Virus Research 166 (2012) 43–53

Contents lists available at SciVerse ScienceDirect

Virus Research

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

A few nucleotide polymorphisms are sufficient to recruit nuclear factors differentially to the intron 1 of HPV-16 intratypic variants

Eduardo López-Urrutia[a], Jesús Valdés[b], Raúl Bonilla-Moreno[a], Martha Martínez-Salazar[c], Martha Martínez-Garcia[d], Jaime Berumen[e,f], Nicolás Villegas-Sepúlveda[a,*]

[a] Depto. Biomedicina Molecular, Centro de Investigación y de Estudios Avanzados-IPN (CINVESTAV-IPN), Unidad Zacatenco, México, D.F. 07360, Mexico
[b] Depto. Bioquímica, Centro de Investigación y de Estudios Avanzados-IPN (CINVESTAV-IPN), Unidad Zacatenco, México, D.F. 07360, Mexico
[c] Subdirección de Investigación básica, Instituto Nacional de Cancerología, Mexico
[d] FES Iztacala UNAM Estado de México, Mexico
[e] Facultad de Medicina, UNAM. 04510, México, D.F., Mexico
[f] Unidad de Medicina Genómica, Hospital General, México, D.F., Mexico

ARTICLE INFO

Article history:
Received 21 October 2011
Received in revised form 28 February 2012
Accepted 28 February 2012
Available online 8 March 2012

Keywords:
HPV-16 variants
SNPs
E6/E7 splicing
C33-A cells
Splicing factors

ABSTRACT

The HPV-16 E6/E7 genes, which contain intron 1, are processed by alternative splicing and its transcripts are detected with a heterogeneous profile in tumours cells. Frequently, the HPV-16 positive carcinoma cells bear viral variants that contain single nucleotide polymorphisms into its DNA sequence. We were interested in analysing the contribution of this polymorphism to the heterogeneity in the pattern of the E6/E7 spliced transcripts. Using the E6/E7 sequences from three closely related HPV-16 variants, we have shown that a few nucleotide changes are sufficient to produce heterogeneity in the splicing profile. Furthermore, using mutants that contained a single SNP, we also showed that one nucleotide change was sufficient to reproduce the heterogeneous splicing profile. Additionally, a difference of two or three SNPs among these viral sequences was sufficient to recruit differentially several splicing factors to the polymorphic E6/E7 transcripts. Moreover, only one SNP was sufficient to alter the binding site of at least one splicing factor, changing the ability of splicing factors to bind the transcript. Finally, the factors that were differentially bound to the short form of intron 1 of one of these E6/E7 variants were identified as TIA1 and/or TIAR and U1-70k, while U2AF65, U5-52k and PTB were preferentially bound to the transcript of the other variants.

1. Introduction

Human papillomaviruses (HPV) are small DNA viruses that infect mucosal and epithelial tissues (zur Hausen, 2008). They are classified as different HPV types if the DNA sequences of the L1 gene vary by more than 10%, as subtypes if they vary between 2 and 10% and as intratypic variants when differences are less than 2% (reviewed in de Villiers et al., 2004). They are also grouped as low- or high-risk (h) viruses according to the frequency in which they are detected in tumour samples. Although almost two hundred HPV types have been described, only 30–40 of them are considered high risk agents of cervical and oral cancers. Strikingly, HPV type 16 (HPV-16) is detected in 50–60% of cervical cancer samples. As a result, HPV-16 is considered to be the major risk agent for cervical cancer (Walboomers et al., 1999). Additionally, hrHPV variants are grouped based on the geographical origin of their isolates and are designated as: European (E) or Asian (As) variants, and so on. HPV-16 variants differ in a number of viral genes, including, but not limited to, E6, E7, E2 and L2. HPV-16 E6/E7 variants have been described in great detail (Ho et al., 1991; Chan et al., 1992). The expression of the E6 and E7 genes from hrHPV induces cellular immortalisation and transformation (Münger et al., 1989; Hawley-Nelson et al., 1989). Therefore, the polymorphic E6/E7 genes might differ in their oncogenic potential (Lichtig et al., 2006) or their capacity to deregulate human keratinocyte apoptosis or differentiation (Zehbe et al., 2009). Thus, the HPV-16 variants may differ in their risks of producing high-grade cervical intraepithelial neoplasias (CIN) or cervical cancer (Xi et al., 2007; Zuna et al., 2009). HPV-16 E6/E7 variants are designated by adding one of the single nucleotide polymorphism (SNP) changes to the identification of the regional origin. For example, the short name of one variant of the European prototype is HPV-16 E-G350 (Zehbe et al., 1998a), with "G350" indicating the SNP change (T350G) with respect to the European prototype HPV-16 sequence (HPV-16 R) that is used as a reference (Yamada et al., 1997). Some of the SNPs in these variants

* Corresponding author at: Av. IPN # 2508, Zacatenco, México City, Apdo. Postal 14-740, CP. 07360, Mexico. Tel.: +52 55 5747 3800x5023; fax: +52 55 5747 3938.
E-mail address: nvillega@cinvestav.mx (N. Villegas-Sepúlveda).

0168-1702/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.virusres.2012.02.026

cause amino acid changes that affect the biological functions of the E6 oncoprotein (Stöppler et al., 1996; Asadurian et al., 2007; Lichtig et al., 2006). So, variants are more prevalent in carcinomas than HPV-16R, with the most prevalent being HPV-16 E-G350 (Zehbe et al., 1998b).

The HPV-16 E6/E7 genes, which contain intron 1, are processed by alternative splicing (AS). The products of the E6/E7 genes are widely observed in cervical tumour cells (Baker et al., 1987; Smotkin et al., 1989), but the E6/E7 transcripts are expressed with heterogeneous profiles in most of the tumours and cervical cancer-derived cell lines (McNicol et al., 1995; De la Rosa-Ríos et al., 2006). Intron 1 is processed by the cellular splicing machinery and produces four transcripts. During splicing, introns are removed by a two-step mechanism catalysed by a large RNA and multi-protein complex known as the spliceosome (reviewed in Sharp, 1994). This multi-protein complex is comprised of five small nuclear ribonucleoprotein complexes (U1-, U2-, U4-, U5- and U6 snRNPs) and at least one hundred accessory or auxiliary proteins that are involved in altering positively or negatively the recognition of sequences that are important for splicing. Interestingly, the mechanism by which the intron 1 is alternatively spliced is special because a remnant of the unspliced E6/E7 pre-mRNA (E6/E7$^{\text{pre-mRNA}}$) is expressed as one of the mature transcripts. A similar mechanism has been observed in HIV and herpesviruses (Dyhr-Mikkelsen and Kjems, 1995; Zabolontny et al., 1997). Such a mechanism may be essential to preserve the coding potential of the viral primary transcript. As is the case with other small viruses, HPV-16 intron 1 has a complex structure, which is comprised of a single splicing donor site (SD), one single branch point sequence (BPs) and three alternative acceptor sites (SA-A to C). The three SA sites yield three alternative transcripts in addition to the unspliced E6/E7$^{\text{pre-mRNA}}$, including E6*I/E7, E6*II/E7 and E6Ê7 (Milligan et al., 2007; Del Moral-Hernández et al., 2010), increasing the E6/E7 coding potential from three to five proteins, with respect to other hrHPVs. The mechanism by which heterogeneity in the expression of the HPV-16 E6/E7 alternative transcripts is produced is not yet understood; however, evidence from our lab indicates that such heterogeneity is produced by the differential use of acceptor sites within intron 1 (Del Moral-Hernández et al., 2010). Although the exact mechanism for the selection of acceptor sites in intron 1 is widely unknown, mainly because most of the spliceosome accessory factors involved in recognition and regulation of the different acceptor sites are largely unknown; the sequence changes observed in the HPV-16 variants may also contribute to the production of a differential splicing profile. Therefore, it is interesting to analyse the impact of single nucleotide polymorphisms (SNPs) within the E6/E7 genes from HPV-16 variants, on the AS of intron 1. The SNP within intron 1 that have been characterised so far are not located on splice sites. Conversely, these changes are localised in sequences that are upstream from the donor and acceptor sites. Therefore, some of the SNPs in the HPV-16 isolates could lie in sequences, which may affect the selection of these splicing sites. As a result, these few SNPs on the E6/E7 splicing profile merit careful analysis.

In this paper, we used three closely related HPV-16 E6/E7 sequences to investigate whether the nucleotide changes within variants were enough to produce a differential E6/E7 splicing profile. The E6/E7 genes were isolated from HPV-16 genomes present in CaSki and SiHa cells or the HPV-16R. Although the former genomes differ from the HPV-16R in 39 and 35 nucleotide changes, respectively, only 4 of the polymorphic sites lie within the E6/E7 regions. C33-A cells were stably transfected with the E6/E7 sequences derived from the three variants, and the results revealed that the E6/E7 variants displayed a differential splicing profile. These data suggest that a few polymorphic changes were sufficient to produce a differential splicing pattern.

RNA–protein, UV cross-linking experiments using HeLa nuclear extracts (NE) also revealed that the E6/E7 transcripts from the genetic variants differentially recruited a few nuclear factors. Moreover, using short E6/E7 RNA probes containing each of the SNPs, we observed that SNPs lying close to the acceptor sites produced differential binding of specific nuclear factors. C33-A cells were stably transfected with E6/E7R mutants containing single or double sequence changes to further analyse the effect of single nucleotide polymorphisms within E6/E7 on the AS of this intron. The results revealed that these mutants also produced a differential splicing profile. The results indicate that one SNP was sufficient to produce a differential splicing profile.

