



HAL
open science

Impact des isoformes E6 sur le métabolisme des ARN dans les cancers associés aux HPV

Philippe Paget-Bailly

► **To cite this version:**

Philippe Paget-Bailly. Impact des isoformes E6 sur le métabolisme des ARN dans les cancers associés aux HPV. Biologie cellulaire. Université Bourgogne Franche-Comté, 2021. Français. NNT : 2021UBFCE004 . tel-03426967

HAL Id: tel-03426967

<https://theses.hal.science/tel-03426967v1>

Submitted on 12 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE DE DOCTORAT DE L'ETABLISSEMENT
UNIVERSITE DE BOURGOGNE FRANCHE-COMTE

Préparée à l'UFR Santé de Besançon

Ecole doctorale n°554 Environnement Santé

Doctorat de Sciences de la Vie et de la Santé

par

M. Philippe PAGET-BAILLY

Impact des isoformes E6 sur le métabolisme des ARN
dans les cancers associés aux HPV

Thèse présentée et soutenue à Besançon, le 6 mai 2021.

Composition du jury :

Dr. Stéfania MILLEVOI	DR, CRCT Oncopôle, Toulouse	Rapporteur
Dr. Ignacio BRAVO	DR, CNRS, Montpellier	Rapporteur
Dr. Gilles TRAVE	DR, IGBMC, Strasbourg	Examineur
Dr. Catherine LLANES	MCF, Univ. Bourgogne Franche-Comté	Examineur
Pr. Jean-Luc PRETET	PU-PH, Univ. Bourgogne Franche-Comté	Examineur
Dr. Aurélie BAGUET	MCF, Univ. Bourgogne Franche-Comté	Directrice de thèse

Remerciements

Ce manuscrit représente l'aboutissement de cinq années de travail menées au sein du laboratoire de l'EA3181 - Carcinogénèse associée aux HPV : facteurs prédictifs et pronostiques. Je tiens à remercier ici toutes les personnes qui m'ont accompagné de près ou de loin dans cette aventure.

Je tiens tout d'abord à remercier les membres de mon jury de thèse,

Madame le docteur Stéfania Millevoi, rapporteur,

Monsieur le docteur Ignacio Bravo, rapporteur,

Monsieur le docteur Gilles Trave, examinateur,

Madame le docteur Catherine Llanes, examinateur,

pour avoir accepté de consacrer du temps à la lecture et à l'analyse de ce travail.

Je tiens à remercier tout particulièrement madame le docteur Aurélie Baguet, mon directeur de thèse. Tu as toujours su m'apporter ton soutien et ton envie de travailler. Je ne te remercierais jamais assez de m'avoir fait rentrer dans le monde de la recherche. J'espère avoir été un bon élève.

J'aimerais remercier chaleureusement les professeurs Christiane Mougin et Jean-Luc Pretet, directeurs successifs de l'EA3181, pour leurs conseils toujours avisés et la rigueur de travail qu'ils m'ont transmis. Merci aussi à mes anciens professeurs et membres de l'équipe de recherche AETIC pour leurs conseils. Je remercie également tous les membres de l'EA3181 sans qui ce travail n'aurait pas été possible. Anne et Julie bien sûr, c'est grâce à vous que les petits stagiaires timides se transforment en docteurs. Merci à Sylvie et Isabelle également pour leur écoute et leurs conseils. Merci aussi aux doctorants devenus docteurs Jérôme et Koceila pour m'avoir formé à la paillasse et pour nos longues conversations existentielles du midi ! Merci aux « stagiaires » Alice, Alex, Sergen, Fanny, Zélie et Thimothée pour les rires à n'en plus finir et à qui j'espère avoir été utile. Merci enfin à toutes les structures et les financements ayant rendu ce travail possible.

Je tiens à remercier tous les chercheurs du petit monde HPV qui ont pu m'apporter leur aide au cours des années, Michael, Alain, Murielle et tant d'autres. Je remercie

particulièrement Gilles, manger un magret de canard en parlant de sciences avec toi jusqu'à minuit, ça n'a pas de prix.

Je remercie aussi mes compagnons plus ou moins professionnels Daniel et Camille, la crème de l'élite bisontine. Toujours fidèle au poste quand il s'agit de boire un coup, de parler de science, ou les deux. Faire une thèse, c'est quand même plus sympa avec des bons potes.

Je remercie bien évidemment Pierre et Lucie, mes éternels coloc' de l'ambiance, inséparables depuis la maternelle et avec qui je formerais toujours un trio d'enfer. Je remercie aussi toute la team d'affreux zozo qui me supporte depuis plus de dix ans, Cécile, Camille, Lucile, Emeline, Quentin, Elie, Pauline et bien d'autres... Merci pour votre amitié (cœur avec les doigts). Merci aussi à toute la clique BASSLIME et Sousoune en particulier pour ces soirées jamais trop longues.

Je remercie enfin ma famille Sophie, Franklin, Suzie, Camille, Rémi et bien sûr Elodie. Quelques lignes ne suffiraient pas à résumer les épreuves que nous avons traversées ces dernières années. En toute simplicité : je vous aime tous très fort. Merci infiniment à mon papa et à ma maman qui ne pourront malheureusement pas lire ce travail.

Table des matières

REMERCIEMENTS	2
LISTE DES FIGURES	6
LISTE DES TABLEAUX	7
LISTE DES ABRÉVIATIONS	8
INTRODUCTION	13
1 LES PAPILOMAVIRUS	14
1.1 ORIGINE ET CLASSIFICATION DES PAPILOMAVIRUS.....	14
1.2 CLASSIFICATION DES HPV A HAUT RISQUE ONCOGENIQUE.....	16
2 HISTOIRE NATURELLE DE L'INFECTION PAR HPV16 MENANT AU CANCER DU COL DE L'UTERUS	18
2.1 LE GENOME D'HPV16.....	18
2.2 CYCLE VIRAL PRODUCTIF D'HPV16	19
2.3 PERSISTANCE DE L'INFECTION PAR HPV16	23
3 ROLES DES ONCOPROTEINES VIRALES DANS LA CARCINOGENESE	27
3.1 LA SUREXPRESSION DE E6 ET E7	27
3.2 L'ONCOPROTEINE E6 D'HPV16.....	29
3.3 L'ONCOPROTEINE E7 D'HPV16.....	36
3.4 EPIDEMIOLOGIE DES CANCERS HPV.....	41
4 EXPRESSION DU GENOME HPV16 : CYCLE VIRAL ET CANCER	45
4.1 REGULATION TRANSCRIPTIONNELLE DES GENES D'HPV16	45
4.2 REGULATION POST-TRANSCRIPTIONNELLE DES TRANSCRITS HPV16	46
4.2.1 <i>Le mécanisme d'épissage alternatif et sa régulation</i>	46
4.2.2 <i>Le transcriptome d'HPV16</i>	49
4.2.3 <i>Régulation de l'épissage des transcrits prématurés E6E7</i>	52
4.2.4 <i>Régulation de la polyadénylation, de l'export et de la stabilité des transcrits HPV</i>	55
4.3 REGULATION TRADUCTIONNELLE DES ISOFORMES E6 ET DE E7.....	56
5 LA PROTEINE TRONQUEE E6*I D'HPV16	60
OBJECTIFS	65
RESULTATS	66
1 IMPACT DES ISOFORMES E6 SUR LE TRANSCRIPTOME CELLULAIRE	67
1.1 RESUME DE LA PUBLICATION	67
1.2 TRAVAUX COMPLEMENTAIRES	91
1.2.1 <i>Obtention des clones stables</i>	91
1.2.2 <i>Caractérisation phénotypique complémentaire des clones stables</i>	92
2 LA PROTEINE E6*I INDUIT LA PRODUCTION DE ROS	95
2.1 E6*I INDUIT LA PRODUCTION DE ROS VIA LES COMPLEXES NOX	95
2.2 DETECTION DE LA PROTEINE E6*I PAR WESTERN BLOTTING	97
2.3 L'INDUCTION DES ROS EST CAUSEE PAR LA PROTEINE E6*I D'HPV16.....	99
2.4 ETUDE DE LA TRADUCTION DE L'ARNM E6*I PAR PROFILAGE POLYSOMIQUE.....	100
3 ETUDE DES EFFETS DES ISOFORMES E6 DANS UNE COHORTE DE CANCERS DE L'OROPHARYNX	105
3.1 EXPRESSION DES TRANSCRITS E6 DANS LES TUMEURS DE L'OROPHARYNX	106

3.2	DEREGULATION DES ARNm CELLULAIRES EN FONCTION DE L'EXPRESSION DES TRANSCRITS E6	107
4	ROLES DE L'EPISSAGE DANS L'EXPRESSION DE L'ONCOPROTEINE E7	110
4.1	LE TRANSCRIT E6*IE7 FAVORISE LA TRADUCTION DE E7	110
4.2	LES ATG DE E6 ET E6*I SONT PREFERENTIELLEMENT RECONNUS PAR LE RIBOSOME	112
4.3	L'ESPACE INTERGENIQUE ENTRE E6*I ET E7 EST INDISPENSABLE A LA TRADUCTION DE E7	114
	DISCUSSION.....	116
1	ETUDE DE L'IMPACT DE E6*I SUR LE TRANSCRIPTOME CELLULAIRE	118
1.1	LES CLONES STABLES EXPRIMANT LES ISOFORMES E6.....	118
1.2	MODIFICATIONS PHENOTYPIQUES INDUITES PAR LES ISOFORMES E6.....	119
1.3	DEREGULATIONS DU TRANSCRIPTOME CELLULAIRE INDUITES PAR LES ISOFORMES E6.....	120
2	L'IMPACT DE E6*I D'HPV16 SUR LE METABOLISME DES ROS.....	121
2.1	ROS ET HOMEOSTASIE CELLULAIRE.....	121
2.2	LES ONCOPROTEINES E6 ET E7 DEREGULENT LE METABOLISME DES ROS.....	122
2.3	E6*I AUGMENTE LA PRODUCTION DE ROS	123
2.4	LE ROLE DE E6*I DANS LA CARCINOGENESE.....	125
3	EXPRESSION DE LA PROTEINE E6*I	126
3.1	DETECTION DE LA PROTEINE E6*I	126
3.2	LA PROTEINE E6*I INDUIT LA PRODUCTION DE ROS	127
3.3	ETUDE DE LA TRADUCTION DE L'ARNm E6*I PAR PROFILAGE POLYSOMIQUE.....	128
4	RELATION DYNAMIQUE ENTRE E6 ET E6*I.....	130
5	L'EPISSAGE DANS LA REGULATION DE L'EXPRESSION DE E7.....	131
	CONCLUSION GENERALE.....	135
	MATERIEL ET METHODES	137
	BIBLIOGRAPHIE.....	144
	ANNEXES.....	174
1	DONNEES COMPLEMENTAIRES (PAGET-BAILLY ET AL., 2019).....	175
2	AUTRES PUBLICATIONS	188
2.1	ETUDE DU METABOLISME ET DE LA REGULATION DES ARN VIRAUX.....	188
2.2	REGULATION DE L'EXPRESSION DE L'ONCOPROTEINE E6 PAR UN AGENT DEMETHYLANT, LE 5AZADC223	

