for my Swedish popular summary and kind help to look for the plate reader. I really like your sweet smile.

In addition, I would like to express my acknowledgement to China Scholarship Council (**CSC**) sponsoring my PhD study. Also, thanks for beautiful environment in **Sweden**, the life in Sweden would be a priceless treasure for my whole life.

I give my special and sincere appreciation to my husband, **Peng**. You are the light in my life. Thanks for your continued support and encouragement, without you, I could not do the best to finish my PhD on time. I really enjoy the life with you! I will be always on your side whenever happy or sad, rich or poor. To my dear baby, **Mingche**. Thanks for giving the opportunity to be your mother. Now you are three years old and are growing up day by day. Thanks to you, the life is increasingly becoming wonderful and fantastic. You also help mom to be a nicer person and understand the true meaning of ¨to be brave¨. I love you! I am so proud of both of you!

感谢我的父母！你们总是那么的支持，鼓励，认可我！在瑞典留学的日子里，无尽的是对彼此的思念和担忧！中国老话说 ¨ 父母在，不远游 ¨ ，很抱歉让你们饱受异国的苦思，希望未来的日子里可以给你们更好的生活和更多的陪伴！妈妈，特别感谢你能来瑞典陪我度过最难的时光，你的康复是我最大的心愿！同时也非常感谢我的公公和婆婆。谢谢婆婆在瑞典的一年陪伴，谢谢你对我的支持！

最后感谢自己不曾放弃！

7. References

1. Crosbie, E.J., et al., *Human papillomavirus and cervical cancer.* The Lancet, 2013. **382**(9895): p. 889-899.
2. Albert, E. and L. Laimins, *Regulation of the Human Papillomavirus Life Cycle by DNA Damage Repair Pathways and Epigenetic Factors.* Viruses, 2020. **12**(7): p. 744.
3. Burd, E.M., *Human papillomavirus and cervical cancer.* Clin Microbiol Rev, 2003. **16**(1): p. 1-17.
4. Accardi, R. and T. Gheit, *Cutaneous HPV and skin cancer.* Presse Med, 2014. **43**(12 Pt 2): p. e435-43.
5. Gheit, T., *Mucosal and Cutaneous Human Papillomavirus Infections and Cancer Biology.* Front Oncol, 2019. **9**: p. 355.
6. Schiffman, M., et al., *Carcinogenic human papillomavirus infection.* Nat Rev Dis Primers, 2016. **2**: p. 16086.
7. zur Hausen, H., *Papillomaviruses and cancer: from basic studies to clinical application.* Nat Rev Cancer, 2002. **2**(5): p. 342-50.
8. Acheson, N.H., *Foundamentals of molecular virology.* 2007: p. 264-271.
9. Petry, K.U., *HPV and cervical cancer.* Scand J Clin Lab Invest Suppl, 2014. **244**: p. 59-62; discussion 62.
10. Doorbar, J., et al., *The biology and life-cycle of human papillomaviruses.* Vaccine, 2012. **30**: p. F55-F70.
11. Moscicki, A.-B., et al., *Updating the natural history of human papillomavirus and anogenital cancers.* Vaccine, 2012. **30**: p. F24-F33.
12. Ho, G.Y., et al., *Natural history of cervicovaginal papillomavirus infection in young women.* New England Journal of Medicine, 1998. **338**(7): p. 423-428.
13. McBride, A.A., *Human papillomaviruses: Diversity, infection and host interactions.* Nature Reviews Microbiology, 2021: p. 1-14.
14. Lee, C. and L.A. Laimins, *The Differentiation-Dependent Life Cycle of Human Papillomaviruses in Keratinocytes*, in *The Papillomaviruses*, R.L. Garcea and D. DiMaio, Editors. 2007, Springer US: Boston, MA. p. 45-67.
15. Cubie, H.A., *Diseases associated with human papillomavirus infection.* Virology, 2013. **445**(1): p. 21-34.
16. Graham, S.V., *Keratinocyte Differentiation-Dependent Human Papillomavirus Gene Regulation.* Viruses, 2017. **9**(9).