Finally, as an initial approach to identify some of the nuclear factors involved in the production of the heterogeneous profile and likely involved in the regulation of AS of intron 1, the proteins bound to the shortest RNA probe of two variants of intron 1 were purified by RNA-affinity chromatography. To this end, two polymorphic E6/E7 RNA probes differing in a single nucleotide (U350G) were used. The results verify the differential recruiting of at least one nuclear factor to the sequence surrounding nt 350. The purified proteins were identified in batch by LC–MS/MS. The TIA1 and TIAR (TIAL) splicing factors were identified as one of the proteins that bind the sequence surrounding the U350 in the E6/E7R transcript. Taken together, the data showed that a single nucleotide change was sufficient to differentially recruit a few auxiliary splicing factors to the intron 1 variants. The implication of the SNPs and some of the identified regulatory splicing factors and its probable role in the AS of the HPV-16 variants are discussed.

2. Materials and methods

2.1. Constructs and site-directed mutagenesis

HPV-16 E6/E7 amplicons were obtained by RT-PCR from the CaSki (ATCC, CRL1550) and SiHa (ATCC, HTB35) cell lines (GenBank ID U89348 and AF001599, respectively) or from the cloned HPV-16 reference genome (GenBank ID NC_001526). The bicistronic amplicons from each variant were cloned into the *HindIII/BamHI* restriction sites of the pEGFP N1 vector, as previously reported for the E6/E7 cassette derived from CaSki cells (Del Moral-Hernández et al., 2010). Mutants were made using the HPV-16R prototype (pE6/E7R-GFP) and the QuikChange site-directed mutagenesis kit (Stratagene). To produce the pT350G and pA442C mutants, the primer pairs T350GS/T350GAS or A442CS/A442CAS were used. The pT350G/A442C mutant was obtained by performing the A442C mutation on pT350G. All constructs were verified by sequencing using the Big Dye Terminator Ready Reaction Kit and analysed in an ABI PRISM 3100 Genetic Analyser System (Perkin Elmer, Branchburg, NJ). All of the primers were purchased from Invitrogen.

2.2. Transfection of C33-A cells and expression analysis

The HPV negative (HPV−) cervical carcinoma-derived cells, C33-A (ATCC, HTB-31), were stably transfected with each of the E6/E7 variants or mutants. Cells were grown in MEM supplemented with 10% FBS and transfected using the $CaCl_2$ method, as previously reported (Del Moral-Hernández et al., 2010). Stable transfectants were produced at least three times for each of E6/E7 constructs and five for the SNP mutants. Transfected cells could be cultured during a six-month period without changes. However, after two months of continuous G418 selection, transfected cells were aliquoted, frozen and kept at −80 °C to avoid long-term effects. We did not observe differences between any of the transfected cells. Total RNA was extracted from confluent cell cultures using TRIzol reagent (Gibco, BRL) following the manufacturer's recommendations and

Table 1
Nucleotide sequence of oligonucleotide primers. Oligonucleotide primers indicating the nucleotide position as numbered in the HPV-16 sequence. Lower case letters show the sequence of the T7 RNA polymerase promoter sequence included in the oligonucleotide primers used to produce the DNA amplicons.

No name	Sequence (5′–3′)	nt position
Primers for UV cross-linking or affinity capture:		
1. E6/E7 Fw	taatacgactcactataggATGCACCAAAAGAGAACT	83–100
2. E6/E7 Rv	TGGTTTCTGAGAACAGATGGG	855–835
3. SD Rv	ATCACATACAGCATATGGATTCCCATCTC	295–267
4. SA-A Fw	taatacgactcactataggACTTTGCTTTTCGGGATTT	234–252
5. SA-A Rv	CAGGACACAGTGGCTTTTG	440–422
6. SA-B Fw	taatacgactcactatagCTGTCAAAAGCCACTGTGTC	418–437
7. SA-B Rv	CAGCTGGGTTTCTCTACGTG	556–537
8. SA-C Fw	taatacgactcactataggGATCTCTACTGTTATGAGC	622–640
9. SA-C Rv	CATTAACAGGTCTTCCAAAGTACG	813–790
10. SA-A Ext Fw	taatacgactcactataggAACAGTTACTGCGACGTGAGGTATATG	207–233
11. SA-A Bio	ACCTCACGTCGTCGCAGTAACTGTttt-Biotin	228–208
Primers for site-directed mutagenesis:		
12. T350GS	TATTCTAAAATTAGTGAGTATAGACATTATTGTTATAGTgTGTATGGAACAACATTAGAA	311–370
13. T350G AS	TTCTAATGTTGTTCCATACAcACTATAACAATAATGTCTATACTCACTAATTTTAGAATA	370–311
14. A442C S	AAAGCCACTGTGTCCTGAcGAAAAGCAAAGACATCTG	424–460
15. A442C AS	CAGATGTCTTTGCTTTTCgTCAGGACACAGTGGCTTT	460–424
Primers for RT PCR		
16. E6-begin	ATGCACCAAAAGAGAACTGCA	83–103
17. E7-end	TGGTTTCTGAGAACAGATGGG	855–835
18. Actin HF[a]	GAAGAGCTACGAGCTGCCTGACG	795–817
19. Actin HR[a]	CACACGGAGTACTTGCGCTC	1103–1084
Primers for nested PCR:		
20. HPVRTFw	ACAGGAGCGACCCAGAAAGTT	118–138
21. HPVRTRw	TGCCCATTAACAGGTCTTCCA	817–797

[a] As numbered at the actin cDNA sequence (GenBank™ accession number NM_001101.3).

subsequently treated with DNase I (DNAfree, Ambion, Inc.). E6/E7 spliced transcripts were detected by RT-PCR, as previously described (De la Rosa-Ríos et al., 2006). Amplicons were resolved by electrophoresis in ethidium bromide-stained agarose gels. Most of the RT-PCR amplicons from transfected cells were cloned in T-protruding pBS SKII+. All the E6/E7 splicing products were identified by colony-PCR, after analysis of 50 clones from each ligation reaction, excepting for E6*I/E7C and E6/E7(S). Subsequently, plasmid for each DNA was prepared and sequenced in both chains. The spliced products from all clones corresponded to their respective E6/E7 variant.

Western blot assays were performed as previously described (Del Moral-Hernández et al., 2010). Membranes were probed overnight using a 1:500 (v/v) dilution of the appropriate antibody. For detection, 1:5000 (v/v) dilutions of HRP anti-rabbit or anti-mouse conjugate antibodies (Zymed) were used. Finally, immune complexes were detected with the SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific). The following antibodies (from Santa Cruz Biotechnology) were used for immuno-detection of splicing factors: anti-U2AF65, anti-hnRNP A1, anti-U70k, anti-hnRNP C1/C2, anti-PTB, anti-SF3B4, anti-TIAR and anti-Sam68.

2.3. RNA–protein, UV cross-linking assays and RNA secondary structure prediction and analysis

RNA probes were produced from DNA amplicons obtained by PCR from each of the E6/E7 variants (nt 83 to nt 855). Amplicons for splicing donor (SD) probes were generated using primers E6/E7 Fw and SD Rv. The acceptor probes were generated with the following primers: SD-A Fw and SD-A Rv for A acceptor; SA-B Fw and SA-B Rv for B acceptor; and SA-C Fw and SA-C Rv for C acceptor (described in Table 1). Uniformly labelled probes were obtained by *in vitro* transcription using the Riboprobe transcription system (Promega), 20 μCi of [α-^{32}P] UTP, and 100 ng of E6/E7 DNA amplicon as template. UV cross-linking assays were carried out as previously described (Del Moral-Hernández et al., 2010). *In silico* RNA secondary structure prediction for the three E6/E7 pre-mRNAs was performed with the Mfold software by using the Rensselaer bioinformatics web server (Zuker, 2003), which was accessed through http://mfold.bioinfo.rpi.edu, Rennselaer Polytechnic Institute. The RNA secondary structure was verified by digesting the RNAs with RNases. Briefly, RNA probes were transcribed using the MEGAScript T7 kit (Ambion) and 1 μg E6/E7 DNA amplicons. The RNAs were then 3′-end labelled using T4 RNA ligase (Ambion) and 20 μCi of pCp [α-^{32}P]-CMP. The RNAs were extensively purified on 8% polyacrylamide gels and extracted using the 'crush and soak' method. For each assay, 125 000 cpm of each probe was heated at 70 °C for 7 min and cooled to 37 °C for 30 min. Then either 1 mU of RNase T1, 1 μl of 10 fg/μl RNase A or 0.4 mU of RNAse V1 (Ambion) was added to the probes and incubated at 37 °C for 5 min. Reactions were halted by adding a mixture of 90% urea, 10% acetic acid, 0.005% bromophenol blue and 0.005% xylene–cyanol. Digestion products were resolved on 20% polyacrylamide and 8 M urea gels, and RNAs were visualised by autoradiography.