Liste des figures

FIGURE 1 : ARBRE PHYLOGENETIQUE DES PAPILOMAVIRUS.	15
FIGURE 2 : ARBRE PHYLOGENETIQUE DES HPVHR.	16
FIGURE 3 : LE GENOME D'HPV16.	18
FIGURE 4 : HISTOLOGIE DU COL UTERIN ET SITES D'INFECTION AUX HPV.	20
FIGURE 5 : CYCLE VIRAL ET EXPRESSION SPATIALE DES PROTEINES VIRALES DANS L'EPITHELIUM.	22
FIGURE 6 : LES VIRIONS HPV.	23
FIGURE 7 : FACTEURS DE CLAIRANCE OU DE PERSISTANCE D'UNE INFECTION PAR UN HPVHR AU NIVEAU DU COL DE L'UTERUS.	24
FIGURE 8 : LA SUREXPRESSION DE E6 ET E7.	28
FIGURE 9 : STRUCTURE DE LA PROTEINE E6 D'HPV16.	30
FIGURE 10 : DEGRADATION DE pRb ET LIBERATION DES FACTEURS E2F.	38
FIGURE 11 : INCIDENCE DES CANCERS CAUSES PAR LES HPV.	42
FIGURE 12 : LES PRINCIPAUX ELEMENTS CIS ET TRANS REGULATEURS DE LA LCR D'HPV16. ...	45
FIGURE 13 : LE MECANISME D'EPISSAGE.	47
FIGURE 14 : LA REGULATION DE L'EPISSAGE.	48
FIGURE 15 : LE TRANSCRIPTOME D'HPV16.	49
FIGURE 16 : CONSERVATION DES SITES D'EPISSAGE E6*I CHEZ LES HPVHR.	51
FIGURE 17 : ELEMENTS CIS REGULATEURS DE L'ORF E6 D'HPV16.	53
FIGURE 18 : LES FACTEURS D'EPISSAGE REGULANT LE DEVENIR DU PREARNM E6E7.	54
FIGURE 19 : REGULATION ET MECANISME DE TRADUCTION DES PROTEINES E6, E7 ET E6*I.	58
FIGURE 20 : LA PROTEINE TRONQUEE E6*I D'HPV16.	61
FIGURE 21 : ETUDE DE LA PROLIFERATION DES CLONES STABLES.	92
FIGURE 22 : ETUDE DE LA VIABILITE DES CLONES STABLES SOUMIS A UN STRESS OSMOTIQUE. .	93
FIGURE 23 : ETUDE DE LA CAPACITE D'INVASION DES CLONES STABLES PAR BLESSURE/CICATRISATION.	94
FIGURE 24 : DETECTION DE L'H ₂ O ₂ PRODUIT PAR LES CLONES STABLES TRAITES AU DPI.	95
FIGURE 25 : DETECTION DE L'H ₂ O ₂ PRODUIT PAR LES CLONES U-E6*I TRAITES AU DPI.	96
FIGURE 26 : DETECTION DES ISOFORMES E6 DANS LES CLONES STABLES APRES INHIBITION DU PROTEASOME.	97
FIGURE 27 : DETECTION DES ISOFORMES E6 EXPRIMES A PARTIR DES VECTEURS MUTES OU NON SUR LES CODONS D'INITIATION DE LA TRADUCTION	98
FIGURE 28 : DETECTION DE L'H ₂ O ₂ PRODUIT PAR LES CELLULES U-2 OS EXPRIMANT LES DIFFERENTS VECTEURS E6.	99
FIGURE 29 : TECHNIQUE DE PROFILAGE POLYSOMIQUE.	100
FIGURE 30 : PROFIL DE POLYSOMES DES CELLULES U-2 OS EXPRIMANT E6 ET E6*I.	101
FIGURE 31 : PRESENCE DES ARNR DANS LES DIFFERENTES FRACTIONS DU GRADIENT.	102
FIGURE 32 : QUANTIFICATION DES TRANSCRITS E6 ET E6*I EXPRIMES TRANSITOIREMENT APRES FRACTIONNEMENT.	103
FIGURE 33 : QUANTIFICATION DES TRANSCRITS E6 ET E6*I APRES FRACTIONNEMENT PAR POLYSOME PROFILING DANS LES CLONES STABLES U-E6 ^{ALL} ET U-E6*I.	104
FIGURE 34 : EXPRESSION DES ARNM E6 ET E6*I DANS UNE COHORTE DE CANCERS DE L'OROPHARYNX HPV16+.	106

FIGURE 35 : EXPRESSION DE TRANSCRITS CELLULAIRES EN FONCTION DE L'EXPRESSION DES TRANSCRITS E6.	108
FIGURE 36 : EXPRESSION DE LA PROTEINE E7 D'HPV16 A PARTIR DES DIFFERENTS TRANSCRITS E6E7 DANS LA LIGNEE U-2 OS.	111
FIGURE 37 : EXPRESSION DE LA PROTEINE E7 D'HPV16 A PARTIR DES TRANSCRITS E6E7 SANS CODON D'INITIATION POUR E6 OU E6*I.....	113
FIGURE 38 : EXPRESSION DE LA PROTEINE E7 D'HPV16 A PARTIR D'UN TRANSCRIT E6*IE7 SANS ESPACE INTERGENIQUE ENTRE LES DEUX ORF.	115

Liste des tableaux

TABLEAU 1 : ANALYSE DES CARACTERISTIQUES DE PATIENTS ET DE LEUR TUMEUR HPV16+ EN FONCTION DE LEUR STATUT TRANSCRIPTIONNELLEMENT ACTIF.....	105
TABLEAU 2 : ANTICORPS PRIMAIRES UTILISES EN WESTERN BLOTTING.....	142
TABLEAU 3 : DIFFERENTIALLY EXPRESSED GENES U-CTL vs. U-E6ALL.....	175
TABLEAU 4 : DIFFERENTIALLY EXPRESSED GENES U-CTL vs. U-E6*I.	184
TABLEAU 5 : PCR AND QPCR PRIMERS.	185

Liste des abréviations

16E6, E6*I : E6, E6*I d'HPV16

18E6, E6*I : E6, E6*I d'hPV18

40S, 60S, 80S : coefficient de sédimentation en Svedberg

A

A2M : alpha-2-macroglobulin

aa : acide aminé

ADGRL4 (ELTD1) : adhesion G protein-coupled receptor L4

ADN : acide désoxyribonucléique

ADNc : ADN complémentaire

ARE : A/U rich element

ARN : acide ribonucléique

ARNcirc : ARN circulaire

ARNg : ARN génomique

ARNm : ARN messager

ARNlnc : long ARN non-codant

ARNnc : ARN non-codants

ARNr : ARN ribosomique

ARNt : ARN de transfert

ARNT2 : Aryl Hydrocarbon Receptor Nuclear Translocator 2

ATR/ATM : ataxia telangiectasia and Rad3-related protein/ataxia telangiectasia mutated

B

Bak : BCL2 antagonist/killer 1

Bax : BCL2 associated X

br : bas risque

BRCA1/2 : breast cancer 1/2

C

cIAP2 : cellular inhibitor of apoptosis 2

C-ter : extrémité carboxyterminale

CAT : catalase

CCL2 (MCP-1) : C-C motif chemokine ligand 2

CD : cluster of differentiation

CDK2 : cyclin dependent kinase 2

CFAP58 : Cilia And Flagella Associated Protein 58

CIN : cervical intraepithelial neoplasia

CKII : casein kinase II

CKIs : CDK inhibitors

CMH : complexe majeur d'histocompatibilité

CNN1 : calponine 1

CO₂ : dioxyde de carbone

CR1/2/3 : régions conservées 1/2/3

CRYAB (HSPB5): Crystallin a B

CTCF : CCCTC-binding factor

CTGF : Connective Tissue Growth Factor

CTL : contrôle

D

DICER1 : Double-Stranded RNA-Specific Endoribonuclease

DISC : death-inducing signaling complex

DLG : discs large homolog 1

DMEM : Dulbecco's Modified Eagle Medium

DNMT1/3a/3b : ADN méthyltransférase 1/3a/3b

DPI : Diphenyleneiodonium

E

E1-8 : early protein 1-8

E1A : adenovirus early region 1A

E2BS : E2 binding site

E2F : E2F transcription factor
E6AP : gène UBE3A, E6 associated protein

EGF : epidermal growth factor
EGFR : epidermal growth factor receptor
EZH2 : enhancer of zeste homolog 2

F
FADD : fas-associated protein with death domain

FAS : fas cell surface death receptor
FPKM : Fragments per Kilobase Million
FT : facteur de transcription

G
GJA1/5 : gap junction protein alpha 1/5, codant la connexine 43/40
GO : Gene Ontology

H
HaCaT : cultured human keratinocyte
HBV : virus de l'hépatite B

HMT : histone méthyltransférases
hnRNP : heterogeneous nuclear ribonucleoproteins

HPV : papillomavirus humain
HPV négatif : sans HPV, sans copie du virus
HPV positif : positif à HPV, avec des copies du génome viral
HPV positif, ARN positif : présence du génome HPV16 et des transcrits E6/E7

HR : haut risque
hTERT : telomerase reverse transcriptase

I
IARC : international agency for research on cancer

IC₅₀ : Concentration inhibitrice médiane
IDH : Indice de Développement Humain
IFN : interféron
IGFBP4 : Insulin Like Growth Factor Binding Protein 4
IL-1/4/6/8/13 : interleukine 1/4/8/13
IRF1/3/7 : interferon regulatory factor 1/3/7

ISG : interferon-stimulated genes

J
JSC : jonction squamo-cylindrique

K
KEGG : Kyoto Encyclopedia of Genes and Genomes
Krt 75 : kératine 75

L
L1/2 : late protein 1/2
LCR : long control region

M
MAGI1 : membrane-associated guanylate kinase inverted 1
MALAT1 : metastasis associated lung adenocarcinoma transcript 1
MCM7 : minichromosome maintenance complex component 7

MDM2 : mouse double minute 2
miR-375 : gène MIR375, microARN-375

miRNA : microARN
MMP3 : Stromelysin 1
MMP7 : matrice métalloprotéinase 7
MMP13 : Collagenase 3
MPT : modification post-traductionnelle
MUPP1 : gène MPDZ, multi-PDZ domain protein 1

MYCN : codant N-myc, neuroblastoma myc

N

N-ter : extrémité aminoterminal

NADP⁺/H : nicotinamide adénine dinucléotide phosphate forme oxydée /forme réduite

NES : nuclear export sequence

NFX1-91/123 : nuclear transcription factor, X-box binding 1 isoforme 91/123

NLS : nuclear localization signal

NMD : nonsense mediated mRNA decay

NOX2 : NADPH oxydase 2

NUP : Nucleoporin

O

Oct1 : gène POU2F1, octamer-binding transcription factor 1

OMS : organisation mondiale de la santé

P

p14^{ARF} : protéine codée par le gène CDKN2A, alternative reading frame

p16^{INK4A} : protéine codée par le gène CDKN2A, inhibitors of CDK4

p21/p21^{CIP1} : codée par CDKN1A, CDK-interacting protein 1

p27^{KIP1} : codée par CDKN1B, cyclin-dependent kinase inhibitor p27

p300/CBP : codée par EP300, E1A binding protein p300

p38^{MAPK} : codée par MAPK14, mitogen-activated protein kinases

P4HA1: Prolyl 4-Hydroxylase Subunit Alpha 1

p53 : codée par TP53

p670 : promoteur tardif d'HPV16

p97 : promoteur précoce d'HPV16

PABPC : cytoplasmic poly(a) binding proteins

pAE/L : site de polyadénylation précoce/tardif

PAGE1 : PAGE family member 1

PARP : poly(ADP-ribose) polymérase

PATJ : PALS1-associated tight junction protein

pb : paire de bases

PBM : PDZ binding motif

PCR : polymerase chain reaction

PDGFB : platelet derived growth factor subunit B

PDZ : PSD-95, DLG et ZO-1

pE8 : promoteur alternatif d'HPV16 (p1127)