17. Graham, S.V., *Human papillomavirus: gene expression, regulation and prospects for novel diagnostic methods and antiviral therapies.* Future microbiology, 2010. **5**(10): p. 1493-1506.
18. Ribeiro, A.L., A.S. Caodaglio, and L. Sichero, *Regulation of HPV transcription.* Clinics (Sao Paulo), 2018. **73**(suppl 1): p. e486s.
19. Doorbar, J., *The papillomavirus life cycle.* J Clin Virol, 2005. **32 Suppl 1**: p. S7-15.
20. Bouvard, V., et al., *A review of human carcinogens—Part B: biological agents.* The Lancet Oncology, 2009. **10**(4): p. 321-322.
21. Pastrana, D.V., et al., *Metagenomic discovery of 83 new human papillomavirus types in patients with immunodeficiency.* Msphere, 2018. **3**(6): p. e00645-18.
22. De Sanjose, S., et al., *Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study.* The lancet oncology, 2010. **11**(11): p. 1048-1056.
23. Graham, S.V. and A.A.A. Faizo, *Control of human papillomavirus gene expression by alternative splicing.* Virus research, 2017. **231**: p. 83-95.
24. Guan, P., et al., *Human papillomavirus types in 115,789 HPV-positive women: a meta-analysis from cervical infection to cancer.* Int J Cancer, 2012. **131**(10): p. 2349-59.
25. Bonelli, P., et al., *The Role of circRNAs in Human Papillomavirus (HPV)-Associated Cancers.* Cancers, 2021. **13**(5): p. 1173.
26. Muñoz, N., et al., *Epidemiologic classification of human papillomavirus types associated with cervical cancer.* New England journal of medicine, 2003. **348**(6): p. 518-527.
27. Arbyn, M., et al., *Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis.* The Lancet Global Health, 2020. **8**(2): p. e191-e203.
28. Woodman, C.B., S.I. Collins, and L.S. Young, *The natural history of cervical HPV infection: unresolved issues.* Nature Reviews Cancer, 2007. **7**(1): p. 11-22.
29. Doorbar, J., et al., *Human papillomavirus molecular biology and disease association.* Reviews in medical virology, 2015. **25**: p. 2-23.
30. So, K.A., et al., *Human papillomavirus genotype-specific risk in cervical carcinogenesis.* J Gynecol Oncol, 2019. **30**(4): p. e52.
31. *<Human Papillomavirus Type 16 Variants and Risk of Cervical Cancer.pdf>*.
32. R.W.Hendrix, S.R.C., *The Double Stranded DNA Viruses*, in *Virus Taxonomy*, C.M. Fauquet, et al., Editors. 2005, Academic Press: San Diego. p. 33-276.
33. UNSPECIFIED, e., *A review of human carcinogens: Part B: Biological agents. IARC monographs on the evaluation of carcinogenic risks to humans.* World Health Organisation, 2012: p. 265-314.
34. Pinidis, P., et al., *Human papilloma virus' life cycle and carcinogenesis.* Maedica, 2016. **11**(1): p. 48.
35. de Sanjose, S., M. Brotons, and M.A. Pavon, *The natural history of human papillomavirus infection.* Best practice & research Clinical obstetrics & gynaecology, 2018. **47**: p. 2-13.

36. Bernard, H.U., *Regulatory elements in the viral genome.* Virology, 2013. **445**(1-2): p. 197-204.
37. Avcı, A., *Genomic organization and proteins of human papillomavirus.* Mikrobiyoloji bulteni, 2012. **46**(3): p. 507-515.
38. Johansson, C. and S. Schwartz, *Regulation of human papillomavirus gene expression by splicing and polyadenylation.* Nat Rev Microbiol, 2013. **11**(4): p. 239-51.
39. Graham, S.V., *Papillomavirus 3' UTR regulatory elements.* Front. Biosci, 2008. **13**(56465663): p. 18.
40. Joyce, J.G., et al., *The L1 major capsid protein of human papillomavirus type 11 recombinant virus-like particles interacts with heparin and cell-surface glycosaminoglycans on human keratinocytes.* Journal of Biological Chemistry, 1999. **274**(9): p. 5810-5822.
41. Giroglou, T., et al., *Human papillomavirus infection requires cell surface heparan sulfate.* Journal of virology, 2001. **75**(3): p. 1565-1570.
42. Kajitani, N. and S. Schwartz, *The role of RNA-binding proteins in the processing of mRNAs produced by carcinogenic papillomaviruses.* Semin Cancer Biol, 2022.
43. McBride, A.A. and A. Warburton, *The role of integration in oncogenic progression of HPV-associated cancers.* PLoS pathogens, 2017. **13**(4): p. e1006211.
44. Ilahi, N.E. and A. Bhatti, *Impact of HPV E5 on viral life cycle via EGFR signaling.* Microb Pathog, 2020. **139**: p. 103923.
45. Huibregtse, J.M. and S.L. Beaudenon, *Mechanism of HPV E6 proteins in cellular transformation.* Seminars in Cancer Biology, 1996. **7**(6): p. 317-326.
46. Pal, A. and R. Kundu, *Human Papillomavirus E6 and E7: The Cervical Cancer Hallmarks and Targets for Therapy.* Frontiers in Microbiology, 2020. **10**.
47. Vats, A., et al., *Human papillomavirus E6 and E7: What remains?* Tumour Virus Res, 2021. **11**: p. 200213.
48. Bodily, J. and L.A. Laimins, *Persistence of human papillomavirus infection: keys to malignant progression.* Trends in Microbiology, 2011. **19**(1): p. 33-39.
49. Klingelhutz, A.J., S.A. Foster, and J.K. McDougall, *Telomerase activation by the E6 gene product of human papillomavirus type 16.* Nature, 1996. **380**(6569): p. 79-82.
50. Oh, S.T., S. Kyo, and L.A. Laimins, *Telomerase activation by human papillomavirus type 16 E6 protein: induction of human telomerase reverse transcriptase expression through Myc and GC-rich Sp1 binding sites.* J Virol, 2001. **75**(12): p. 5559-66.
51. Veldman, T., et al., *Transcriptional activation of the telomerase hTERT gene by human papillomavirus type 16 E6 oncoprotein.* J Virol, 2001. **75**(9): p. 4467-72.
52. Miller, J., et al., *HPV16 E7 protein and hTERT proteins defective for telomere maintenance cooperate to immortalize human keratinocytes.* PLoS Pathog, 2013. **9**(4): p. e1003284.
53. Howie, H.L., R.A. Katzenellenbogen, and D.A. Galloway, *Papillomavirus E6 proteins.* Virology, 2009. **384**(2): p. 324-34.
54. Vaziri, H., et al., *Loss of telomeric DNA during aging of normal and trisomy 21 human lymphocytes.* American journal of human genetics, 1993. **52**(4): p. 661.