2.4. Isolation and identification of proteins bound to E6/E7 RNAs

Proteins were purified using 214 nt RNA probes from intron 1 of each HPV-16 variant. Probes were hybridised to a small, 3′-end-biotinylated DNA oligonucleotide complementary to the first 27 nts of each RNA probe (Harris et al., 2006). Briefly, 500 pmol of each biotinylated oligonucleotide (Table 1) were mixed with 500 pmol of the corresponding RNA probe dissolved in 10 mM HEPES (pH = 7.0), 50 mM KCl, 10% glycerol, 1 mM EDTA, 1 mM DTT, 0.5% Triton X-100 and 0.15 μg/ml of yeast tRNA supplemented with RNAsin (Promega Co) and protease inhibitors (Complete, Roche). The reaction mixtures were heated at 95 °C for 2 min, and the annealing of the hybrid DNA/RNA was allowed to proceed for 1 h at 37 °C with gentle agitation. Subsequently, a mixture of 500 μg of HeLa NE and 100 μl of magnetic beads (μMACS Streptavidin MicroBeads; Miltenyi Biotec) were added, in batch, and incubated for 1 h at 4 °C. The mixture was passed through a μColumn (Miltenyi Biotec) and subjected to a magnetic field to capture the proteins that were bound to the RNA/DNA hybrids. The batch mixture was washed four times with the HEPES buffer. Proteins were eluted from the column using the above HEPES buffer supplemented with 1 M NaCl. The eluted

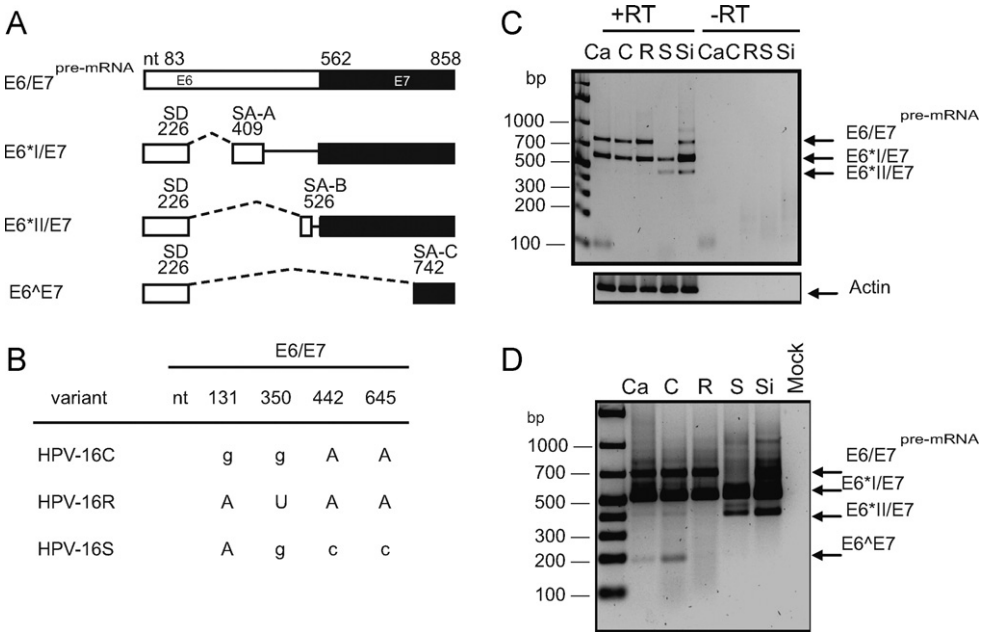

Fig. 1. Transfection of E6/E7 genes from polymorphic HPV-16 variants produces a heterogeneous pattern of spliced transcripts. (A) A schematic representation of the mature transcripts that are expressed from the HPV-16 E6/E7 genes. Dashed lines represent the three alternative ways of processing intron 1. Numbers above the E6/E7$^{pre\text{-}mRNA}$ indicate the beginning and end of the E6/E7 cassette and the splice donor (SD) or acceptor sites (SA-A, SA-B or SA-C). (B) Single nucleotide variation among the three E6/E7 sequences is shown. (C) Pattern of the alternative E6/E7 transcripts obtained by RT-PCR from CaSki (Ca) or SiHa cells (Si), and C33-A cells transfected with the sequences from the three variants: CaSki (C) Reference (R) or SiHa (S). (D) Nested PCR from C33-A cells transfected with the sequences from the three variants: CaSki (C), Reference (R) or SiHa (S). Actin was used as loading control. The nested PCR shows the presence of the minor E6*II transcript. C33-A cells transfected with the pEGFP-N1 empty plasmid were used as mock cells.

proteins were resolved by 12% SDS-PAGE and stained with silver using the mass spectrometry-compatible SilverQuest™ kit (Invitrogen). Proteins were identified by MS/MS analysis in the Columbia Protein Core Facility (Columbia University College of Physicians & Surgeons). Briefly, tryptic peptides were extracted from gel slices and injected into a reverse-phase LC column for stratification. Proteins were identified by ESI–LC–MS/MS analysis. Data files were processed using the MassLynx ProteinLynx software, and PkI data were submitted to www.matrixscience.com for analysis using the Mascot algorithm. All of the RNA probes used for protein purification were similar to those used for UV cross-linking but with an extended upstream region of 27 nts. We purchased biotinylated DNA oligonucleotides that were complementary to the extended region of the RNA probes that were mentioned above. The SA-AExt (extended) probe was obtained using the SA-AExtFw and the SA-A Rv primers (see Table 1).

3. Results

3.1. HPV-16 genetic variation contributes to produce heterogeneity of E6/E7 spliced transcripts

The HPV-16 E6/E7 genes produce four mature transcripts, three of which are spliced and one of which is the unspliced pre-mRNA (Fig. 1). Previous work has shown that the profile of the four transcripts is heterogeneous; however, the mechanism that produces the transcript variation is not fully understood. To analyse the contribution of DNA polymorphisms within the E6/E7 genes of HPV-16 to the heterogeneous splicing profile, the E6/E7 sequences derived from three closely related variants (isolated from the reference HPV-16 or CaSki and SiHa cells) were used to transfect C33-A cells. The genes from CaSki (E6/E7C) differ from HPV-16R (E6/E7R) in only 2 nts (positions 131 and 350); while, the genes from SiHa (E6/E7S) differ from E6/E7R in 3 nts (positions 350, 442 and 645). Similarly, E6/E7C and E6/E7S differ in 3 nts (131, 442 and 645) between them, as detailed in Fig. 1.

Interestingly, after stably transfecting C33-A cells using the three E6/E7 sequences, the results revealed that the profile of spliced transcripts was different among them. Notably, the major transcript, E6*I/E7, was differentially expressed in all cells, with the highest level corresponding to cells transfected with the E6/E7S cassette. Conversely, the unspliced E6/E7$^{pre\text{-}mRNA}$ was detected in the higher amount in the E6/E7R cassette. Differences in minor transcripts were not obvious, it likely due to their low levels of expression. As a result, we carried out a nested PCR to better detect these mRNAs, despite the fact that the accuracy of the estimates of expression levels was sacrificed. The results revealed that the E6*II/E7 transcript was only detected for the E6/E7C and E6/E7S cassettes (Fig. 1). By using three polymorphic HPV-16 E6/E7 variants, we were able to reproduce the heterogeneity in the transcript pattern in a similar fashion to that observed in tumour samples; however, the level of the E6/E7 transcripts seems to be slightly higher for the endogenous HPV-16 in the CaSki and SiHa cells. These data suggest that one or more of the SNPs may be involved in the production of the heterogeneous profile.

3.2. Two single nucleotide changes are sufficient to produce transcript heterogeneity

To assess whether the variation in the splicing profile could be attributable to a single nucleotide change in the E6/E7 sequences, we constructed G350 and C442 single nucleotide mutants containing only one of the polymorphic nucleotides with respect to the E6/E7R sequence. Additionally, we also constructed a third mutant (G350/C442) containing both of the changes. All of mutants were verified by sequencing (Fig. 2). C33-A cells were stably transfected with the mutants, and the results also revealed a differential E6/E7 splicing profile. These results shown that the expression of E6*II was increased in the two mutants containing the C442

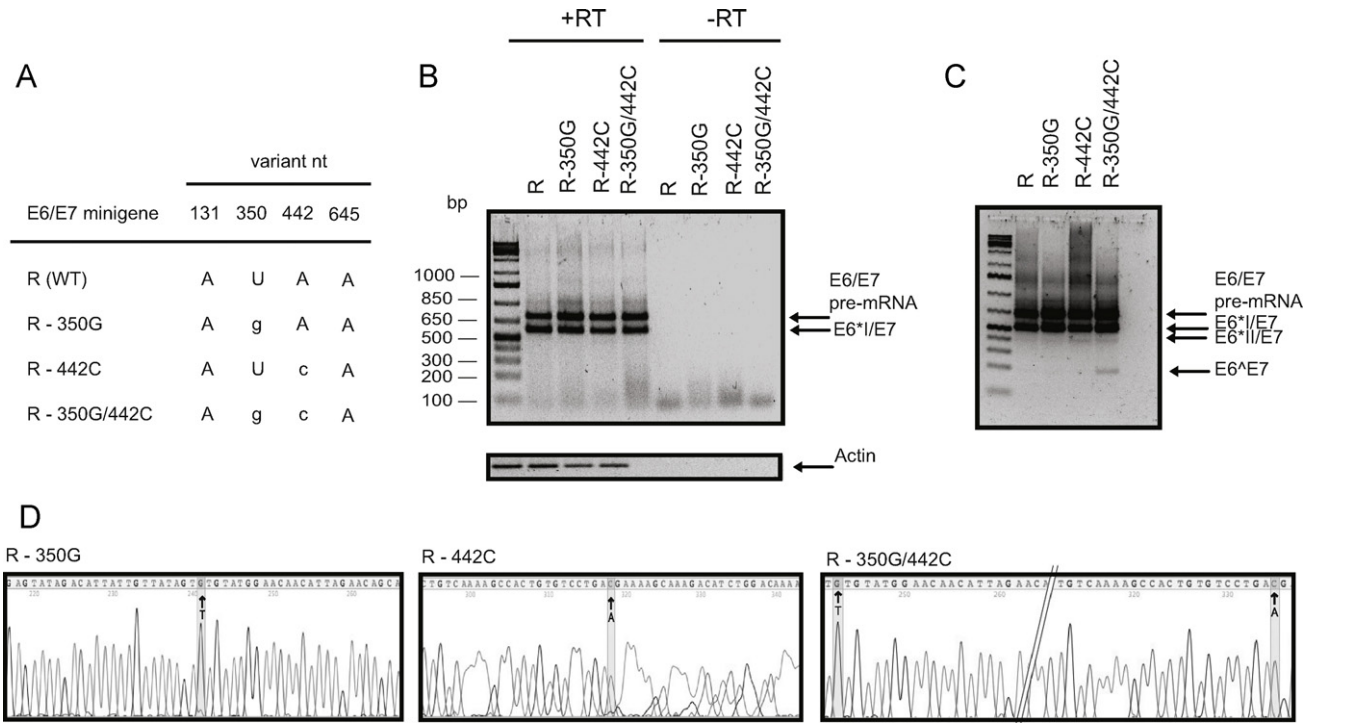

Fig. 2. Single nucleotide differences produce a differential splicing pattern in the processing of the E6/E7$^{pre\text{-}mRNAs}$. (A) Nucleotide differences between wild-type HPV-16R E6/E7 and the mutant minigene constructs are shown. (B) Profile of spliced E6/E7 transcripts obtained by RT-PCR or (C) nested PCR from C33-A cells transfected with the E6/E7 mutants and the reference sequence. (D) Electropherograms showing the single nucleotide changes introduced in each of the mutants. Single nucleotide changes are showed in lowercase.