PFKFB4 : 6-phosphofructo-2-kinase / fructose-2,6-biphosphatase 4

PICSAR (NLC1-C) : P38 Inhibited Cutaneous Squamous Cell Carcinoma Associated LincRNA

PIK3CA : phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha

PML : ProMyelocytic Leukemia

pRb : codée par le gène RB1, retinoblastoma

PRC1/2 : polycomb repressive complex 1/2

PTEN : phosphatase and tensin homolog

PTPN3 : protein tyrosine phosphatase, non-receptor type 3

PTPN13 : protein tyrosine phosphatase, non-receptor type 13

PV : papillomavirus

Q

qPCR : PCR quantitative

R

RAC2 : Rac Family Small GTPase 2

RISC : RNA-induced silencing complex

ROS : Reactive Oxygen Species, espèces réactives de l'oxygène

RPLP0 : Ribosomal Protein Lateral Stalk Subunit P0

S

SA : site accepteur d'épissage

SCRIB : scribbled planar cell polarity protein

SD : site donneur d'épissage

siARN : small interfering ARN

SMAD4 : MAD homolog 4

SMOC1 : SPARC related modular calcium binding 1

SR : serin arginin rich protein

STC1 : Stanniocalcin 1

SV40 : virus simien 40

SYCP2 : Synaptonemal Complex Protein 2

SYP : Synaptophysine

T

tBid : protéine Bid tronquée

TBP : TATA box binding protein

TCGA : Tumor Cancer Genome Atlas

TEM: transition épithélio
mésenchymateuse

TERC : telomerase RNA component

TGFβ : transforming growth factor beta

TLR9 : toll like receptor 9

TNFAIP6 : TNF alpha induced protein 6

TNFR1 : gène TNFRSF1A, tumor necrosis factor receptor 1

TNFα : tumor necrosis factor alpha

TNNC1 : Troponine C

TP53/63 : gène codant p53/63

TRAIL : gène TNFSF10, tumor-necrosis-factor related apoptosis inducing ligand

U

U-2 OS : human bone osteosarcoma epithelial

USF1/2 : upstream transcription factor 1

UTR : régions non traduites des ARNm

UVB : ultraviolets B

V

VADS : voies aérodigestives supérieures

VEGF : vascular endothelial growth factor

VIH : virus de l'immunodéficience humaine

X

XRCC1 : X-Ray repair cross complementing 1

Y

YY1 : yin yang 1

Z

ZO-1 : gène TJP1, zonula occludens 1

Introduction

1 Les papillomavirus

1.1 Origine et classification des papillomavirus

Les papillomavirus (PV) sont des petits virus non-enveloppés à ADN circulaire double brin. La famille des *Papillomaviridae* se distingue par une coévolution extrêmement ancienne avec ses espèces hôtes. Des dizaines de PV infectent des mammifères autres que l'Homme mais aussi d'autres vertébrés tel que des reptiles et des oiseaux (McBride, 2017). La découverte récente d'un PV chez une espèce de poisson repousse encore d'avantage cette coévolution avec les vertébrés et fait remonter l'évolution des PV à au moins 420 millions d'années (López-Bueno et al., 2016).

Aujourd'hui, plus de 300 génotypes de PV sont connus dont au moins 200 infectant spécifiquement l'Homme : les Papillomavirus Humain (HPV) (Van Doorslaer, 2013). Plusieurs organismes comme l'International Committee on Taxonomy of Viruses (ICTV) ou le Papillomavirus Episteme (PaVE) ont la charge d'établir une base de données regroupant chaque PV, leur séquence nucléotidique annotée et leur phylogénie. Il est fort probable que bien plus de génotypes existent pour chaque espèce identifiée jusque-là. La grande diversité de la famille des *Papillomaviridae* est représentée par un arbre phylogénétique en figure 1. Une nomenclature universelle de la phylogénie a été proposée par (de Villiers et al. 2004; Bernard et al. 2010). On y retrouve la hiérarchisation typique en famille>genre>espèce>type>sous-type>variant. HPV16 fait partie du genre alpha espèce 9 type 16 avec des variants européen et africain par exemple. La classification est établie en fonction du niveau de divergence nucléotidique des génomes des différents PV, historiquement du gène *L1* (Late). Les limites arbitraires sont placées à 45%, 10%, 2% et moins de 2% de divergence entre deux génomes pour un genre, type, sous-type et variant, respectivement (de Villiers, 2013) (Bzhalava et al., 2015). L'arbre phylogénétique en figure 1 est réalisé sur la divergence des séquences des gènes E1, E2, L1 et L2.

Il est intéressant de noter que certains genres contiennent des HPV ainsi que des PV non-humains, indiquant que la coévolution avec l'espèce hôte n'est pas la seule variable expliquant la divergence des PV. Une spécification et une adaptation ancestrale en fonction du site anatomique infecté, et précédant la coévolution, explique cette proximité atypique entre des PV infectant différentes espèces (Bravo and Alonso, 2007). Les HPV présentent en effet un tropisme pour les épithéliums malpighiens, cutanés ou muqueux. Les genres beta et gamma regroupent la majeure partie des HPV à tropisme cutané avec une centaine de types.

Néanmoins, des HPV à tropisme cutané sont retrouvés dans les cinq principaux genres. Seul le genre alpha comprend les HPV capables d'infecter les muqueuses. Ces muqueuses sont les voies aéro-digestives supérieures avec l'oropharynx, et la région ano-génitale avec le col de l'utérus et le canal anal notamment, mais aussi le vagin, la vulve, le pénis et le scrotum (Doorbar et al., 2012).

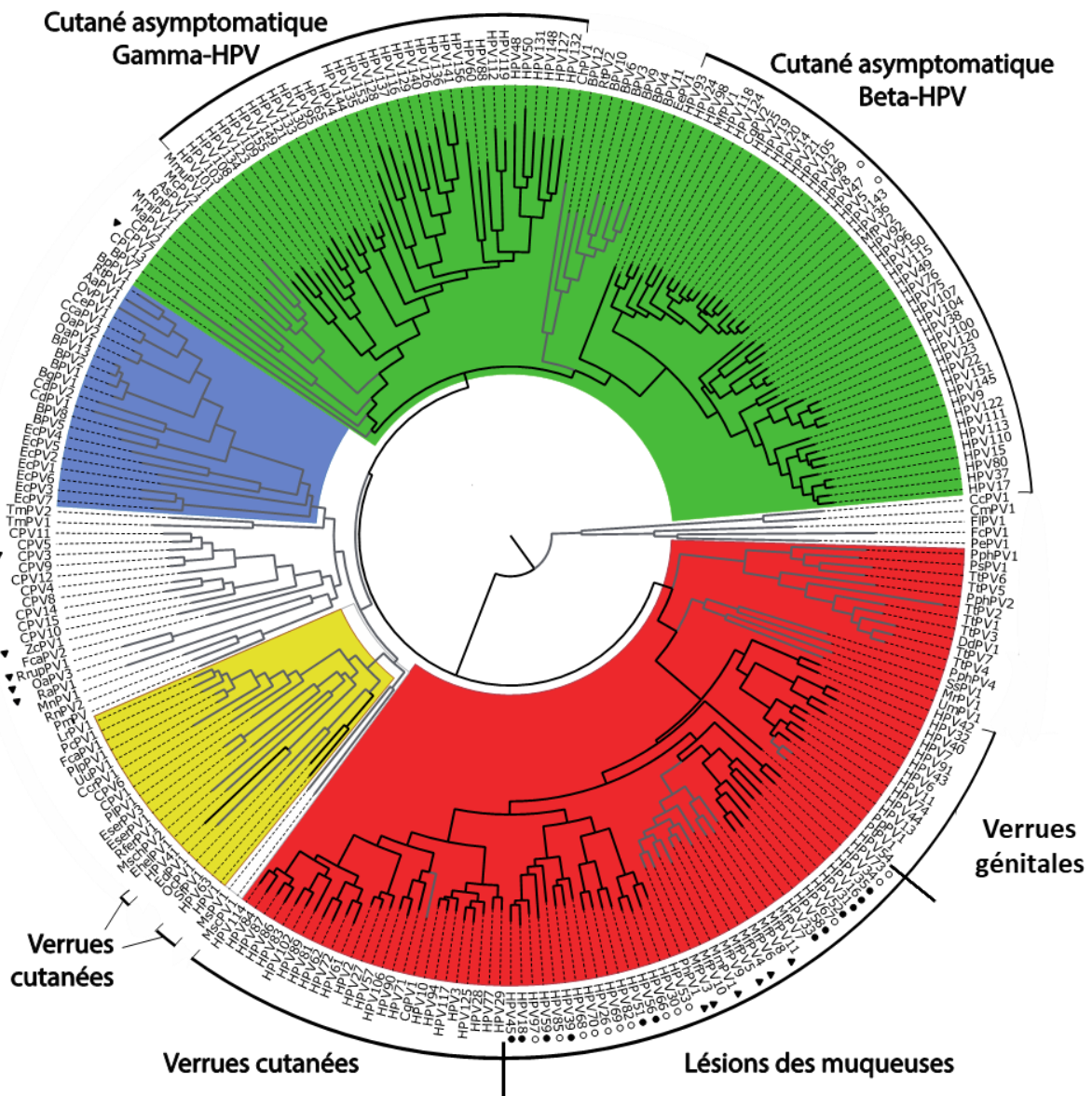


Figure 1 : Arbre phylogénétique des papillomavirus.

L'arbre est généré par comparaison nucléotidiques des séquences E1E2L1L2 de 263 PV. Les couleurs représentent, en rouge Alpha et Omikron PV ; en vert Beta, Gamma et Xi PV ; en jaune Lambda et Zeta PV et en Bleu Delta et Zeta PV. Sur l'arbre, les HPV sont en noir et les autres PV en gris. Le potentiel carcinogène d'un PV ou HPV est indiqué par un triangle ou un rond, respectivement en regard du type. Pour les HPV, on distingue les carcinogènes démontrés (rond noir plein), des carcinogènes probables (cercle noir) selon l'IARC (International Agency for Research on Cancer) (adapté de Bravo et Félez-Sánchez, 2015).

1.2 Classification des HPV à haut risque oncogénique

Au début des années 1980 émergent les premières preuves d'un lien entre les HPV et certains cancers génitaux car de l'ADN d'HPV est retrouvé dans de nombreux échantillons tumoraux (Gissmann et al., 1984). Il sera découvert par la suite que certains HPV du genre alpha sont les agents étiologiques de ces cancers. De fait, l'étude des HPV muqueux du genre alpha domine largement la recherche clinique et fondamentale car ils sont à l'origine de la quasi-totalité des cancers induits par HPV.