55. Greider, C.W., *Telomere length regulation.* Annual review of biochemistry, 1996. **65**(1): p. 337-365.
56. Castro-Muñoz, L.J., et al., *The Human Papillomavirus (HPV) E1 protein regulates the expression of cellular genes involved in immune response.* Scientific Reports, 2019. **9**(1): p. 13620.
57. McBride, A.A., *The papillomavirus E2 proteins.* Virology, 2013. **445**(1-2): p. 57-79.
58. Bergvall, M., T. Melendy, and J. Archambault, *The E1 proteins.* Virology, 2013. **445**(1-2): p. 35-56.
59. McKinney, C.C., K.L. Hussmann, and A.A. McBride, *The role of the DNA damage response throughout the papillomavirus life cycle.* Viruses, 2015. **7**(5): p. 2450-2469.
60. Graham, S.V., *Human Papillomavirus E2 Protein: Linking Replication, Transcription, and RNA Processing.* J Virol, 2016. **90**(19): p. 8384-8.
61. Doorbar, J., *The E4 protein; structure, function and patterns of expression.* Virology, 2013. **445**(1-2): p. 80-98.
62. Chang, J.-L., et al., *The expression of HPV-16 E5 protein in squamous neoplastic changes in the uterine cervix.* Journal of biomedical science, 2001. **8**(2): p. 206-213.
63. Di Domenico, F., et al., *Expression of human papilloma virus type 16 E5 protein in amelanotic melanoma cells regulates endo-cellular pH and restores tyrosinase activity.* J Exp Clin Cancer Res, 2009. **28**: p. 4.
64. Liao, S., et al., *Human papillomavirus 16/18 E5 promotes cervical cancer cell proliferation, migration and invasion in vitro and accelerates tumor growth in vivo.* Oncology reports, 2013. **29**(1): p. 95-102.
65. Chen, X.S., et al., *Structure of Small Virus-like Particles Assembled from the L1 Protein of Human Papillomavirus 16.* Molecular Cell, 2000. **5**(3): p. 557-567.
66. Schwartz, S., *Regulation of human papillomavirus late gene expression.* Upsala Journal of Medical Sciences, 2000. **105**(3): p. 171-192.
67. Buck, C.B., et al., *Arrangement of L2 within the papillomavirus capsid.* J Virol, 2008. **82**(11): p. 5190-7.
68. Buck, C.B., P.M. Day, and B.L. Trus, *The papillomavirus major capsid protein L1.* Virology, 2013. **445**(1-2): p. 169-174.
69. Wang, J.W. and R.B. Roden, *L2, the minor capsid protein of papillomavirus.* Virology, 2013. **445**(1-2): p. 175-86.
70. Roden, R.B.S. and P.L. Stern, *Opportunities and challenges for human papillomavirus vaccination in cancer.* Nat Rev Cancer, 2018. **18**(4): p. 240-254.
71. Chow, L.T., T.R. Broker, and B.M. Steinberg, *The natural history of human papillomavirus infections of the mucosal epithelia.* APMIS, 2010. **118**(6-7): p. 422-49.
72. Stanley, M., *Immunobiology of HPV and HPV vaccines.* Gynecologic oncology, 2008. **109**(2): p. S15-S21.
73. Westrich, J.A., C.J. Warren, and D. Pyeon, *Evasion of host immune defenses by human papillomavirus.* Virus Research, 2017. **231**: p. 21-33.