SNP, whereas the mutant containing both changes (G350/C442) increased also the expression of the E6Ê7 transcript. The resulting transcript profiles were reminiscent to those observed between the three natural variants as clearly shown by a nested PCR (Fig. 2) and also suggested the importance of these two SNPs in the heterogeneity. Taken together, all of the data clearly show that the splicing profiles are heterogeneous even in the E6/E7 mutants containing only a single nucleotide polymorphism, which confirms the participation of single SNPs in generating the heterogeneity of the E6/E7 transcript profiles in cervical carcinoma cells.

3.3. Single SNP produces structural changes in E6/E7 transcript folding

The SNPs are located upstream and far from the splicing donor (SD) or acceptor sites (SA) [SNP 131 (95 nts upstream of the SD); SNP 350 (59 nts upstream of the SA-A); SNP 442 (74 nts upstream of the SA-B) and SNP 650 (92 nts upstream of the SA-C)]. This fact raised the possibility that the polymorphic nucleotides caused structural changes in the variant pre-mRNAs that altered the recognition of the splice sites. To identify structural changes in the three E6/E7 variant transcripts, we analysed their secondary structures. *In silico* analyses of the E6/E7 sequences revealed minor, but obvious, changes in the predicted secondary structures of the polymorphic variants and mutants. By analysis of the sequences surrounding each of the SNPs, we observed that the U350G and A442C SNPs produce the observed structural changes. Not structural changes were observed for A131G or A645C SNPs (data not shown). However, due to the complexity of such structures, we only focused our efforts on the first acceptor of intron 1, SA-A (short intron 1). The predicted structure for the E6/E7R transcript showed that U350 was located within a bulged region that was in the middle of two short, double-chain stem structures. Conversely, the G350 variant in the E6/E7 transcript was included in an internal loop structure, which was also within two short stems (Fig. 3). A similar structure was observed for the CaSki sequence, which also contains the 350G SNP (data not shown). These plausible structural changes merited more careful analyses because the RNA secondary structures suggested by *in silico* analyses are not necessarily stable structures in solution and under physiological conditions. As a result, we performed RNA structure mapping using RNases A, T1 and V1. With this purpose in mind, 207 nt RNAs corresponding to the E6/E7 transcripts (nts 234–440) from HPV-16R (probe R) and SiHa (probe S) were used for mapping assays. The respective transcripts from SiHa and CaSki short intron 1 (from donor to acceptor A), were exactly equal in nucleotide sequences, which enabled us to make a direct comparison using only two of the three sequences (probes R and S). The RNA probes contained the branch point sequence (BP), the polymorphic nucleotide at position 350, the polypyrimidine tract (PPT) and the splicing acceptor-A (SA-A). The RNAse digestion profiles revealed two additional G bands in probe S digested with a low amount of RNAse T1 when compared to probe R that does not contain G in position 350. For the SiHa RNA, two very sensitive guanines were detected by digestion with RNAse T1. The first site corresponds to G350, and the second is consistent with higher sensitivity of the next upstream G (G348), which is located at the loop. No digestion was observed at position 350 when using the reference probe, which contains a uridine (Fig. 3). However, the RNAse V1 digestion profile of both RNA probes suggested higher double-strandedness around G350, which was unexpected. A reasonable explanation is that an alternative structure may be produced, in such structure, the GU pairs in loop c′ might be paired producing an increased double-strandedness. In both cases, our results suggest structural change in the nucleotides surrounding the polymorphic site at position 350. The structural changes observed in the E6/E7 variants may contribute to alter the recognition of the sequences surrounding nt 350 by splicing factors, therefore, as initial approach to identify protein factors that could be involved in the differential

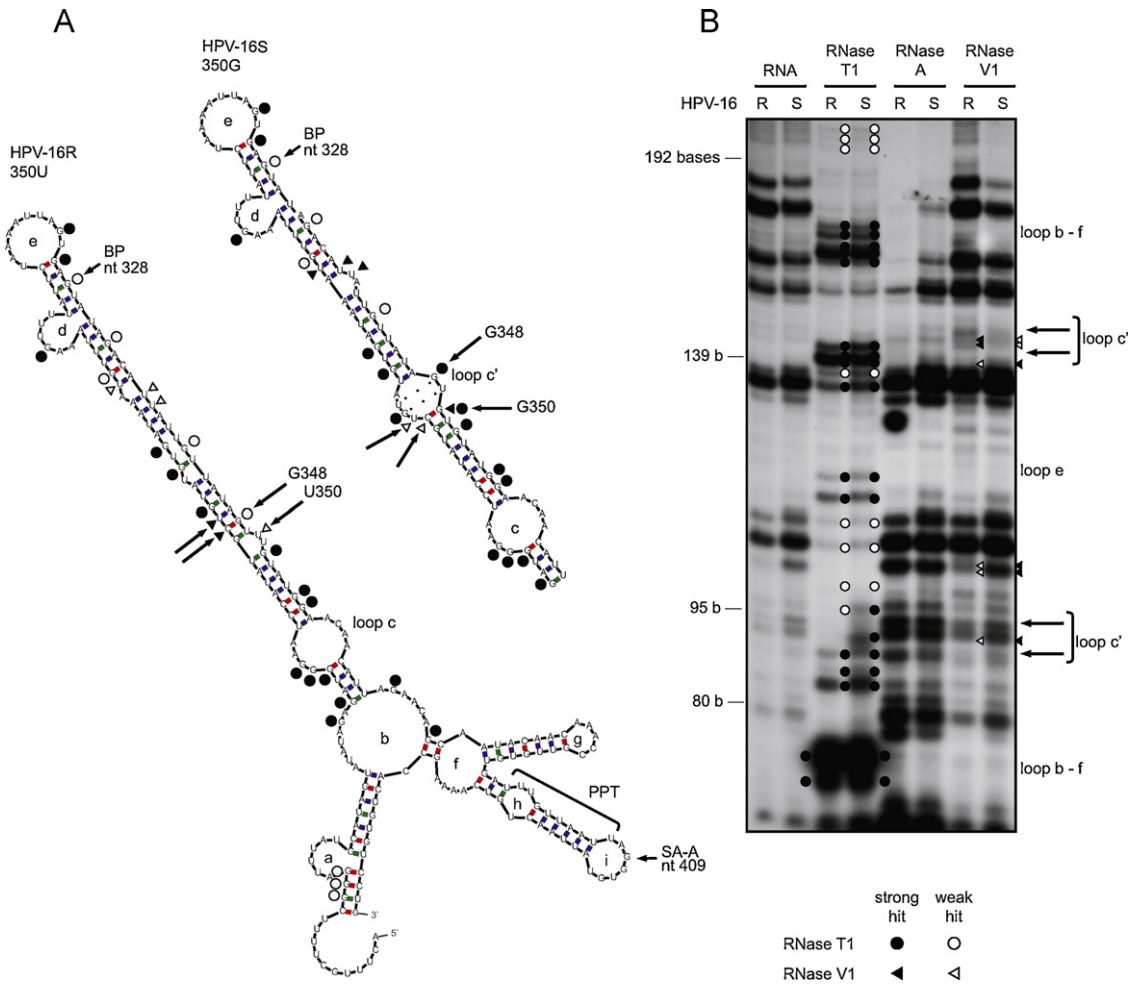

Fig. 3. Nucleotide polymorphisms produce changes in the secondary structure of the first splicing acceptor site. (A) Secondary structure prediction of the short intron 1 sequence containing only the splicing acceptor A in the HPV-16 R transcript. Only the different region is shown for the short intron 1 of HPV-16S. Predicted RNA loops were marked by lowercase letters; the presence of an additional loop (c′) in the SiHa RNA is shown. (B) To map the RNA secondary structure of the short intron 1, *in vitro* labelled RNA probes of HPV-16R or S were digested with low amounts of RNase T1 (cleaves 3′ of single stranded Gs), RNase A (cleaves 3′ of C and U) and RNase V1 (cleaves 3′ of base-paired nucleotides). Dark and light symbols indicate strong or weak cleavage and the circles and triangles indicate RNase T1 or V1 cleavage sites, respectively. Arrows indicate changes in RNase cleavage intensities that correlate with differences in the predicted secondary structure. The possibility of base pairing between GU residues within the loop c′ is signalled by a dotted line.

recognition of these SNPs; we set out to investigate the effect of the SNPs on the recruitment of nuclear factors to the transcripts containing each of the sequence changes.