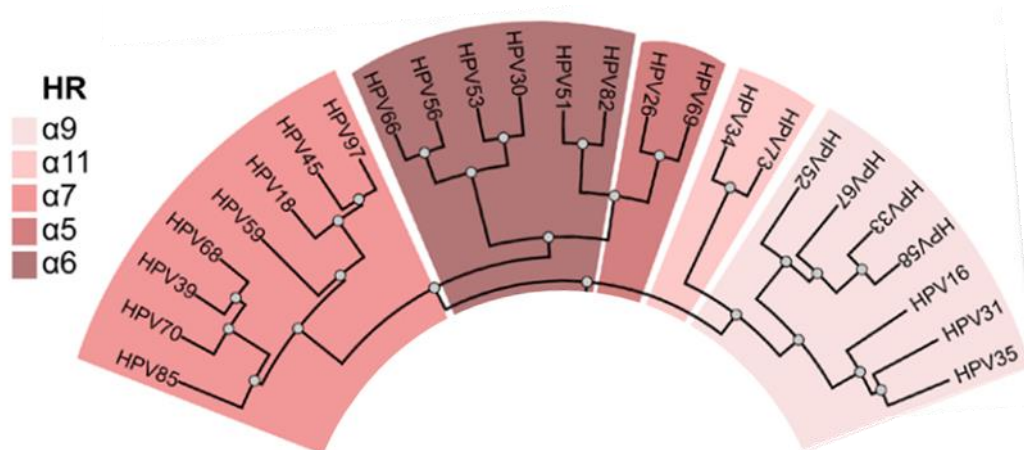


Figure 2 : Arbre phylogénétique des HPVhr.

L'arbre représente les HPV du genre alpha avec les 12 principaux types carcinogènes ainsi que les autres types possiblement haut risque, comme le laisse suggérer leur proximité génétique avec les douze premiers. Les couleurs représentent l'espèce à laquelle appartient un génotype donné, les HPVhr étant tous du genre α (adapté de Auslander et al. 2019).

C'est pour cela qu'en plus de la classification phylogénétique, les HPV du genre alpha sont autrement classés en fonction de leur potentiel carcinogène. On distingue les HPV à haut risque oncogène (HPVhr) associés à un risque de cancer invasif et les HPV à bas risque (HPVbr) oncogène associés à des lésions bénignes. La première classification par l'IARC des HPVhr en 1995 comprenait uniquement HPV16 et 18, plus quelques autres types proches considérés comme potentiellement carcinogènes. Ce nombre n'a cessé de croître au fur et à mesure des études épidémiologiques et des découvertes fondamentales sur les caractéristiques oncogéniques des HPV et l'étiologie des cancers associés. Les HPVhr étaient au nombre de 12 avérés dans la classification de l'IARC en 2005. Aujourd'hui, en recoupant les informations phylogéniques et épidémiologiques avec de plus en plus d'études sur les autres localisations de cancers induits par HPV que le col de l'utérus, notamment l'oropharynx et le canal anal, ils sont encore plus nombreux (WHO-IARC, 2017). Les HPVhr sont à l'heure actuelle classés en groupes 1 carcinogènes avérés avec les 16, 18, 31, 33, 35, 39, 45, 51, 52,

56, 58, 59, groupe 2A possiblement carcinogène avec le 68, et groupe 2B avec les 26, 30, 34, 53, 66, 67, 69, 70, 73, 82, 85, 97 (Bouvard et al., 2009). Le pouvoir carcinogène des types du groupe 2B est difficile à évaluer du fait de leur très faible prévalence.

La découverte du lien entre cancer du col de l'utérus et HPV a abouti à l'obtention du prix Nobel de physiologie et de médecine par le professeur Harald zur Hausen en 2008. Pr Zur Hausen avait pressenti l'implication des HPV dès 1974. L'ICTV dans sa taxonomie virale a aussi nommé à titre honorifique l'ordre regroupant la famille des *Papillomaviridae* en *Zurhausenvirales*. On sait aujourd'hui la place plus qu'importante qu'occupent les pathogènes dans le développement des cancers chez l'Homme, qu'il soit sain ou immunodéprimé, et que le pathogène soit le facteur causal (zur Hausen, 2009) ou un cofacteur pro-tumoral (Porta et al., 2011). Au total, 20% des cancers sont directement causés par des pathogènes.

La famille de virus la plus proche des *Papillomaviridae* est celle des *Polyomaviridae*. Elle regroupe une centaine de génotypes dont plus d'une dizaine sont connus pour infecter l'Homme. Bien qu'ayant un impact moindre sur la santé humaine (la plupart des infections sont asymptomatiques), dans de rares cas les polyomavirus sont également à l'origine d'infections pouvant mener à des cancers (Moens et al., 2017).

Sur le plan pathologique, les quelques 200 autres génotypes HPVbr ne sont pas en reste. Ils sont à l'origine d'une comorbidité considérable, peuvent favoriser l'apparition de cancers ou en être clairement à l'origine chez les populations fragiles, immunodéprimées ou à contexte génétique propice. Citons notamment les verrues génitales et les verrues vulgaires ; les cas de papillomes, de papillomatoses respiratoires, et d'épidermodysplasie verruciforme par les HPV 5 et 8 ; les rares cas de cancers ne présentant qu'une infection par un HPVbr, généralement HPV6 ou 11 (Cornall et al., 2013) (Egawa and Doorbar, 2017).

Le génotype HPV16 est le plus fréquemment retrouvé dans les cancers. Il cause à lui seul 60% des cancers du col de l'utérus et 90% des cancers de l'oropharynx et du canal anal induits par les HPVhr.

2 Histoire naturelle de l'infection par HPV16 menant au cancer du col de l'utérus

2.1 Le génome d'HPV16

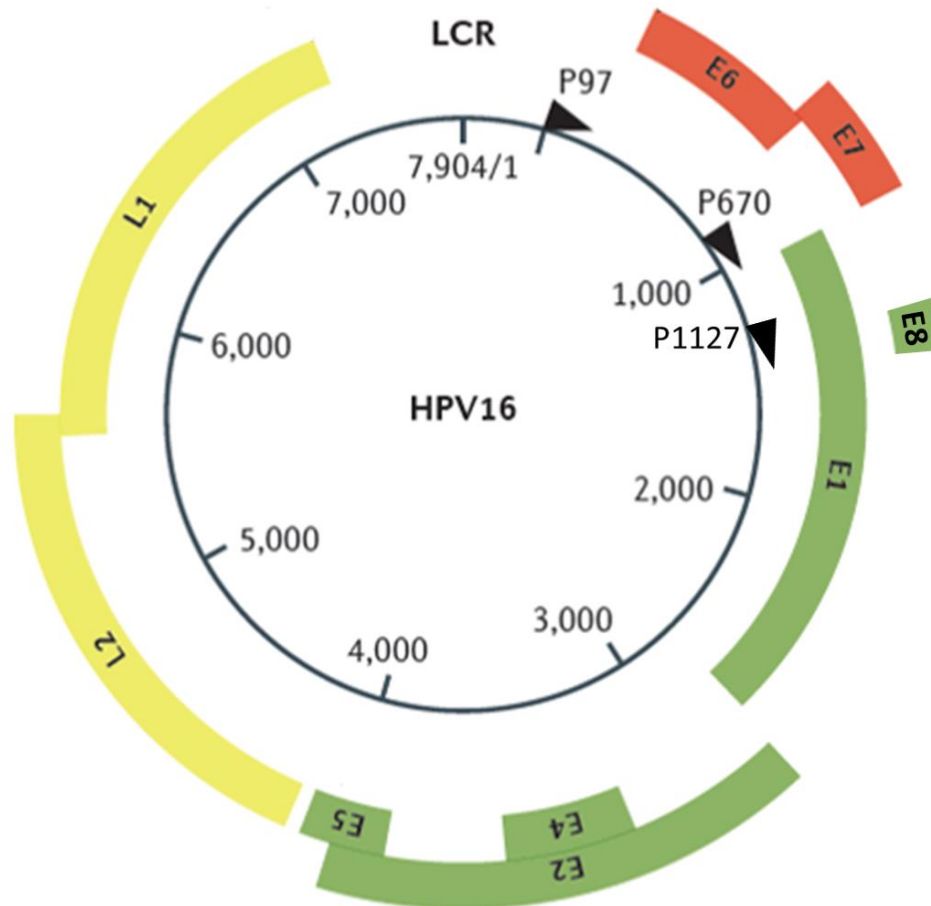


Figure 3 : Le génome d'HPV16.

Un seul brin de l'ADN viral est codant. Après la région cis-régulatrice LCR qui contient l'origine de réplication et le promoteur p97, on retrouve en rouge les ORF des oncogènes E6 et E7, puis les ORF des autres gènes précoces E1, E2, E4 et E5. Le promoteur tardif p670 est situé dans l'ORF E7. Enfin, viennent les ORF des gènes tardifs L2 et L1. Les sites de polyadénylation précoce et tardif se trouvent en aval de l'ORF E5 et L1, respectivement. Adapté de (Doorbar, 2006).

Le génome de 7906 pb (paires de bases) d'HPV16 est composé de 7 gènes précoces E6, E7, E1, E2, E4 et E5 (E pour early), de 2 gènes tardifs L1 et L2 (L pour Late), et de la LCR (Long Control Region) non codante. L'ORF (Open Reading Frame) E8 et son promoteur n'ont été découvertes que récemment (Straub et al., 2015). Les gènes précoces codent les protéines nécessaires au cycle de vie du virus et/ou à la carcinogenèse, comme il sera détaillé par la suite. Les gènes tardifs codent les protéines formant la capside et jouant un rôle dans l'entrée du virus dans la cellule. La LCR quant à elle, contient l'origine de réplication et la majorité des éléments *cis* régulateurs nécessaires à la régulation transcriptionnelle des 9 ORF situées en aval. Elle se termine par le promoteur p97 où débute la transcription des gènes

précoces. La transcription des gènes tardifs se fait à partir du promoteur p670 dans l'ORF E7. La transcription des ORF viraux se termine pour les gènes précoces et tardifs au niveau du site de polyadénylation pAE en position 4215 et pAL en position 7321, respectivement (Groves and Coleman, 2015).

La présence de seulement trois promoteurs pour neuf ORF implique la présence de nombreux éléments *cis* régulateurs au sein même des ORF afin d'assurer la régulation post-transcriptionnelle de l'expression des gènes viraux. Les ORF possèdent notamment de très nombreux sites accepteurs et donneurs d'épissage qui permettront la production d'un transcriptome viral vaste et complexe (J. Chen et al. 2014).

2.2 Cycle viral productif d'HPV16

Deux modèles permettent d'expliquer l'entrée des HPVhr au niveau du col de l'utérus. Le premier à avoir été décrit propose que des microlésions dans la « zone de transformation » permettent au virus d'accéder aux cellules épithéliales basales. Cette zone est composée d'un épithélium malpighien pluristratifié appelé exocol (Roberts et al., 2007) (figure 4). Le deuxième propose que les HPVhr présentent un tropisme pour une population discrète de progéniteurs épithéliaux située à la JSC (Jonction Squamo-Cylindrique) formant une monocouche de « cellules cuboïdes » (Herfs et al., 2012). Le développement des lésions précancéreuses au niveau de la JSC et la correspondance des profils moléculaires entre cellules cuboïdes et cellules de carcinomes/adénocarcinomes associés aux HPVhr semblent indiquer que cette population est à l'origine des cancers cervicaux (Herfs et al., 2013a). Il reste cependant vrai que la zone de transformation est le lieu de nombreuses infections par HPV, mais dans l'ensemble pas celles évoluant en cancer (Herfs et al., 2013b).