74. Tindle, R.W., *Immune evasion in human papillomavirus-associated cervical cancer.* Nature Reviews Cancer, 2002. **2**(1): p. 59-64.
75. Sasagawa, T., H. Takagi, and S. Makinoda, *Immune responses against human papillomavirus (HPV) infection and evasion of host defense in cervical cancer.* Journal of Infection and Chemotherapy, 2012. **18**(6): p. 807-815.
76. Steinbach, A. and A.B. Riemer, *Immune evasion mechanisms of human papillomavirus: An update.* Int J Cancer, 2018. **142**(2): p. 224-229.
77. Kanodia, S., L.M. Fahey, and W.M. Kast, *Mechanisms used by human papillomaviruses to escape the host immune response.* Current cancer drug targets, 2007. **7**(1): p. 79-89.
78. Meites, E., et al., *Human papillomavirus vaccination for adults: updated recommendations of the Advisory Committee on Immunization Practices*. 2019, Wiley Online Library.
79. Moody, C., *Mechanisms by which HPV Induces a Replication Competent Environment in Differentiating Keratinocytes.* Viruses, 2017. **9**(9).
80. Ghittoni, R., et al., *The biological properties of E6 and E7 oncoproteins from human papillomaviruses.* Virus Genes, 2010. **40**(1): p. 1-13.
81. Day, D. and M.F. Tuite, *Post-transcriptional gene regulatory mechanisms in eukaryotes: an overview.* The Journal of endocrinology, 1998. **157**(3): p. 361-371.
82. Busch, A. and K.J. Hertel, *Evolution of SR protein and hnRNP splicing regulatory factors.* Wiley Interdiscip Rev RNA, 2012. **3**(1): p. 1-12.
83. Herbert, A. and A. Rich, *RNA processing and the evolution of eukaryotes.* Nature genetics, 1999. **21**(3): p. 265-269.
84. Ragheb, H.S., *Post-transcriptional modification.*
85. Crouch, R.J., *Ribosomal RNA processing in eukaryotes*, in *Processing of RNA*. 2017, CRC Press. p. 213-226.
86. Licatalosi, D.D. and R.B. Darnell, *RNA processing and its regulation: global insights into biological networks.* Nature Reviews Genetics, 2010. **11**(1): p. 75-87.
87. Schwartz, S., *Papillomavirus transcripts and posttranscriptional regulation.* Virology, 2013. **445**(1-2): p. 187-96.
88. Mole, S., et al., *Regulation of splicing-associated SR proteins by HPV-16*. 2006, Portland Press Ltd.
89. Cerasuolo, A., et al., *The role of RNA splicing factors in cancer: regulation of viral and human gene expression in human papillomavirus-related cervical cancer.* Frontiers in cell and developmental biology, 2020. **8**: p. 474.
90. Ajiro, M. and Z.-M. Zheng, *Oncogenes and RNA splicing of human tumor viruses.* Emerging microbes & infections, 2014. **3**(1): p. 1-16.
91. Nilsen, T.W. and B.R. Graveley, *Expansion of the eukaryotic proteome by alternative splicing.* Nature, 2010. **463**(7280): p. 457-463.
92. Wilkinson, M.E., C. Charenton, and K. Nagai, *RNA splicing by the spliceosome.* Annu. Rev. Biochem, 2020. **89**(1): p. 359-388.