3.4. E6/E7 genetic variation produces a differential binding of nuclear factors in the intron 1

To analyse the effect of the genetic variation of the E6/E7 variants on the protein factor recruitment, we performed RNA-protein UV cross-linking experiments using a 772 nts E6/E7$^{\text{pre-mRNA}}$ probe from each variant and HeLa NEs. After comparing the binding profiles among transcripts of the E6/E7 variants, the results revealed a few differences in the binding of at least seven factors with apparent molecular weights of 32, 45, 50, 54, 58, 62 and 70 kDa (Fig. 4). Conversely, when the cross-linking patterns from RNA fragments containing the three alternative acceptors from the HPV-16R intron were compared, the results revealed that many nuclear factors were recruited differentially to three splicing acceptors (Fig. 4A). The data showed that a difference in a few nucleotides among pre-mRNAs was sufficient to cause differential recruitment of nuclear factors. To determine the extent to which each of the isolated SNPs contributes to differences in binding of the nuclear factors, we performed UV cross-linking with RNA probes from each of the E6/E7 variants that contain the first SNP (SNPs A131G) and the splicing donor (SD). The results revealed that the three binding patterns were similar. The patterns suggest that this SNP, which is near the donor site, does not contribute substantially to the differential binding of nuclear factors to E6/E7 variant transcripts (Fig. 4C). Therefore, we designed a set of additional RNA probes encompassing each of the SNPs and the splicing acceptor sites. First, the three polymorphic probes containing the first acceptor, SA-A (SNPs U350G), were used for cross-linking assays. The result showed striking differences. The probe R, which contains a U at position 350, showed a differentially bound protein (p50) that was not detected in the probes C and S (Fig. 4D). Conversely, a 54 kDa band was observed bound to the three RNA probes. On the other hand, a 62 kDa band was found bound to the polymorphic sequence containing the SA-B transcript (SNP A442C) in the CaSki and reference probes (containing A 442), but the same band was barely visible in the SiHa probe that contains a C at this position (Fig. 4E). The analysis of the three 192 nt polymorphic probes with the third acceptor, SA-C (SNP A650C), showed other differential weakly binding proteins, however these protein bands were not sharp enough to be analysed.

To identify the differentially recruited protein factors, we carried out protein-RNA affinity capture experiments to isolate nuclear

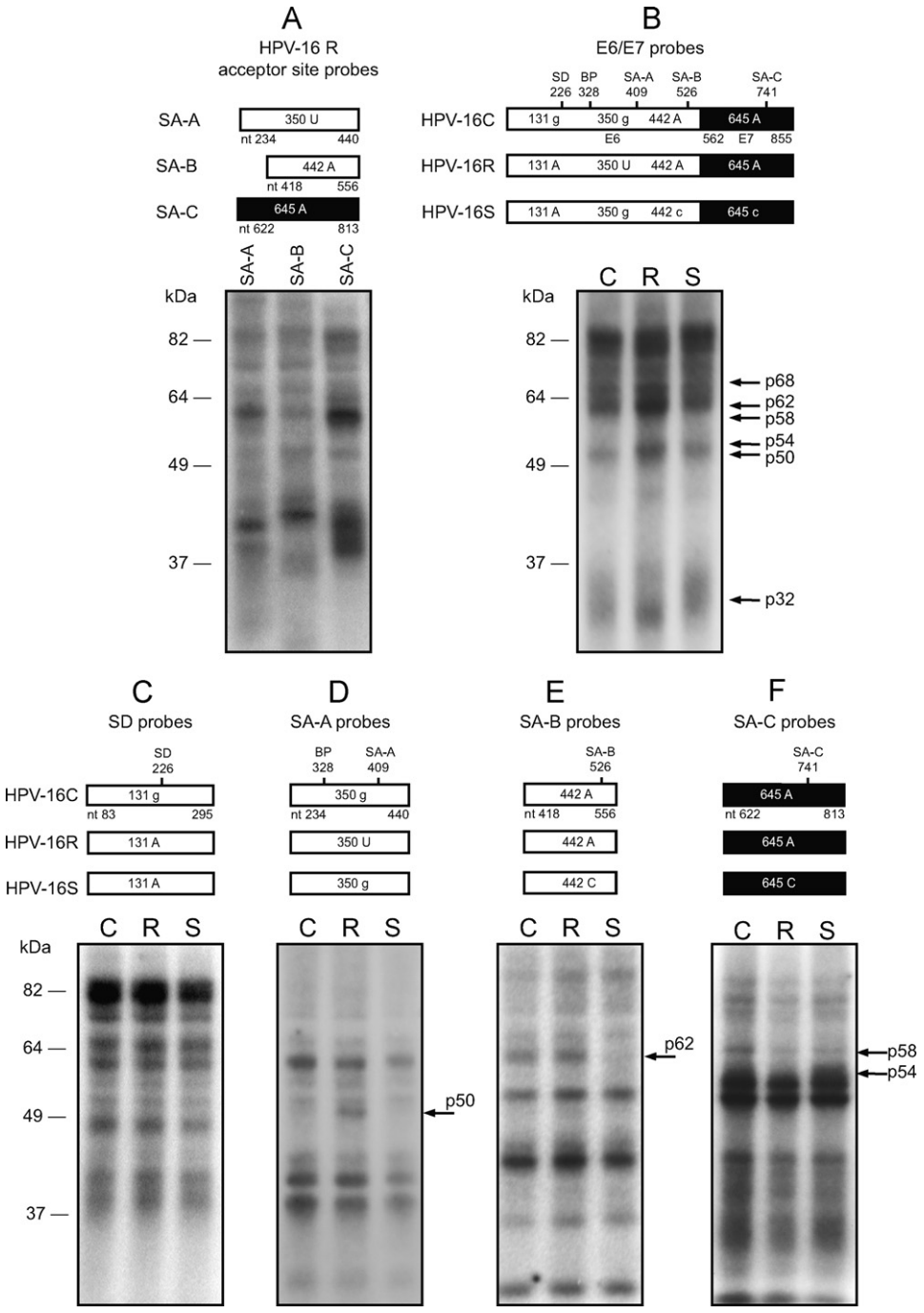

Fig. 4. The E6/E7$^{pre\text{-}mRNAs}$ from polymorphic HPV-16 variants differentially recruited splicing factors. (A) RNA-protein, UV cross-linking experiments comparing the splicing factors bound to RNA probes containing each of the acceptor sites from HPV-16R. A diagram of the RNA probes used for each assay is shown in the top of each panel. Numbers indicate the beginning and end of each RNA probe, and the nucleotide variation is shown inside of each cartoon. (B) Comparison of accessory splicing factors that bind to the E6/E7$^{pre\text{-}mRNAs}$ (nts 83–855) from the HPV-16C, HPV-16R and HPV-16S variants. At least six differentially bound proteins were detected in cross-linking experiments. UV cross-linking experiments using RNA probes for: (C) the splicing donor; (D) A acceptor; (E) B acceptor and (F) C acceptor. Arrows indicate the proteins that are differentially bound to each RNA probe.

proteins that bind to the HPV-16 E6/E7 regions of interest. We first used the two RNA probes containing the short intron 1 structure along with the respective SNPs (U350G). Both probes constituted the minimal sequence from the SD up to the first splicing acceptor site (SA-A). Only probes for the reference and SiHa variants were used for capture experiments because this portion of the CaSki and SiHa sequences were identical. The results revealed a unique pattern of proteins eluted from both RNA-affinity columns (Fig. 5) and confirmed that the proteins of 50 and 54 kDa were bound to the RNA probes containing the U and G polymorphisms, respectively. The purified factors were used to further confirm the protein profiles that were previously obtained in the RNA–protein, UV cross-linking assays. Importantly, the results resembled the data that were observed in previous experiments using total HeLa NE (Fig. 5). These experiments revealed that protein purification by RNA affinity chromatography was a useful method to identify the nuclear factors that are bound differentially to UV cross-linking probes. Although the bands at 50 and 54 kDa were excised from SDS-PAGE and slices were used for identification using mass spectroscopy, these assays failed to identify the proteins differentially

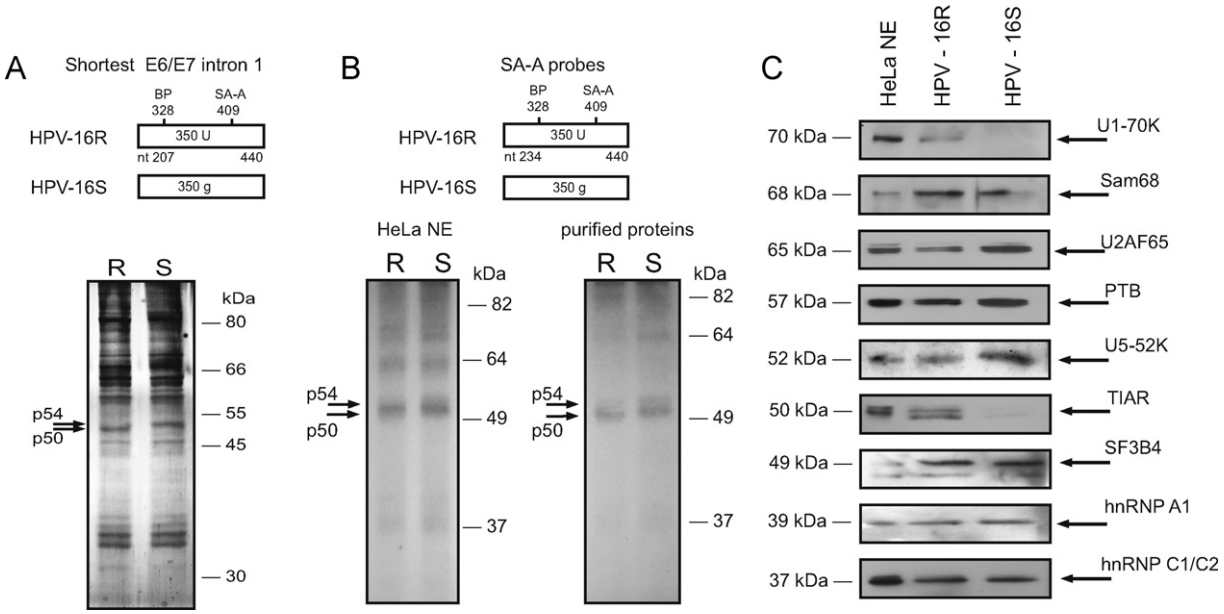

Fig. 5. Splicing factors bound differentially to the shortest form of intron 1 from the polymorphic variants, HPV-16R and HPV-16S. (A) Top: diagrammatic representation of the short intron 1 probes. Bottom: separation pattern of the RNA affinity-purified proteins bound to the intron 1 probe from the HPV-16R and HPV-16S variants. (B) Top: diagrammatic representation of the SA-A probes. Bottom: UV cross-linking experiments using HeLa NE (left) or affinity-purified proteins. (C) Relative levels of splicing factors bound to the short intron 1 probe from the HPV-16R or HPV-16S variants, as detected by Western blotting. Input HeLa NE was used as a control. Membranes were stripped and reprobed with antibodies from the different splicing factors.