Concernant les autres sites anatomiques d'infection par un HPV muqueux, les cellules permettant l'entrée du virus n'ont pas été clairement identifiées. La signature moléculaire propre à la JSC n'est pas retrouvée dans des cellules issues de lésions vaginales, vulvaires ou péniennes positives à HPV (Herfs et al., 2012). La jonction anorectale est similaire à la JSC du col utérin et partage une origine embryologique commune avec une signature moléculaire proche mais on n'y retrouve pas cette zone discrète unistratifiée de cellules cuboïdes (Yang et al., 2015). Contrairement à la carcinogenèse du col, la survenue d'un cancer de l'oropharynx ne semble pas précédée par le développement de lésions précancéreuses. Nous manquons d'une compréhension de l'étiologie des cancers de l'oropharynx induits par HPV alors qu'ils deviennent prédominants dans les pays développés, dépassant maintenant parfois les cas de

cancer cervicaux (Berman and Schiller, 2017). L'organisation histologique et le phénotype spécifique des cellules cuboïdes du col utérin les rendent extrêmement propices aux infections se transformant en cancer et semblent expliquer pourquoi les cancers cervicaux sont sur-représentés par rapport aux autres cancers associés aux HPV.

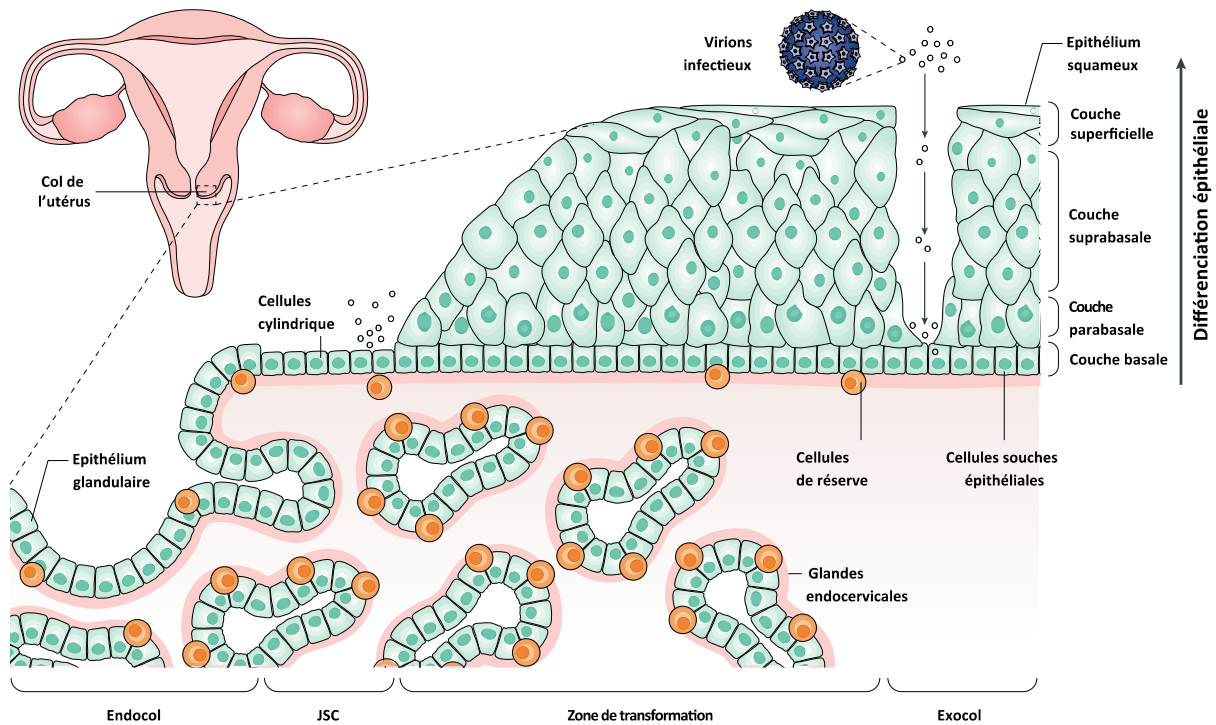


Figure 4 : Histologie du col utérin et sites d'infection aux HPV.

En haut à gauche, représentation du tractus génital féminin. La zone en pointillés est représentée schématiquement zoomée et sous forme de coupe frontale du col de l'utérus et de son voisinage. On y trouve de gauche à droite, i) l'endocol et son épithélium monostratifié de type glandulaire, ii) la JSC avec les cellules cuboïdes, et iii) l'exocol pluristratifié. Les virions infectieux (cercles bleus) sont représentés au niveau des sites d'infections, à savoir la lame basale de l'exocol mise à nu suite à un microtraumatisme, et les cellules cuboïdes de la JSC. Un troisième site d'infection possible au niveau de l'épithélium glandulaire de l'endocol n'est pas montré. Adapté de (Schiffman et al., 2016).

L'infection par HPV16 va donc débiter par la reconnaissance à la surface des progéniteurs épithéliaux d'héparanes sulfates servant d'ancrage au virus en interagissant avec la protéine de capsid L1 (Johnson et al., 2009). La protéine L2 est alors isomérisée de sorte que son extrémité N-terminale soit rendue accessible à des protéases (furine et ProteinConvertase5/6) (Kines et al., 2009). La nouvelle extrémité se lie alors à un récepteur secondaire transmembranaire. Plusieurs ont été identifiés dont l'intégrine $\alpha 6$, le récepteur à l'EGF (Epidermal Growth Factor), les tétraspanines ou encore l'annexine A2 (Graham, 2017). Le virus est ensuite endocyté par un mécanisme non-canonique de micropinocytose (endocytose non spécifique de milieu extra-cellulaire). Le virus se retrouve alors dans des endosomes lissent indépendants des clathrines et cavéolines et fusionnent avec les endosomes

multivésiculaires. Les conditions de plus en plus acides entraînent la dissociation de la capsid. Les protéines L2 restent attachées au génome viral, ainsi que quelques copies de L1. Leurs interactions avec le complexe rétrograde et la γ -secrétase permettent alors au complexe protéine/ADN viral d'échapper à la voie de dégradation lysosomale et de continuer leur transport rétrograde vers le Trans Golgi. Le rôle du transport transgolgien reste encore mal compris. Ensuite, la localisation nucléaire finale des épisomes HPV va nécessiter le processus de mitose pour deux raisons. Premièrement, la réorganisation mitotique du Golgi place le virus dans une vésicule et la vésicule est transportée grâce aux microtubules pour se positionner au plus proche du noyau (DiGiuseppe et al., 2017). Deuxièmement, le complexe HPV ADN/protéine est incapable de passer la membrane nucléaire par les CPN (Complexe de Pore Nucléaire) (Aydin et al., 2014). Ce qui donne un argument de plus en faveur du tropisme des HPV pour les cellules basales non-différenciées capables de se diviser (Pyeon et al., 2009).

Une fois dans le noyau, le génome viral s'associe aux corps nucléaires PML (ProMyelocytic Leukemia) et devient enfin transcriptionnellement actif (Day et al., 2004). Tout commence avec la production de la protéine E2 (et E1). E2 va remplir trois rôles *via* son action *trans*-régulatrice en se fixant sur différents sites de la LCR. Elle va d'abord réguler négativement l'expression de toutes les protéines précoces dans les cellules basales. Pour cela E2 se fixe sur E2BS#3 et #4 (E2 Binding Site) ce qui empêche la fixation des facteurs généraux de la transcription et réprime fortement le promoteur précoce p97 (Maglennon et al., 2011). Cette faible activité virale est une des raisons pour lesquelles l'infection à HPVhr est peu immunogène (Westrich et al., 2017) (Peh et al., 2002). En se fixant sur E2BS#2, E2 permet ensuite le recrutement de E1 qui possède une activité hélicase et permet la réplication du génome viral, en coopération avec la machinerie de réplication cellulaire (Sanders and Stenlund, 1998). Enfin, E2 permet l'ancrage des épisomes viraux aux chromosomes cellulaires (Iftner et al., 2017). Ainsi, les épisomes répliqués pendant la mitose de la cellule hôte sont répartis équitablement entre les cellules filles. Une cellule infectée reste donc basale et l'autre va migrer et se différencier dans l'épithélium pour poursuivre le cycle viral. Les protéines E6 et E7 vont être faiblement exprimées dans les cellules basales (figure 6) et vont participer au maintien des épisomes viraux (Park and Androphy, 2002).

Dans les couches parabasales et suprabasales, le cycle viral se poursuit par une étape d'amplification du génome viral pour arriver à 1000 copies par cellule environ (Groves and Coleman, 2015). Pour se faire, le changement de phénotype des kératinocytes en

différenciation et migrant dans l'épithélium va permettre l'expression accrue des protéines E6 et E7 (Egawa et al., 2015) (Stoler et al., 1992). E6 va lier et entrainer la dégradation de p53 (Scheffner et al., 1993). Cela va permettre une dérégulation des points de contrôle du cycle cellulaire et va empêcher l'apoptose des cellules infectées normalement non prolifératives. E6 va également se lier et entrainer la dégradation de protéines à domaine PDZ, régulant la signalisation et la polarité cellulaire pour favoriser le cycle viral (James and Roberts, 2016) (Nicolaidis et al., 2011). E7 va lier et entrainer la dégradation de pRb (protéine du Rétinoblastome), libérant le FT (Facteur de Transcription) E2F (Dyson et al., 1989). E2F est alors libre d'activer l'expression des cyclines A et E et des autres gènes nécessaires à l'activation du cycle cellulaire (Jones et al., 1997). Nous détaillerons les deux oncoprotéines et leurs effets de manière plus exhaustive dans la partie carcinogénèse.

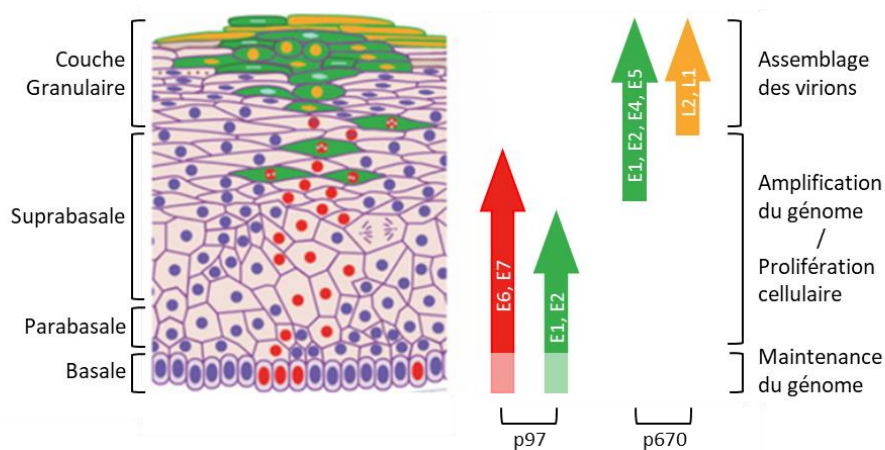


Figure 5 : Cycle viral et expression spatiale des protéines virales dans l'épithélium.