93. Shi, Y., *Mechanistic insights into precursor messenger RNA splicing by the spliceosome.* Nature reviews Molecular cell biology, 2017. **18**(11): p. 655-670.
94. Matlin, A.J. and M.J. Moore, *Spliceosome assembly and composition.* Advances in experimental medicine and biology, 2007. **623**: p. 14-35.
95. Moore, M.J., C.C. Query, and P.A. Sharp, *Splicing of precursors to mRNAs by the spliceosome.* Cold Spring Harbor Monograph Series, 1993. **24**: p. 303-303.
96. Wahl, M.C., C.L. Will, and R. Lührmann, *The spliceosome: design principles of a dynamic RNP machine.* cell, 2009. **136**(4): p. 701-718.
97. Martinez-Contreras, R., et al., *8 hnRNP proteins and splicing control.* Advances in Experimental Medicine & Biology, 2008. **623**: p. 123.
98. Herzel, L., et al., *Splicing and transcription touch base: co-transcriptional spliceosome assembly and function.* Nature reviews Molecular cell biology, 2017. **18**(10): p. 637-650.
99. Seraphin, B. and M. Rosbash, *Identification of functional U1 snRNA-pre-mRNA complexes committed to spliceosome assembly and splicing.* Cell, 1989. **59**(2): p. 349-358.
100. Burge, C.B., T. Tuschl, and P.A. Sharp, *Splicing of precursors to mRNAs by the spliceosomes.* Cold Spring Harbor Monograph Series, 1999. **37**: p. 525-560.
101. Fu, X.D. and M. Ares, Jr., *Context-dependent control of alternative splicing by RNA-binding proteins.* Nat Rev Genet, 2014. **15**(10): p. 689-701.
102. Zheng, C.L., X.-D. Fu, and M. Gribskov, *Characteristics and regulatory elements defining constitutive splicing and different modes of alternative splicing in human and mouse.* Rna, 2005. **11**(12): p. 1777-1787.
103. Kim, E., A. Goren, and G. Ast, *Alternative splicing: current perspectives.* Bioessays, 2008. **30**(1): p. 38-47.
104. Blencowe, B.J., *Alternative splicing: new insights from global analyses.* Cell, 2006. **126**(1): p. 37-47.
105. Black, D.L., *Mechanisms of alternative pre-messenger RNA splicing.* Annual Reviews of Biochemistry, 2003. **72**: p. 291-336.
106. Wang, Z., et al., *General and specific functions of exonic splicing silencers in splicing control.* Molecular cell, 2006. **23**(1): p. 61-70.
107. Sassone-Corsi, P. and E. Borrelli, *Transcriptional regulation by trans-acting factors.* Trends in Genetics, 1986. **2**: p. 215-219.
108. Lim, K.H., et al., *Using positional distribution to identify splicing elements and predict pre-mRNA processing defects in human genes.* Proc Natl Acad Sci U S A, 2011. **108**(27): p. 11093-8.
109. Wang, Y., et al., *Mechanism of alternative splicing and its regulation.* Biomed Rep, 2015. **3**(2): p. 152-158.
110. Blencowe, B.J., *Exonic splicing enhancers: mechanism of action, diversity and role in human genetic diseases.* Trends in biochemical sciences, 2000. **25**(3): p. 106-110.
111. Wang, Z. and C.B. Burge, *Splicing regulation: from a parts list of regulatory elements to an integrated splicing code.* RNA, 2008. **14**(5): p. 802-13.

112. Fairbrother, W.G., et al., *Predictive identification of exonic splicing enhancers in human genes.* Science, 2002. **297**(5583): p. 1007-1013.
113. Stoneley, M. and A.E. Willis, *Cellular internal ribosome entry segments: structures, trans-acting factors and regulation of gene expression.* Oncogene, 2004. **23**(18): p. 3200-3207.
114. Pichon, X., et al., *RNA binding protein/RNA element interactions and the control of translation.* Current Protein and Peptide Science, 2012. **13**(4): p. 294-304.
115. Kim, M.-Y., J. Hur, and S.-J. Jeong, *Emerging roles of RNA and RNA-binding protein network in cancer cells.* BMB reports, 2009. **42**(3): p. 125-130.
116. Fay, J., et al., *Increased expression of cellular RNA-binding proteins in HPV-induced neoplasia and cervical cancer.* Journal of medical virology, 2009. **81**(5): p. 897-907.
117. Hentze, M.W., et al., *A brave new world of RNA-binding proteins.* Nature reviews Molecular cell biology, 2018. **19**(5): p. 327-341.
118. Cook, K.B., et al., *RBPDB: a database of RNA-binding specificities.* Nucleic acids research, 2010. **39**(suppl_1): p. D301-D308.
119. Han, N., W. Li, and M. Zhang, *The function of the RNA-binding protein hnRNP in cancer metastasis.* Journal of cancer research and therapeutics, 2013. **9**(7): p. 129.
120. Graveley, B.R., *Sorting out the complexity of SR protein functions.* Rna, 2000. **6**(9): p. 1197-1211.
121. Cáceres, J.F., et al., *Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity.* The Journal of cell biology, 1997. **138**(2): p. 225-238.
122. Manley, J.L., et al., *SR proteins and splicing control 1569.* Genes Dev, 1996.
123. Kelemen, O., et al., *Function of alternative splicing.* Gene, 2013. **514**(1): p. 1-30.
124. Mo, L., et al., *An analysis of the role of HnRNP C dysregulation in cancers.* Biomarker Research, 2022. **10**(1): p. 1-10.
125. Lin, S. and X.-D. Fu, *SR proteins and related factors in alternative splicing.* Advances in experimental medicine and biology, 2007. **623**: p. 107-122.
126. Zahler, A.M., et al., *Distinct functions of SR proteins in alternative pre-mRNA splicing.* Science, 1993. **260**(5105): p. 219-222.
127. Cartegni, L. and A.R. Krainer, *Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1.* Nature genetics, 2002. **30**(4): p. 377-384.
128. Simard, M.J. and B. Chabot, *SRp30c is a repressor of 3' splice site utilization.* Mol Cell Biol, 2002. **22**(12): p. 4001-10.
129. Chen, Q., et al., *Functions of heterogeneous nuclear ribonucleoproteins in stem cell potency and differentiation.* Biomed Res Int, 2013. **2013**: p. 623978.
130. Krecic, A.M. and M.S. Swanson, *hnRNP complexes: composition, structure, and function.* Current opinion in cell biology, 1999. **11**(3): p. 363-371.
131. Geuens, T., D. Bouhy, and V. Timmerman, *The hnRNP family: insights into their role in health and disease.* Human genetics, 2016. **135**(8): p. 851-867.