bound to the transcripts. This result may have been due to the low content of the single proteins or the proteins were bound to both RNAs, but in different amounts. Additionally, all the proteins bound to each of the transcripts were digested and identified in batch, using LC–MS/MS analyses. The result revealed that more than 90 spliceosomal proteins were bound to intron 1, Around, 80 of the identified proteins were canonical splicing factors of the spliceosome assemble complexes A and B, and other 25–30 proteins that include RNA helicases, hnRNPs and RBD regulatory proteins, most which have functions in splicing that are not fully understood. Additionally, other non spliceosomal proteins were identified most of them are part of the transcription machinery or are involved in RNA processing or translation (data not shown).

The detailed mechanisms by which these proteins exert their effects on the splicing machinery remain to be determined. Intriguingly, by using this approach, only a few serine-arginine rich proteins (SR proteins) were detected bound to the short form of intron 1 from the E6/E7 transcript: SRP-53 and -55, as well as canonical spliceosomal components as SC35 and U2AF65. Interestingly, many of the proteins that were identified were known accessory or regulatory splicing factors that recognise poly-uridine stretches in pre-mRNAs, including TIA1, TIAR (TIAL1), U2AF65, PTB and hnRNP C1/C2, elav-like. Other poly-uridine binding proteins were co-purified, but we decided to investigate only the factors that were previously suggested to be involved in AS regulation of other human polymorphic transcripts. To verify that some of these proteins were differentially bound to the two E6/E7 polymorphic transcripts, similar amounts of proteins obtained by RNA affinity capture experiments were assayed by Western blotting. The results revealed that TIA1 and TIA-related protein (TIAL1/TIAR) were preferentially bound to the RNA containing a U at position 350 but were scarcely detected in proteins bound to the other RNA probe. Interestingly, U1-70K was also preferentially bound to the reference variant. Conversely, several proteins, including hnRNP A1 and hnRNPC1/C2, were bound in equal amounts to both probes. A similar result was observed for SF3b subunit 4, a component of the U2 snRNP, and Sam68. Interestingly, U2AF65, U5-52k and PTB1 were detected in slightly higher amounts bound to the SiHa probe (G-containing) than proteins bound to the reference probe. Our result revealed that U1-70k and TIA1/TIAR, a protein with an apparent molecular weight of 50 kDa, was preferentially bound to the sequence containing U350. The 54 kDa protein bound to the G350 transcript remains to be identified, although, U2AF65, U5-52k and PTB1 are good candidates.

Taken together, all of the data have shown that one single nucleotide polymorphism was sufficient to differentially recruit nuclear factors to intron 1 of HPV-16 and to produce structural changes in the intron 1 that may favour the binding of different splicing factors. The sequences surrounding nucleotides 350 and 442 were involved in differentially recruiting at least the U70k and TIA1/TIAR, or U2AF65 and PTB splicing factors in the intron 1 from the reference and SiHa variants, respectively.

4. Discussion

In this work, we analysed whether sequence polymorphisms within the HPV-16 variants contributed to the heterogeneity of the E6/E7 splicing profile. By using three E6/E7 sequences isolated from closely related HPV-16 variants, we have shown that a few nucleotide changes are sufficient to produce a differential splicing profile. Furthermore, using mutants that contained a single nucleotide change, we also showed that one SNP was sufficient to reproduce the heterogeneous splicing profile. Additionally, a difference of two or three SNPs among these viral sequences was sufficient to differentially recruit several nuclear factors to the polymorphic E6/E7 transcripts. Moreover, only one SNP was sufficient to alter the binding site of at least one nuclear or splicing factor, changing the ability of the transcript to bind some factors. Finally, some of the factors bound to the reference probe, or the variant U350, were identified as TIA1 and/or TIAR and U170k. Conversely, in the SiHa transcript the splicing factors, U2AF65 and PTB were preferentially bound.

The complete mechanism of the AS of the HPV-16 E6/E7 genes is not fully understood; however, previous work has shown that

E6*I/E7 is the major splicing product, whereas the other three minor transcripts are produced in low and variable levels in cervical cancer cells (Shirasawa et al., 1991; De la Rosa-Ríos et al., 2006) and also indicates that the three acceptor sites within intron 1 are suboptimal, and each one contains a different polypyrimidine tract (Zhao et al., 2005). Our unpublished data suggest that these acceptor sites are differentially recognised by nuclear factors. We have previously proposed that the heterogeneity in the spliced transcripts is produced, at least in part, by a differential usage of the acceptor sites (Del Moral-Hernández et al., 2010). Therefore, differential recruiting of nuclear factors onto the E6/E7 pre-mRNA may contribute to the variable recognition of the splicing acceptors. Previously, it has been suggested that some splicing factors vary in tumour cells and cell lines (Grosso et al., 2008; McFarlane and Graham, 2010; Del Moral-Hernández et al., 2010) which may mean that differential availability of splicing factors in tumour cells contributes to producing the heterogeneous splicing profile. Here, we have shown that also the HPV-16 polymorphism is crucial for producing a variation in the E6/E7 splicing profile, as it have been shown for several human genes (reviewed in Pagani and Baralle, 2004). Therefore, it is reasonable to suggest that the heterogeneity in splicing of the HPV-16 E6/E7 transcript is produced by a complex interaction between the HPV-16 polymorphic variants and the availability of canonical and accessory splicing factors in cervical carcinoma cells, which combine to result in differential use of the sub-optimal splicing acceptors within intron 1. However, the identities of accessory and regulatory splicing factors involved in the recognition of HPV-16 intron 1 are completely ignored.

Accessory splicing factors reinforce or impair the recognition of optimal and suboptimal alternative splice sites (Singh and Valcárcel, 2005). HPV introns are processed by the host cell spliceosome, and canonical splicing factors are involved in the recognition of the donor and acceptor sites within the exons; however, intron 1 is alternatively spliced by the selection of three different 3′ splice sites. Therefore, each of the three intron/exon junctions might be recognised by a different combination of accessory and canonical splicing factors. During recognition, auxiliary splicing factors form a bridge between canonical splicing factors over the intron or the exons, with the latter depending on the length of the intron (Schellenberg et al., 2008). Unfortunately, most of accessory splicing factors that bind to the intron 1 have not been identified yet.

As an initial approach to identify the factors bound to the intron 1, two RNA probes (from HPV-16R and SiHa) containing only the shortest form of intron 1 were used for purification of nuclear factors by RNA-affinity chromatography. The RNA probe encompassed the donor site (begins 20 nts upstream of the 5′ splice site) and splice site acceptor A (ends 32 nts downstream of the PyAG). Around 100 nuclear proteins were identified by MS/MS; approximately 80 core spliceosomal proteins were bound to each of the E6/E7 intron 1 constructs. Around two dozens of them were splicing regulatory or accessory factors, whose significance and function in the splicing of intron 1 remains to be understood (data not shown). Several factors differential bound to the R- and S-RNA probes were identified by using cross-linking assays, however, the factors involved in the differential recognition of the E6/E7 variants were not easily identified by LC/MS/MS; this last might be due to a differences in the apparent molecular mass of the nuclear proteins deduced by their migration in the UV cross-linking assays and the real molecular weight of these factors. Thus, identifying the binding sites and the role of the factors involved in the AS of intron 1 requires very difficult and hard future work to be done; however, several of the splicing factors that were identified deserve special attention, including: Sam68, TIAR (TIAL), UASP56, PUF60, PTB1, U1-70K, U2AF65, CD2BP2 (U5-52K), Nono, and hnRNP C1/C2, most of which are polypyrimidine binding proteins. Poly U binding proteins have been recently involved in regulation of AS for introns containing suboptimal splice sites (Le Guiner et al., 2001), as is the case for intron 1 of HPV-16.

Enhancer and silencer sequences have been described in many exons and introns that produce intron/exon bridges and positively or negatively affect AS. The short names of exonic splicing enhancers and exonic splicing silencers are ESE or ESS, respectively, and the short names of intronic splicing enhancers or intronic splicing silencers are ISE or ISS, respectively (reviewed in Pagani and Baralle, 2004). No enhancer or silencer sequences have been identified in intron 1 thus far; however, two accessory splicing factors enhancing the usage of the first acceptor site in HPV-16 in the presence of epidermal growth factor (EGF), were recently described. Although the exact position of a potential enhancer sequence was not mapped, it was clearly shown that correct recognition of the first acceptor site A (SA-A) depends on the level of the factors, Brm and Sam68. At the same time, the use of this acceptor was negatively regulated by hnRNP A1 and hnRNP A2. Yet, the binding sequences of all these factors have not been mapped or located (Rosenberger et al., 2010).