E2 réprime l'expression des gènes précoces à partir du promoteur précoce p97 dans les cellules basales et assure la maintenance des épisomes viraux. L'expression des gènes précoces augmente avec la migration au niveau para et suprabasale entrainant la prolifération des cellules infectées et l'amplification des épisomes viraux. La différenciation cellulaire change la régulation transcriptionnelle et posttranscriptionnelle des messagers viraux. Dans les couches supérieures, E1, E2, E4 et E5 assurent la fin de l'amplification virale tandis que l'expression de L1 et L2 permet la formation de nouveaux virions infectieux. Adapté de (Doorbar, 2006).

La différenciation des kératinocytes entraine aussi l'expression des protéines E4 et E5, amplifiant encore d'avantage la réplication virale par un mécanisme de « cercle roulant » (Flores and Lambert, 1997). En effet, E1^E4 va permettre de bloquer le cycle cellulaire en G2 et va s'associer avec E1 et E2 pour découpler la réplication virale (Egawa et al., 2017). E5 joue également un rôle dans cette amplification (Genther et al., 2003), même si son rôle principal est plus axé sur l'échappement immunitaire..

Enfin, quand les kératinocytes atteignent la couche granulaire, on observe la production des protéines tardives L2 puis L1. Elles sont importées dans le noyau et

s'assemblent en capsomères pentamérique. Le génome viral est encapsidé et il y a formation du virion en forme d'icosaèdre à 72 faces (figure 7) (Graham, 2017). E4 va également intervenir dans cette étape finale en restructurant le cytosquelette d'actine pour favoriser l'excrétion des virions (Wang et al., 2004). Les virions sont alors libérés sans lyse cellulaire quand les kératinocytes meurent et desquament. Ces nouveaux virions vont potentiellement pouvoir infecter de nouveaux épithéliums muqueux.

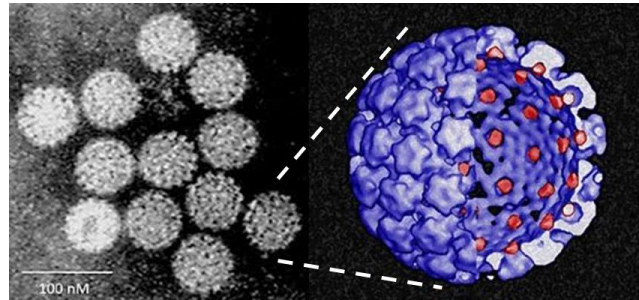


Figure 6 : Les virions HPV.

A gauche, des virions HPV observés par microscopie électronique. A droite, représentation bio-informatique de la capside HPV. En bleu sont représentés les capsomères constitués d'un pentamère de L1 et en rouge sur la coupe une copie de la protéine L2 au centre de chacun des capsomères.

On ne peut que constater ici l'adaptation du virus à son hôte, capable de détourner les processus cellulaires et l'organisation du tissu épithélial à son avantage en étant faiblement détectable par le système immunitaire.

2.3 Persistance de l'infection par HPV16

Une infection par un HPV au niveau cervical dure en moyenne un an avant d'être éliminée (Einstein and Burk, 2001). Sur la totalité des infections, 90% vont être éliminées avec succès par l'hôte en moins de 3 ans (Shulzhenko et al., 2014). Mais une fraction de ces infections ne va pas être éliminée par l'hôte et va persister. Plus l'infection persiste, plus l'expression des protéines virales dans les cellules basales va perturber l'homéostasie cellulaire. Des anomalies cytologiques apparaissent, ainsi que des lésions dysplasiques précancéreuses au niveau histologique. Ces dysplasies appelées CIN (Cervical Intra epithelial Neoplasia) sont classées par grade : légères (CIN 1), moyennes (CIN 2) et sévères (CIN 3) (Schiffman et al., 2016). Ces lésions se traduisent par un changement progressif de l'épithélium avec un envahissement par des cellules proliférantes de moins en moins différenciées. La perte d'homéostasie tissulaire rend à terme l'infection non-productive et environ 1% des CIN1 aboutissent à un cancer cervical.

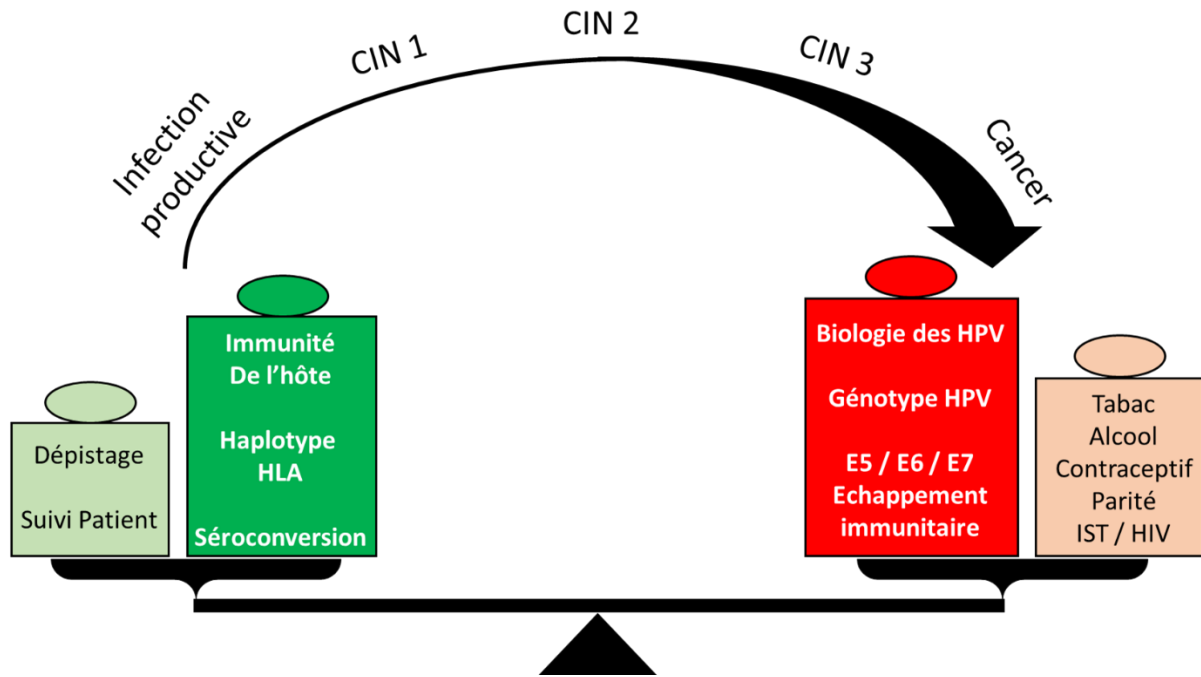


Figure 7 : Facteurs de clairance ou de persistance d'une infection par un HPVhr au niveau du col de l'utérus.

Un ensemble de facteurs liés au virus est à l'origine de la persistance. Le premier est évidemment le génotype HPV, hr ou non, à l'origine de l'infection mais aussi la charge virale. Viennent ensuite les facteurs liés à la biologie des HPV. L'expression des protéines virales est très faible. E2 va réprimer l'expression des protéines précoces dans les cellules basales en inhibant le promoteur p97 (Maglennon et al., 2011). La production des protéines tardives plus immunogènes pouvant être présentées par le CMH (Complexe Majeur d'Histocompatibilité) ne se fait que dans les couches supérieures de l'épithélium, là où les cellules immunitaires sont les moins présentes (Stanley, 2012). L'utilisation de codons rares par les HPV va aussi permettre une expression minimale des protéines virales. En effet, les huit ORF HPV codent leurs acides aminés à l'aide de codons peu présents dans les gènes humains. Ce « biais d'usage de codons » fait du contexte cellulaire en ARNt (ARN de transfert) un facteur limitant la traduction des gènes viraux (Zhao and Chen, 2011). Ce biais remplit plusieurs fonctions dans la biologie des HPV (cf. 4.3. Régulation traductionnelle de E6 et E7). Plusieurs études ont d'ailleurs démontré que l'optimisation de ces codons pour augmenter l'expression des protéines virales, mais sans changer leur séquences, entraîne une meilleure réponse immunitaire (Samorski et al., 2006) et séroconversion (Cheung et al., 2004) dans des modèles murins.

Les protéines virales vont aussi activement supprimer ou empêcher la réponse immunitaire en inhibant la réponse antivirale innée, l'inflammation et la présentation

d'antigène. E6 va réprimer l'expression des IFN (Interférons) en interagissant avec IRF-3 et ISGF3. E6 et E7 vont également réprimer l'expression de TLR-9 (Toll-Like Receptor 9) reconnaissant les ADNdb exogènes (Hasan et al., 2007). E6 et E7 vont aussi inhiber directement les voies de signalisation IFN, TNF α et TGF- β et d'autres cytokines pro-inflammatoires comme IL-8 (InterLeukine-8), IL-6 et IL-10 (Westrich et al., 2017) (Ghittoni et al., 2010). E7 va également réprimer la protéine TAP1 (Transporteur Associated with antigen Protein 1) impliquée dans le transport et chargement d'antigènes au CMH de classe I (Vambutas et al., 2000). E5 de son côté va aussi réprimer la présentation d'antigènes par le CMH de classe I, évitant la reconnaissance par les lymphocytes T cytotoxiques (Ashrafi et al., 2006), et réprimer CD1d (Cluster of Differentiation-1d) empêchant l'activation des cellules NK (Natural Killer) (Miura et al., 2010).

Le rôle de l'immunité de l'hôte dans le développement du cancer cervical est également prédominant. La variabilité des haplotypes HLA entre les individus joue aussi un rôle dans la susceptibilité à la persistance (Paaso et al., 2019). Le site d'infection est aussi important dans l'absence de détection immunitaire. En effet, HPV16 infecte les cellules épithéliales au-dessus de la lame basale, là où les cellules immunitaires sont les moins présentes (Sasagawa et al., 2012). De plus, aucune des protéines virales n'est sécrétée et l'infection productive libère les virions néoformés par desquamation de la couche cornée supérieure de l'épithélium, sans cycle lytique immunogène.

Viennent enfin les facteurs environnementaux. Pour les cancers cervicaux sont retrouvés le tabac et l'alcool (Weiderpass et al., 2001) (Smith et al., 2012) comme pour la majorité des cancers, mais aussi la prise de contraceptifs hormonaux et le nombre de grossesses arrivées à terme (Thun et al., 2017). D'autres infections sexuellement transmissibles favorisent le développement des cancers cervicaux, le VIH (Virus de l'Immunodéficience Humaine) notamment. Une étude sur des patientes immunodéprimées VIH positives sans traitement antirétroviral a montré que ces femmes présentes un taux d'infection à HPVhr quatre fois supérieur à la normal et développent également beaucoup de cancers cervicaux *via* d'autres génotypes que le 16 et le 18 (Sasagawa et al., 2012). Pour finir, les facteurs socio-économiques, l'éducation, l'hygiène, dans une moindre mesure les pratiques sexuelles, mais surtout l'accès à la prévention et aux soins sont des facteurs liés aux cancers HPV (IARC).