132. Liu, Y. and S.L. Shi, *The roles of hnRNP A2/B1 in RNA biology and disease.* Wiley Interdisciplinary Reviews: RNA, 2021. **12**(2): p. e1612.
133. Matlin, A.J., F. Clark, and C.W. Smith, *Understanding alternative splicing: towards a cellular code.* Nat Rev Mol Cell Biol, 2005. **6**(5): p. 386-98.
134. Nasim, M.T., et al., *HnRNP G and Tra2β: opposite effects on splicing matched by antagonism in RNA binding.* Human molecular genetics, 2003. **12**(11): p. 1337-1348.
135. Dreyfuss, G., et al., *hnRNP proteins and the biogenesis of mRNA.* Annual review of biochemistry, 1993. **62**: p. 289-289.
136. Xu, N., C.-Y.A. Chen, and A.-B. Shyu, *Versatile role for hnRNP D isoforms in the differential regulation of cytoplasmic mRNA turnover.* Molecular and cellular biology, 2001. **21**(20): p. 6960-6971.
137. Loflin, P., C.-Y.A. Chen, and A.-B. Shyu, *Unraveling a cytoplasmic role for hnRNP D in the in vivo mRNA destabilization directed by the AU-rich element.* Genes & development, 1999. **13**(14): p. 1884-1897.
138. White, E.J., G. Brewer, and G.M. Wilson, *Post-transcriptional control of gene expression by AUF1: mechanisms, physiological targets, and regulation.* Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms, 2013. **1829**(6-7): p. 680-688.
139. Kiledjian, M., et al., *Identification of AUF1 (heterogeneous nuclear ribonucleoprotein D) as a component of the alpha-globin mRNA stability complex.* Molecular and cellular biology, 1997. **17**: p. 4870 - 4876.
140. DeMaria, C.T., et al., *Structural determinants in AUF1 required for high affinity binding to A+ U-rich elements.* Journal of Biological Chemistry, 1997. **272**(44): p. 27635-27643.
141. Fialcowitz, E.J., et al., *A hairpin-like structure within an AU-rich mRNA-destabilizing element regulates trans-factor binding selectivity and mRNA decay kinetics.* J Biol Chem, 2005. **280**(23): p. 22406-17.
142. Yoon, J.H., et al., *PAR-CLIP analysis uncovers AUF1 impact on target RNA fate and genome integrity.* Nat Commun, 2014. **5**: p. 5248.
143. Geuens, T., D. Bouhy, and V. Timmerman, *The hnRNP family: insights into their role in health and disease.* Hum Genet, 2016. **135**(8): p. 851-67.
144. Shin, K.-H., et al., *hnRNP G elicits tumor-suppressive activity in part by upregulating the expression of Txnip.* Biochemical and biophysical research communications, 2008. **372**(4): p. 880-885.
145. Cho, Y., et al., *RBMX is a component of the centromere noncoding RNP complex involved in cohesion regulation.* Genes Cells, 2018. **23**(3): p. 172-184.
146. Moursy, A., F.H. Allain, and A. Clery, *Characterization of the RNA recognition mode of hnRNP G extends its role in SMN2 splicing regulation.* Nucleic Acids Res, 2014. **42**(10): p. 6659-72.
147. Sebbag-Sznajder, N., et al., *Regulation of alternative splicing within the supraspliceosome.* J Struct Biol, 2012. **177**(1): p. 152-9.