SA-A is the most frequently used of the three splicing acceptors within intron 1. As mentioned above, Sam68 is an important factor involved in enhancing the splicing of intron 1 at this acceptor site. Sam68 is a member of the STAR splicing factors that binds single strand homopolymer stretches of U or A nucleotides (Lukong and Richard, 2003). Moreover, Sam68 interacts with TIAR1 within exon v5 of CD44 to promote the recruitment of U1 snRNP and U6 snRNP (Zhu et al., 2003) by way of the interaction of TIAR1 with U1C (Förch et al., 2002) and to reinforce the recognition of weak donor sites containing downstream poly-U stretches. The lack of U-stretches or TIA1/TIAR binding, results in exon skipping in several eukaryotic (Förch et al., 2000; Le Guiner et al., 2001; Izquierdo et al., 2005) and viral transcripts (Choi and Pintel, 2009). The same exon, v5, binds Brm, a member of the SWI/SNF complex, which is also a known regulator of AS. Brm interacts with Prp 6 and SAP 130/SF3a120a, which are components of the U5 snRNP (Allemand et al., 2008) and U2 snRNP (Underhill et al., 2000), respectively. These interactions suggest that Brm is involved in the recruitment of the U2/U5.U6 snRNP complex, which promotes assembly of the spliceosome. Interestingly, all of these auxiliary splicing factors (except for Brm) were identified in the RNA-affinity capture experiments using the short forms of intron 1-R and -S, containing only the splicing donor and SA-A (data not shown). It has also been shown that Sam68 binds upstream from the BPs in CD44 exon v5 and increases recognition of the downstream splicing acceptor site, by increasing binding of U2AF (Batcshé et al., 2006) and also its effect is exerted in a signal transduction-mediated manner (Tisserant and König, 2008).

The observed changes in the splicing of the E6/E7 variant transcripts containing SNPs suggests the presence of positive and negative regulatory sequences upstream and downstream of acceptor site A. In summary, the change from A to C at nt 442 (A442C) seems to negatively alter the recognition of acceptor site A, which simultaneously enhances the recognition of acceptors B and increases the level of the E6*II/E7. Therefore, it is reasonable to suggest that an enhancer is affected by this SNP and that nucleotide U350 is part of this regulatory sequence. Interestingly, the only experimentally detected BP (De la Rosa-Ríos et al., 2006) lies at nt 328. Therefore, it is likely that this positive regulatory sequence may increase the recognition of the acceptor site and the BPs by U2AF and U2 snRNP and its associated splicing factors. Intriguingly, the sequence changes at nts 350/442 (U to G/A to C) enhanced the usage of acceptor site B and C and resulted in increased levels of the E6*II/E7 and E6Ê7 transcripts, which may also reduce the efficiency of acceptor site A use, likely by simple competition between the acceptors. These data suggest that polymorphism

442 affects putative silencer sequences that are involved in the negative control of the distal acceptor.

In silico analyses of the E6/E7 sequence from nts 234 to 440 reveals a large stem-loop structure with the U350 in the middle of three bulged nucleotides. The change from U to G produces a putative alternative folding of this RNA sequence, which places the G350 in the middle of an internal loop. RNAse T1 mapping of this HPV-16 E6/E7 sequence suggested higher amounts of single-stranded RNA in G350 constructs than in U350 transcripts. It is reasonable to speculate that the U350G change alters the recognition of auxiliary splicing factors involved in the functioning of the HPV-16 E6/E7 enhancer, because splicing regulatory sequences (enhancers or silencers) require the presence of single stranded RNA regions (Hiller et al., 2007). Similarly, *in silico* analyses showed an obvious structural change in the secondary structure of nts 418–556 within the E6/E7 RNA, where single-stranded regions were increased around nt 442 in the SiHa sequence containing a C at position 442 (data not shown). However, the RNAse V1 digestion profile of both RNA probes suggested slight increases in double-strandedness around G350, which was unexpected. A reasonable explanation is that alternative RNA structures may coexist in this RNA in solution. These findings suggest a complex interaction between regulatory sequences that dictate use of the three splicing acceptors in the HPV-16 E6/E7 RNA, this probably due to all three acceptor sharing the only BP sequence located upstream of acceptor A. As we have shown, a 50 kDa splicing factor interacts with the RNA motif containing nt U350 (guUugu), and this factor is substituted by a protein of 54 kDa when this region contains a G at the same position. This latter unidentified protein is bound at least to the guGugu RNA stretch. Interestingly, the combination of both changes, 350G and 442C, increases the usage of both B and C acceptors, which suggests a very complex interaction between the putative regulatory motifs. These interactions between splice sites, auxiliary splicing factors and regulatory sequences merit a hard and careful future work to determine the exact mechanism by which these elements exert their effects and why distal nucleotide changes that alter the recognition of accessory factor neighbouring the A acceptor, may positively or negatively affect the use of B and C acceptor sites.

The E6/E7$^{pre\text{-}mRNA}$ level is displayed differently in CaSki and SiHa cells; CaSki cells expresses around a three-fold with respect to SiHa cells (Jeon and Lambert, 1995). This fact correlates with the low level of the E6 protein (Shai et al., 2007) observed in SiHa cells; similar differences are observed for the E7 protein (Wang et al., 2007). The low levels of the two oncoproteins observed in SiHa, also correlate with the low level of the E6/E7$^{pre\text{-}mRNA}$, that was consistently observed in its E6/E7splicing profile. Since E6/E7$^{pre\text{-}mRNA}$ RNA is the main source of the E6 and E7 oncoproteins (Stacey et al., 1995; Del Moral-Hernández et al., 2010); an alternative explanation is that, heterogeneity of the splicing profile, specifically on the level of the E6/E7$^{pre\text{-}mRNA}$ is at least in part responsible for the different expression of the E6 and E7 proteins observed between these cells. A differential expression of the oncoproteins dependent on the HPV-16 variant could have important consequences in tumour cell development.

In summary, we showed that single nucleotide changes in the RNA sequence of a viral RNA were sufficient to differentially recruit accessory spliceosomal proteins to the intron 1 sequences. In addition, several splicing factors were identified that bound to the polymorphic introns: TIA1/TIAR and U1-70k for the E6/E7R or U2AF65, U5-52k and PTB for E6/E7S. In the future, the analyses of the interactions of some of these factors with *cis*-acting regulatory elements in the HPV-16 intron 1, and their role in AS, will shed light on the mechanism that controls the alternative usage acceptor sites in this intron, and the consequences of the heterogeneity over the expression level of the HPV-16 oncoproteins.

Acknowledgements

We are grateful to Beatríz Tinoco for expert technical assistance. We thank, Dr Mary Ann Gawinowicz, from the protein core facility in Columbia University College of Physicians & Surgeons, for the mass spectrometry identification of proteins. E.L-U. is indebted to CONACyT, Mexico for a doctoral fellowship. This work was supported in part by funds allocated to N.V-S. by CINVESTAV and by grants 38217N and 81611 from CONACyT.