L'enjeu de santé publique mondial est donc de pouvoir dépister quelles infections vont persister et à terme se transformer en cancer afin de les traiter. Nous disposons de moyens de prévention secondaire pour lutter contre le développement des lésions précancéreuses. Il n'existe pas de dépistage spécifique pour les autres cancers associés aux HPV. L'organisation mondiale de la santé (OMS, WHO) recommande le dépistage des lésions du col chez les femmes de 30 à 50 ans par l'une des méthodes suivantes : i) test visuel par colposcopie à l'acide acétique faisant apparaître les dysplasies, ii) examen cytologique sur frottis cervico-utérin ou, iii) « test HPV » sur frottis cervico-utérin. Ces trois tests n'ont qu'une valeur indicative permettant le triage des femmes, soient négatives ne nécessitant pas un suivi, soient positives nécessitant une surveillance. La comparaison des pays ayant mis en place une politique de dépistage (Bray et al., 2005) et ceux qui ne l'ont pas fait (Bray et al., 2013) montre clairement leur efficacité à diminuer la prévalence des lésions précancéreuses, des cancers cervicaux et des décès associés. Le traitement chirurgical des lésions du col détectées à temps suffit généralement à stopper leur progression (Bray et al., 2018).

Tous ces éléments permettent donc à certaines infections de persister. C'est durant ce temps prolongé que les oncogènes E6 et E7 vont être surexprimés et entraîner des modifications réversibles et irréversibles provoquant d'abord des lésions précancéreuses et à terme l'apparition d'un cancer. En définitive, les propriétés d'échappement immunitaire de E6 et E7 (et E5) sont premières et au moins aussi importante que leurs propriétés oncogéniques.

3 Rôles des oncoprotéines virales dans la carcinogénèse

3.1 La surexpression de E6 et E7

L'activité transformante d'HPV16 dans des fibroblastes de souris a été décrite pour la première fois en 1986 par (Yasumoto et al., 1986). Cette transformation repose sur la surexpression et les effets des deux oncoprotéines E6 et E7 (zur Hausen, 2000). Plusieurs évènements peuvent être à l'origine de cette surexpression. Dans 30% des cancers cervicaux, la totalité des épisomes viraux s'intègrent au génome cellulaire (Dürst et al., 1985) ce qui entraîne une perte de l'ORF E2 (Schwarz et al., 1985). La perte de E2 lève l'inhibition du promoteur p97 et aboutit à la surexpression de E6 et E7. De plus, l'ARNm bicistronique E6E7 transcrit à partir des génomes intégrés est aussi transcrit avec une 3'UTR de fusion humaine qui augmente la stabilité de l'ARNm (Jeon and Lambert, 1995) (Ehrig et al., 2019). A noter que ces études ne vont pas jusqu'à montrer un effet sur l'expression des oncogènes au niveau protéique.

Si au moins 80% des tumeurs HPV+ présentent de l'intégration, tous sites et cancers confondus (McBride and Warburton, 2017), une part non-négligeable n'en présente pas ou en présente partiellement. La perte de E2 n'est donc pas le seul élément expliquant la surexpression. Pour ces cas, la fonction de répression de E2 peut être empêchée par la méthylation des E2BS qui rend impossible sa fixation et la répression du promoteur p97 (Thain et al., 1996) (McBride and Warburton, 2017). L'apparition de mutations dans la LCR au niveau des E2BS ou d'autre éléments *cis* régulateurs, ou dans l'ORF E2 directement peut aussi mener à la surexpression en rendant E2 non-fonctionnelle (Veress et al., 1999) (Lace et al., 2009). Pour appuyer ces observations, (Häfner et al., 2008) ont montré sur une cohorte de 55 carcinomes que l'état physique du génome HPV16 ne corrèle pas avec le taux de transcrits E6E7. La fonction répressive de E2 peut aussi être dérégulée par son interaction avec E6 (Grm et al., 2005) et E7 (Smal et al., 2009). La charge virale corrèle aussi avec l'expression de E6 et E7 (Hong et al., 2017). Dans la majorité des cas de cancers cervicaux, la surexpression provient vraisemblablement d'un ensemble de ces facteurs qui aboutit à une surexpression des oncogènes viraux pouvant varier d'un facteur 100 d'un échantillon à un autre, comme (Hong et al., 2017) et nos propres données (Paget-Bailly et al., 2019) l'ont montré.

Si la surexpression de E6 et E7 est nécessaire à la transformation, elle est au moins aussi importante à la survie des cellules transformées. En effet, les lignées de cancers

cervicaux SiHa et HeLa entrent en quiescence ou en apoptose lorsque l'on réprime l'expression des oncoprotéines E6 et E7 par l'expression ectopique de E2 (Nishimura et al., 2000), d'ARN interférent (Butz et al., 2003) ou de microARN (Bonetta et al., 2015) dirigés contre E6 et E7.

Il est intéressant de noter que les conditions propices à la surexpression de E6 et E7 sont mises en place indirectement par les protéines E6 et E7 elles-mêmes. En effet, l'instabilité génomique nécessaire à l'intégration des épisomes viraux peut s'expliquer par les effets de E6 sur les voies de réparations de l'ADN, de même que la méthylation des E2BS provient des dérégulations épigénétiques provoquées par E7 (cf. 3.2. L'oncoprotéine E6 d'HPV16 et 3.3. L'oncoprotéine E7 d'HPV16). Il est donc incorrect de parler d'évènement transformant au sens ponctuel, lorsque l'on parle de la carcinogenèse induite par HPV. Il s'agit en fait d'un processus lent pendant lequel les effets et l'expression de E6 et E7 se renforcent au cours du temps jusqu'à un basculement vers un phénotype cancéreux (figure 8).

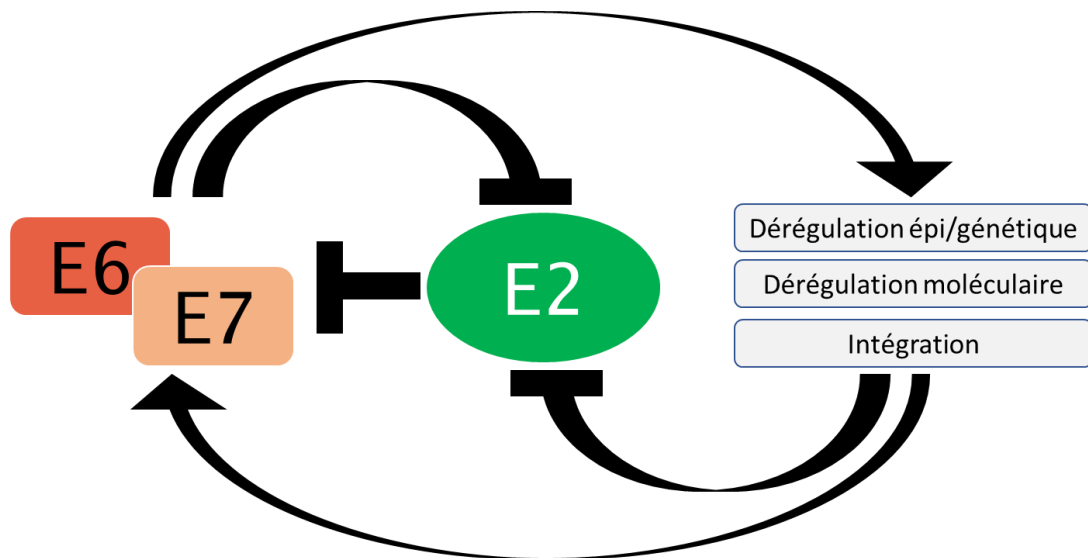


Figure 8 : La surexpression de E6 et E7.

La répression de l'expression de E6 et E7 par E2 va disparaître au cours d'une infection persistante.

Des travaux récents ayant analysé les séquences d'HPV16 retrouvés dans 5000 cas d'infections, de pré-cancers et de cancers cervicaux, révèlent une hypovariation de la séquence des oncoprotéines qui corrèle avec la sévérité de la pathologie. Autrement dit, l'aboutissement d'une infection par HPV16 en cancer n'est possible qu'avec la conservation des séquences, et donc des fonctions, spécifiques des oncoprotéines E6 et E7. Il est intéressant de noter que peu importe le groupe contrôle ou pré/cancer, aucune mutation n'est jamais observée dans le motif LxCxE de E7 nécessaire à son association avec pRb (Mirabello et al., 2017). Par ailleurs, la séquence de E7 possède deux fois moins de mutations que la séquence

de E6 dans le groupe des cancers cervicaux. E6 est tout de même la séquence la moins mutée dans ce groupe après E7. Ces nouvelles données cliniques peuvent sembler contradictoires par rapport aux études moléculaires réalisées sur les deux oncoprotéines ayant montré que E7 est plus tolérante aux mutations que E6 dans la conservation de ses propriétés oncogéniques (Roman and Munger, 2013) (Zanier et al., 2012). En revanche, les propriétés de E7 sont belles et bien plus importantes que celles de E6 dans le processus de transformation cellulaire (Matlashewski et al., 1987) (Vousden et al., 1988) (Phelps et al., 1988) (Bedell et al., 1989).

3.2 L'oncoprotéine E6 d'HPV16

- Structure de l'oncoprotéine E6 d'HPV16

E6 d'HPV16 est une protéine basique composée de 151 a.a. (Figure 9) (Androphy et al., 1987). Sa séquence contient 4 motifs CxxC s'arrangeant en 2 doigts de *Zinc*. Ces deux doigts de *Zinc*, l'un dans la portion N-terminale (ZD1), l'autre C-terminale (ZD2), sont séparés par 35 aa. ZD1 permet d'abord la liaison aux protéines à motif LxxLL. Ce motif riche en Leucine est retrouvé dans diverses protéines cellulaires et notamment des récepteurs nucléaires (d'où son nom alternatif de NR box). Ce motif est également retrouvé dans l'ubiquitine E3 ligase E6AP (Associated Protein) qui est recruté par 16E6 (Chen et al., 1998). La manière dont E6 recrute p53 en revanche n'est pas totalement définie. La résolution de la structure de la partie C-terminale de E6 dans un premier temps (Nominé et al., 2006), puis de la protéine complète plus récemment, montre que E6 pourrait former un homodimère (Zanier et al., 2012). Cette homodimérisation n'est cependant pas indispensable à l'interaction avec p53 (Conrady et al., 2020). ZD2 présente quant à lui une région riche en résidus chargés positivement, lui conférant la capacité de lier les acides nucléiques. E6 reconnaît ainsi les « jonctions ADN 4 voies » (Nominé et al., 2003) et inhibe la réplication virale (Grm et al., 2005). E6 est ainsi retrouvée plus précisément au niveau de la péri-chromatine, haut lieu de transcription, et des granules d'inter-chromatine riches en ribonucléoprotéines. E6 colocalise aussi avec les ribosomes dans le cytoplasme, ce qui semble indiquer que ZD2 permet la liaison aux ARN (Masson et al., 2003). E6 a d'ailleurs été montrée comme interagissant avec son propre ARNm, une caractéristique qui n'est pas partagée avec les HPVbr (Bodaghi et al., 2009). E6 contient également trois séquences NLS (Nuclear Localisation Sequence) (Tao et al., 2003). Sa localisation est majoritairement nucléaire à l'instar des E6 des HPVbr. En revanche on la retrouve aussi dans le cytosol, ce qui pourrait expliquer le plus grand nombre

d'interacteurs de 16E6 par rapport aux protéines E6 des HPVbr (Guccione et al., 2002). Enfin, un motif PBM (PDZ Binding Domain) à l'extrémité C-terminale de séquence ETQL permet la liaison aux protéines cellulaires à domaine PDZ, acronyme de PSD95, Dlg et ZO-1, les premiers membres de cette famille protéique à avoir été identifiés (Ghittoni et al., 2010).