148. Heinrich, B., et al., *Heterogeneous nuclear ribonucleoprotein G regulates splice site selection by binding to CC(A/C)-rich regions in pre-mRNA.* J Biol Chem, 2009. **284**(21): p. 14303-15.
149. Ma, L., et al., *X-Linked RNA-Binding Motif Protein Modulates HIV-1 Infection of CD4(+) T Cells by Maintaining the Trimethylation of Histone H3 Lysine 9 at the Downstream Region of the 5' Long Terminal Repeat of HIV Proviral DNA.* mBio, 2020. **11**(2).
150. Nasim, M.T., et al., *HnRNP G and Tra2beta: opposite effects on splicing matched by antagonism in RNA binding.* Hum Mol Genet, 2003. **12**(11): p. 1337-48.
151. Takemoto, T., et al., *RBMX is a novel hepatic transcriptional regulator of SREBP-1c gene response to high-fructose diet.* FEBS letters, 2007. **581**(2): p. 218-222.
152. Elliott, D.J., et al., *RBMX family proteins connect the fields of nuclear RNA processing, disease and sex chromosome biology.* The international journal of biochemistry & cell biology, 2019. **108**: p. 1-6.
153. Adamson, B., et al., *A genome-wide homologous recombination screen identifies the RNA-binding protein RBMX as a component of the DNA-damage response.* Nature cell biology, 2012. **14**(3): p. 318-328.
154. Kai, M., *Roles of RNA-Binding Proteins in DNA Damage Response.* Int J Mol Sci, 2016. **17**(3): p. 310.
155. Wu, C., N. Kajitani, and S. Schwartz, *Splicing and Polyadenylation of Human Papillomavirus Type 16 mRNAs.* Int J Mol Sci, 2017. **18**(2).
156. Ajiro, M. and Z.-M. Zheng, *E6^ E7, a novel splice isoform protein of human papillomavirus 16, stabilizes viral E6 and E7 oncoproteins via HSP90 and GRP78.* MBio, 2015. **6**(1): p. e02068-14.
157. Pim, D., V. Tomaic, and L. Banks, *The human papillomavirus (HPV) E6* proteins from high-risk, mucosal HPVs can direct degradation of cellular proteins in the absence of full-length E6 protein.* Journal of virology, 2009. **83**(19): p. 9863-9874.
158. Zheng, Z.-M., et al., *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.* Journal of molecular biology, 2004. **337**(5): p. 1091-1108.
159. Zheng, Y., et al., *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.* Journal of Virology, 2020. **94**(20): p. e01509-20.
160. Tang, S., et al., *The E7 oncoprotein is translated from spliced E6* I transcripts in high-risk human papillomavirus type 16-or type 18-positive cervical cancer cell lines via translation reinitiation.* Journal of virology, 2006. **80**(9): p. 4249-4263.
161. Stacey, S.N., et al., *Leaky scanning is the predominant mechanism for translation of human papillomavirus type 16 E7 oncoprotein from E6/E7 bicistronic mRNA.* Journal of virology, 2000. **74**(16): p. 7284-7297.
162. Zheng, Y., et al., *Efficient production of hpv16 e2 protein from hpv16 late mrnas spliced from sd880 to sa2709.* Virus Research, 2020. **285**: p. 198004.