References

Allemand, E., Batsché, E., Muchardt, C., 2008. Splicing, transcription, and chromatin: a ménage à trois. Curr. Opin. Genet. Dev. 18, 145–151.
Asadurian, Y., Kurilin, H., Lichtig, H., Jackman, A., Gonen, P., Tommasino, M., Zehbe, I., Sherman, L., 2007. Activities of human papillomavirus 16 E6 natural variants in human keratinocytes. J. Med. Virol. 79, 1751–1760.
Baker, C.C., Phelps, W.C., Lindgren, V., Braun, M.J., Gonda, M.A., Howley, P.M., 1987. Structural and transcriptional analysis of human papillomavirus type 16 sequences in cervical carcinoma cell lines. J. Virol. 61, 962–971.
Batcshé, E., Yaniv, M., Muchrardt, C., 2006. The human SWI/SNF subunit Brm is a regulator of alternative splicing. Nat. Struct. Mol. Biol. 13, 22–29.
Chan, S-Y., Ho, L., Ong, C-K., Chow, V., Drescher, B., Dürst, M., ter Meulen, J., Villa, L., Luande, J., Mgaya, H.N., Bernard, H-U., 1992. Molecular variants of human papillomavirus type 16 from our continent suggest ancient pandemic spread of the virus and its coevolution with humankind. J. Virol. 66, 2057–2066.
Choi, E-Y., Pintel, D., 2009. Splicing of the large intron present in the nonstructural gene of minute virus of mice is governed by TIA-1/TIAR binding downstream of the nonconsensus donor. J. Virol. 83, 6306–6311.
De la Rosa-Ríos, M.A., Martínez-Salazar, M., Martínez-Garcia, M., Gonzalez-Bonilla, C., Villegas-Sepúlveda, N., 2006. The intron I of HPV16 has a suboptimal branch point at a guanosine. Virus Res. 118, 46–54.
de Villiers, E.M., Fauquet, C., Broker, T.R., Bernard, H.-U., zur Hausen, H., 2004. Classification of papillomaviruses. Virology 324, 17–27.
Del Moral-Hernández, O., López-Urrutia, E., Bonilla-Moreno, R., Martínez-Salazar, M., Aréchaga-Ocampo, E., Berumen, J., Villegas-Sepúlveda, N., 2010. The HPV-16 E7 oncoprotein is expressed mainly from the unspliced E6/E7 transcript in cervical carcinoma C33-A cells. Arch. Virol. 155, 1959–1970.
Dyhr-Mikkelsen, H., Kjems, J., 1995. Inefficient spliceosome assembly and abnormal branch site selection in splicing of an HIV-1 transcript in vitro. J. Biol. Chem. 270, 24060–24066.
Förch, P., Puig, O., Kedersha, N., Martínez, C., Granneman, S., Séraphin, B., Anderson, P., Valcárcel, J., 2000. The apoptosis-promoting factor TIA-1 is a regulator of alternative pre-mRNA splicing. Mol. Cell 6, 1089–1098.
Förch, P., Puig, O., Martinez, C., Seraphin, B., Valcarcel, J., 2002. The splicing regulator TIA-1 interacts with U1-C to promote U1 snRNP recruitment to 5′ splice sites. EMBO J. 21, 6882–6892.
Grosso, A.R., Martins, S., Carmo-Fonseca, M., 2008. The emerging role of splicing factors in cancer. EMBO Rep. 9, 1087–1093.
Harris, D., Zhang, Z., Chaubey, B., Pandey, V.N., 2006. Identification of the cellular factors associated with the 3'-nontranslated region of the hepatitis c virus genome. Mol. Cell Proteomics 5, 1006–1018.
Hawley-Nelson, P., Voudsen, K.H., Hubbert, N.L., Lowy, D.R., Schiller, J.T., 1989. HPV16 E6 and E7 proteins cooperate to immortalizate human foreskin keratinocytes. EMBO J. 8, 3905–3910.
Hiller, M., Zhang, Z., Backofen, R., Stamm, S., 2007. Pre-mRNA secondary structures influence exon recognition. PLOS Genet. 3, 2147–2155.
Ho, L., Chan, S-Y., Chow, V., Chong, T., Tay, S-K., Villa, L.L., Bernard, H-U., 1991. Sequence variants of human papillomavirus type 16 in clinical samples permit verification and extension of epidemiological studies and construction of a phylogenetic tree. J. Clin. Microbiol. 29, 1765–1772.
Izquierdo, J.M., Majós, N., Bonna, S., Martínez, C., Castelo, R., Guigó, R., Bilbao, D., Valcárcel, J., 2005. Regulation of Fas alternative splicing by antagonistic effects of TIA-1 and PTB on exon definition. Mol. Cell 19, 475–484.
Jeon, S., Lambert, P.F., 1995. Integration of human papillomavirus type 16 DNA into the human genome leads to increased stability of the E6 and E7 mRNAs: Implications for cervical carcinogenesis. Proc. Natl. Acad. Sci. U.S.A. 92., 1654–1658.
Le Guiner, C., Lejeune, F., Galiana, D., Kister, L., Breathnach, R., Stévenin, J., Del Gatto-Konczak, F., 2001. TIA-1 and TIAR activate splicing of alternative exons with weak 5′ splice sites followed by a U-rich stretch on their own pre-mRNAs. J. Biol. Chem. 276, 40638–40646.
Lichtig, H., Algrisi, M., Botzer, L.E., Abadi, T., Verbitzky, Y., Jackman, A., Tommasino, M., Zehbe, I., Sherman, L., 2006. HPV16 natural variants exhibit different activities in funcional assays relative to the carcinogenic potential of E6. Virology 350, 216–227.
Lukong, K.E., Richard, S., 2003. Sam68, the KH domain-containing superSTAR. Biochim. Biophys. Acta 1653, 73–86.
McFarlane, M., Graham, S.V., 2010. Human papillomavirus regulation of SR proteins. Biochem. Soc. Trans. 38, 1116–1121.
McNicol, P., Guijon, F., Wayne, S., Hidajat, R., Paraskevas, M., 1995. Expression of human papillomavirus type 16 E6-E7 open reading frame varies quantitatively

in biopsy tissue from different grades of cervical intraepithelial neoplasia. J. Clin. Microbiol. 33, 1169–1173.
Milligan, S.G., Veerapraditsin, T., Ahamet, B., Mole, S., Graham, S.V., 2007. Analysis of novel human papillomavirus type 16 late mRNAs in differentiated W12 cervical epithelial cells. Virology 360, 172–181.
Münger, K., Phelps, W.C., Bubb, V., Howley, P.M., Schlegel, R., 1989. The E6 and E7 genes of human papillomavirus type 16 together are necessary and sufficient for transformation of primary human keratinocytes. J. Virol. 63, 4417–4421.
Pagani, F., Baralle, F.E., 2004. Genomic variants in exons and introns: identifying the splicing spoilers. Nat. Rev. Genet. 5, 389–396.
Rosenberger, S., De-Castro, Arce, J., Langbein, L., Steenbergen, R.D.M., Rösl, F., 2010. Alternative splicing of human papillomvirus type-16 E6/E6*I early mRNA is coupled to EGF signaling via Erk1/2 activation. Proc. Natl. Acad. Sci. U.S.A. 107, 7006–7011.
Shai, A., Brake, T., Somoza, C., Lambert, P.F., 2007. The human papillomavirus E6 oncogene dysregulates the cell cycle and contributes to cervical carcinogenesis through two independent activities. Cancer Res. 67, 1626–1635.
Schellenberg, M.J., Ritchie, D.B., MacMillan, A.M., 2008. Pre-mRNA splicing: a complex picture in higher definition. Trends Biochem. Sci. 33, 243–246.
Sharp, P.A., 1994. Split genes and RNA splicing. Cell 77, 805–816.
Shirasawa, H., Tanzawa, H., Matsunaga, T., Simizu, B., 1991. Quantitative detection of spliced E6-E7 transcripts of human papillomavirus type 16 in cervical premalignant lesions. Virology 184, 795–798.
Singh, R., Valcárcel, J., 2005. Building specificity with nonspecific RNA-binding proteins. Nat. Struct. Mol. Biol. 12, 645–653.
Smotkin, D., Prokoph, H., Wettstein, F.O., 1989. Oncogenic and nononcogenic human genital papillomaviruses generate the E7 mRNA by different mechanisms. J. Virol. 63, 1441–1447.
Stacey, S.N., Jordan, D., Snijders, P.J.F., Mackett, M., Walboomers, J.M.M., Arrand, J.R., 1995. Translation of the human papillomavirus type 16 E7 oncoprotein from bicistronic mRNA is independent of splicing events within the E6 open reading frame. J. Virol. 69, 7023–7031.
Stöppler, M.C., Ching, K., Stöppler, H., Clancy, K., Schlegel, R., Icenogle, J., 1996. Natural variants of the human papillomavirus type 16 E6 differ in their abilities to alter keratinocyte differetiation and to induce p53 degradation. J. Virol. 70, 6987–6993.
Tisserant, A., König, H., 2008. Signal-regulated pre-mRNA occupancy by the general splicing factor U2AF. PLOS One 3, e1418.
Underhill, C., Qutob, M.S., Yee, S.P., Torchia, J., 2000. A novel nuclear receptor corepressor complex, N-CoR, contains components of the mammalian SWI/SNF complex and the corepressor Kap-1. J. Biol. Chem. 275, 40463–40470.
Walboomers, J.M.M., Jacobs, M.V., Manos, M.M., Bosch, F.X., Kummer, J.A., Shah, K.V., Snijders, P.J.F., Peto, J., Meijer, C.J.L.M., Muñoz, N., 1999. Human papillomavirus is necessary cause of invasive cercal cancer worldwide. J. Pathol. 189, 12–19.
Wang, X.G., Revskaya, E., Bryan, R.A., Strickler, H.D., Burk, R.D., Casadevall, A., Dadachoca, E., 2007. Treating cancer as an infectious disease-viral antigens as novel targeting for treatment and potential prevention of tumors of viral etiology. Plos one 10, e1114.
Xi, L.F., Koutsky, L.A., Hildesheim, A., Galloway, D.A., Wheeler, C.M., Winer, R.L., Ho, J., Kiviat, N.B., 2007. Risk for high-grade cervical intraepithelial neoplasia associated with variants of human papillomavirus types 16 and 18. Cancer Epidem. Biomar. 16, 4–10.
Yamada, T., Manos, M.M., Peto, J., Greer, C.E., Muñoz, N., Bosch, F.X., Wheeler, C.M., 1997. Human papillomavirus type 16 sequence variation in cervical cancers: A worldwide perspective. J. Virol. 71, 2463–2472.
Zabolontny, J.M., Krummenacher, C., Fraser, N.W., 1997. The herpes simplex virus type 1 2.0 kilobase latency-associated transcript is a stable intron which branches at a guanosine. J. Virol. 71, 4199–4208.
Zehbe, I., Voglino, G., Delius, H., Wilander, E., Tommasino, M., 1998a. Risk of cervical cancer and geographical variations of human papillomavirus 16 E6 polymorphisms. Lancet 352, 1441–1442.
Zehbe, I., Wilander, E., Delius, H., Tommasino, M., 1998b. Human papillomavirus 16 E6 variants are more prevalent in invasive cervical carcinoma than the prototype. Cancer Res. 58, 829–833.
Zehbe, I., Richard, C., De Carlo, C.A., Shai, A., Lambert, P.F., Lichtig, H., Tommasino, M., Sherman, L., 2009. Human papillomavirus 16 E6 variants differ in their dysregulation of human keratinocyte differentiation and apoptosis. Virology 383, 69–77.
Zhao, X., Oberg, D., Rush, M., Fay, J., Lambkin, H., Schwartz, S., 2005. A 57-nucleotide upstream early polyadenylation element in human papillomavirus type 16 interacts with hFip1, CstfF64, hnRNP C1/C2, and polypyrimidine tract binding protein. J. Virol. 79, 4270–4288.
Zhu, H., Hasman, R.A., Young, K.M., Kedersha, N.L., Lou, N.L., 2003. U1 snRNP-dependent function of TIAR in the regulation of alternative RNA processing of the human calcitonin/CGRP pre-mRNA. Mol. Cell Biol. 23, 5959–5971.
Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 31, 3406–3415.
Zuna, R.E., Moore, W.E., Shanesmith, R.P., Dunn, S.T., Wang, S.S., Schiffman, M., Blakey, G.L., Teel, T., 2009. Association of HPV16 E6 variants with diagnostic severity in cervical cytology samples of 354 women in a US population. Int. J. Cancer 125, 2609–2613.
zur Hausen, H., 2008. Papillomairuses—to vaccination and beyond. Biochemistry (Mosc.) 73, 498–503.