163. Jackson, S.P. and J. Bartek, *The DNA-damage response in human biology and disease.* Nature, 2009. **461**(7267): p. 1071-1078.
164. Ciccia, A. and S.J. Elledge, *The DNA damage response: making it safe to play with knives.* Molecular cell, 2010. **40**(2): p. 179-204.
165. Nakad, R. and B. Schumacher, *DNA Damage Response and Immune Defense: Links and Mechanisms.* Frontiers in Genetics, 2016. **7**.
166. Welsh, C., et al., *Reduced levels of XPA, ERCC1 and XPF DNA repair proteins in testis tumor cell lines.* International journal of cancer, 2004. **110**(3): p. 352-361.
167. Pan, M.-R., et al., *Connecting the dots: from DNA damage and repair to aging.* International journal of molecular sciences, 2016. **17**(5): p. 685.
168. Mendoza, J., et al., *Association between ERCC1 and XPA expression and polymorphisms and the response to cisplatin in testicular germ cell tumours.* British journal of cancer, 2013. **109**(1): p. 68-75.
169. Anacker, D.C. and C.A. Moody, *Modulation of the DNA damage response during the life cycle of human papillomaviruses.* Virus research, 2017. **231**: p. 41-49.
170. Abraham, R.T., *Cell cycle checkpoint signaling through the ATM and ATR kinases.* Genes & development, 2001. **15**(17): p. 2177-2196.
171. Huang, R.-X. and P.-K. Zhou, *DNA damage response signaling pathways and targets for radiotherapy sensitization in cancer.* Signal Transduction and Targeted Therapy, 2020. **5**(1): p. 60.
172. Mehta, K. and L. Laimins, *Human papillomaviruses preferentially recruit DNA repair factors to viral genomes for rapid repair and amplification.* MBio, 2018. **9**(1): p. e00064-18.
173. Bristol, M.L., D. Das, and I.M. Morgan, *Why Human Papillomaviruses Activate the DNA Damage Response (DDR) and How Cellular and Viral Replication Persists in the Presence of DDR Signaling.* Viruses, 2017. **9**(10).
174. Spriggs, C.C. and L.A. Laimins, *Human Papillomavirus and the DNA Damage Response: Exploiting Host Repair Pathways for Viral Replication.* Viruses, 2017. **9**(8): p. 232.
175. Nilsson, K., C. Wu, and S. Schwartz, *Role of the DNA Damage Response in Human Papillomavirus RNA Splicing and Polyadenylation.* Int J Mol Sci, 2018. **19**(6).
176. Abjaude, W., et al., *ATM Pathway Is Essential for HPV–Positive Human Cervical Cancer-Derived Cell Lines Viability and Proliferation.* Pathogens, 2022. **11**(6): p. 637.
177. Lenzken, S.C., A. Loffreda, and S.M. Barabino, *RNA splicing: a new player in the DNA damage response.* Int J Cell Biol, 2013. **2013**: p. 153634.
178. Nilsson, K., et al., *The DNA damage response activates HPV16 late gene expression at the level of RNA processing.* Nucleic Acids Res, 2018. **46**(10): p. 5029-5049.
179. Shkreta, L. and B. Chabot, *The RNA splicing response to DNA damage.* Biomolecules, 2015. **5**(4): p. 2935-2977.
180. Zhao, X., M. Rush, and S. Schwartz, *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**(20): p. 10888-905.

181. Li, X., et al., *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 research, 2013. **41**(22): p. 10488-10508.
182. Li, X., et al., *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**(9): p. e72776.
183. Yu, H., et al., *hnRNP G prevents inclusion on the HPV16 L1 mRNAs of the central exon between splice sites SA3358 and SD3632.* Journal of General Virology, 2018. **99**(3): p. 328-343.
184. Collier, B., et al., *Specific inactivation of inhibitory sequences in the 5' end of the human papillomavirus type 16 L1 open reading frame results in production of high levels of L1 protein in human epithelial cells.* Journal of virology, 2002. **76**(6): p. 2739-2752.
185. Forslund, O., et al., *A novel human in vitro papillomavirus type 16 positive tonsil cancer cell line with high sensitivity to radiation and cisplatin.* BMC cancer, 2019. **19**(1): p. 1-13.
186. Johansson, C., et al., *Acetylation of intragenic histones on HPV16 correlates with enhanced HPV16 gene expression.* Virology, 2015. **482**: p. 244-259.
187. Hartel, N.G., et al., *Deep protein methylation profiling by combined chemical and immunoaffinity approaches reveals novel PRMT1 targets.* Molecular & Cellular Proteomics, 2019. **18**(11): p. 2149-2164.
188. Cai, T., et al., *Deletion of RBMX RGG/RG motif in Shashi-XLID syndrome leads to aberrant p53 activation and neuronal differentiation defects.* Cell reports, 2021. **36**(2): p. 109337.
189. Ouyang, Y.Q., et al., *Expression levels of hnRNP G and hTra2-beta1 correlate with opposite outcomes in endometrial cancer biology.* Int J Cancer, 2011. **128**(9): p. 2010-9.
190. Hirschfeld, M., et al., *HNRNP G and HTRA2-BETA1 regulate estrogen receptor alpha expression with potential impact on endometrial cancer.* BMC Cancer, 2015. **15**(1): p. 86.

Printed by Media-Tryck, Lund 2022 NORDIC SWAN ECOLABEL 3041 0903

LUND UNIVERSITY

FACULTY OF MEDICINE

Department of Laboratory Medicine

Lund University, Faculty of Medicine
Doctoral Dissertation Series 2022:102
ISBN 978-91-8021-263-2
ISSN 1652-8220