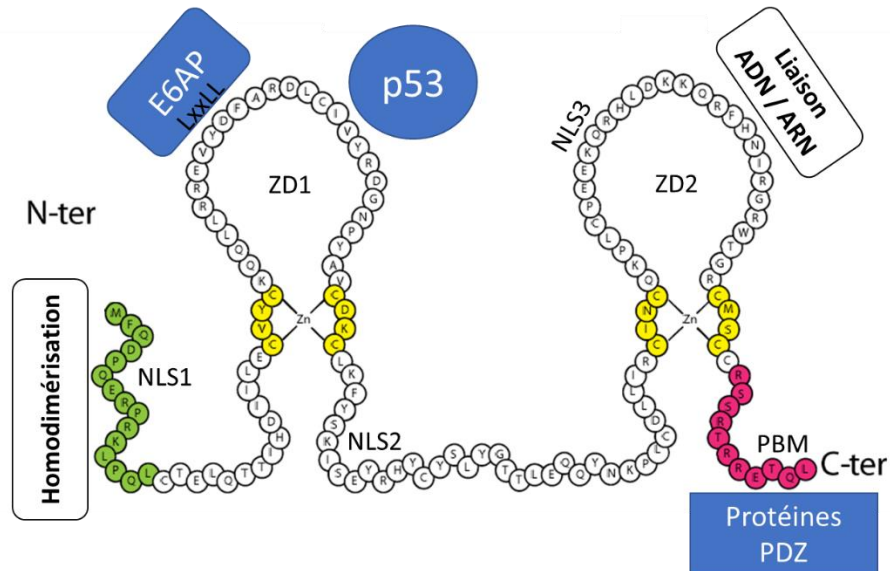


Figure 9 : Structure de la protéine E6 d'HPV16.

Les aa (acides aminés) en vert sont impliqués dans l'homodimérisation de la protéine E6, où se trouve également le NLS1. En jaune sont représentés les aa avec les résidus Cystéine formant les 2 doigts de Zinc, ZD1 et ZD2. En violet sont représentés les aa du PBM de E6. Les protéines schématisées en bleu sont disposées en regard de leur zone d'interaction avec E6. Certains a.a. chargés positivement du ZD2 sont impliqués dans la capacité de E6 à lier les acides nucléiques.

De manière intéressante, aucun des domaines d'interaction de E6 (LxxLL, p53, ADN/ARN et PDZ) ne se « chevauche » permettant potentiellement plusieurs combinaisons d'interactions possibles. N'ayant pas d'activité catalytique, E6 exerce ses fonctions en séquestrant des protéines cellulaires ou en forçant des interactions entre des protéines cellulaires. Enfin, E6 interagit aussi avec d'autres protéines ne contenant ni domaine PDZ, ni motif LxxLL, comme p53. Leur liaison à E6 se fait peut-être *via* des domaines non-identifiés de E6 ou plus probablement de manière indirecte par E6AP ou d'autres protéines associées à E6 (Howie et al., 2009). Cette interaction avec E6AP stabilise aussi la protéine E6 (Tomaić et al., 2009). La résolution récente de la structure de E6 laisse entrevoir une nouvelle ère thérapeutique par l'identification *in silico* et *in vivo* de petites molécules inhibitrices spécifiques de ses interactions (Celegato et al., 2020).

- Dégradation du suppresseur de tumeur p53

La protéine p53 est un suppresseur de tumeur et un facteur de transcription qui régule l'expression d'au moins 300 gènes cellulaires, répartis dans plusieurs voies contrôlant le cycle

cellulaire, la réparation de l'ADN, l'apoptose extrinsèque et intrinsèque, l'autophagie, la traduction et le métabolisme en général. Son activation est principalement induite par les stress cellulaires liés aux dommages à l'ADN, à la réplication, aux protéines mal repliées et à l'hypoxie (Fischer, 2017). Ces stress activent des protéines senseurs qui vont à leur tour activer p53 par des modifications post-traductionnelles. Comme pour toutes les voies de réponse au stress, l'activation de la voie p53 va d'abord tenter de rétablir l'homéostasie cellulaire et si cela s'avère impossible, la mort par apoptose est alors déclenchée.

La dégradation de p53 par E6 est le mécanisme le plus connu et étudié de l'oncogenèse induite par HPV16. Il y a dans un premier temps formation du complexe tripartite E6/E6AP/p53. Il en ressort la poly-ubiquitination de p53 par E6AP et son adressage au protéasome 26S pour dégradation. L'interaction de p53 avec E6 a été démontrée pour la première fois en 1990 (Scheffner et al., 1990). Il a ensuite été montré que E6 recrute l'Ubiquitine E3 ligase E6AP (Huibregtse et al., 1991) et forme le complexe menant à la dégradation de p53 (Scheffner et al., 1993). La simple interaction de E6 avec p53 va d'abord suffire à empêcher son action *trans* activatrice. C'est de cette manière que les HPVbr bloquent l'action de p53 (Lechner and Laimins, 1994). Le pouvoir carcinogène des HPVhr provient notamment de leur capacité exclusive à dégrader p53. Cette affinité de E6 pour p53 va aussi entraîner la relocalisation de p53 au cytoplasme, l'interaction masquant son NLS (Mantovani and Banks, 1999).

Mais E6 dérégule également les senseurs aux stress et activateurs de p53 en amont, ainsi que les effecteurs chargés de rétablir l'homéostasie ou d'induire la mort en aval. E6 va interagir avec CBP/p300 afin d'empêcher l'acétylation de p53 qui augmente normalement son affinité pour ses gènes cibles. CBP/p300 est normalement activée par les dommages à l'ADN, c'est une des premières causes d'accumulation de mutations génétiques (Barlev et al., 2001). De manière similaire, E6 entraîne la dégradation de hADA3, une autre acétyltransférase chargée d'activer p53 (Kumar et al., 2002). hADA3 est aussi lié à l'inhibition de p14/ARF. ARF s'associe normalement avec MDM2, une ubiquitine ligase et principal répresseur de p53, pour empêcher sa dégradation (Shamanin et al., 2008). L'interaction de E6 avec ces acétyltransférases, ainsi qu'avec Tip600 et les histones méthyltransférases CARM1, SET7 et PRMT1 (Wallace and Galloway, 2015) également remodelateurs de la chromatine vont provoquer des changements drastiques dans le transcriptome cellulaire qui ne se limitent pas aux gènes régulés par p53.

Les effets de la perte de p53 sont détaillés ci-après avec les autres effets de E6 et sont regroupés en fonction des propriétés majeures des cellules tumorales (Hanahan and Weinberg, 2011). L'échappement au système immunitaire induit par E6 (et E7) a déjà été traité (cf. 2.3. Persistance).

- Accumulation de mutations génétiques

Privée de p53, carrefour de nombreuses voies protégeant l'intégrité cellulaire, la cellule est libre de survivre en accumulant les mutations génétiques nécessaires à sa transformation complète. Le stress induit par les oncoprotéines s'ajoute aux quelques 200000 lésions génétiques auxquelles une cellule doit faire face chaque jour (Atamna et al., 2000). E6 participe aussi à l'accumulation de nouvelles mutations génétiques en dégradant XRCC1 qui répare normalement les cassures simple brin (Iftner et al., 2002), MGMT qui restaure les guanines alkylées (Srivenugopal and Ali-Osman, 2002), et en relocalisant BRCA1 au cytoplasme qui répare les cassures double brin (Zhang et al., 2005). E6 va aussi favoriser la dégradation de MCM7 (Kühne and Banks, 1998) impliqué dans le contrôle de la réplication cellulaire et de la stabilité chromosomique. En conséquence, il ressort un ensemble de mutations somatiques formant la signature des cancers cervicaux. A noter que l'intégration des épisomes viraux peut aussi se faire au sein d'un gène cellulaire s'ajoutant à la charge mutationnelle provoquée par E6 (Dürst et al., 1987) (Bodelon et al., 2016). Les données du TCGA (The Cancer Genome Atlas) (Burk et al., 2017) montre le gène *PIK3CA* comme muté dans 40% des cas, *EGFR* (30%), *SMAD4* (25%), *EP300* (p300, 20%), *FBXW7* (20%) et *PTEN* (15%). On observe également des duplications de pro-oncogènes comme *TP63* (75%), *PIK3CA* (50%) et *TGFBR2* (30%), ainsi que des délétions de suppresseurs de tumeurs comme *FOXP1* (40%) et *PTEN* (30%) (Litwin et al., 2017). On constate évidemment très peu de mutations dans *TP53* et *RBI* compte tenu des effets respectifs de E6 et E7 sur p53 et pRb. On retrouve ainsi fréquemment des aberrations cytogénétiques augmentant avec la sévérité des lésions et au stade cancer, le gain du bras 3q (60%), 1q (40%), 20q (30%), 1p (25%) et la perte du bras 3p (40%), 2q (35%) et 4p (35%) (Shulzhenko et al., 2014). Ces réarrangements globaux vont aussi participer au phénotype tumoral.

- Inhibition de l'apoptose

Outre la dégradation de p53, E6 va rendre inopérant la plupart des points clés de l'apoptose. En effet, E6 va inhiber l'apoptose intrinsèque en favorisant la dégradation de Bak (Underbrink et al., 2008). Suite à un stress cellulaire, Bak forme un hétérodimère avec Bax et

s'insère dans la membrane mitochondriale pour provoquer le relargage du cytochrome c et de facteurs induisant l'apoptose. Cela mène à la formation de l'apoptosome, à l'activation de la Caspase 9 puis des caspases 3 et 7 entraînant à leur tour la mort cellulaire (Leverrier et al., 2007). L'action de E6 sur d'autres protéines pro- ou anti-apoptotiques que Bak est plus discutée (Howie et al., 2009). E6 bloque également l'apoptose extrinsèque en empêchant la transduction du signal par les récepteurs de mort. Les deux voies Fas et TRAIL sont bloquées en interagissant avec TNFR-1 (Filippova et al., 2002) et l'adaptateur FADD (Filippova et al., 2004). Plus en aval, E6 inhibe la Caspase 8 (Filippova et al., 2007). Se faisant, E6 empêche une communication entre les deux voies de l'apoptose car cette Caspase peut activer tBid qui stimule la voie intrinsèque. Pour finir, E6 entraîne la surexpression de cIAP-2 qui inhibe ces deux voies (James et al., 2006a).

- Immortalisation cellulaire

Une des caractéristiques essentielles que doit acquérir une cellule pro tumorale est la capacité de proliférer indéfiniment. Sans cette immortalisation, les cellules tumorales hyperprolifératives finiraient par mourir par sénescence répllicative. En effet, 200pb sont perdues aux extrémités des chromosomes à chaque division cellulaire. Chaque cellule somatique a donc un nombre de divisions fini appelée « limite de Hayflick » avant de perdre de l'information et son intégrité. Seules les cellules souches et les cellules cancéreuses sont capables d'inverser cette horloge biologique par l'activation du complexe ribonucléoprotéique de la télomérase. Ce complexe est composé de la protéine hTERT à activité rétrotranscriptase, et de TERC l'ARNnc (ARN non codant) servant de matrice à hTERT pour reformer les extrémités télomériques (Shay and Wright, 2019). E6 permet l'immortalisation cellulaire par la surexpression de hTERT. *Via* E6AP, E6 va recruter des HAT (histones acétyltransférases) pour décompacter la chromatine au niveau de *hTERT* et faciliter sa transcription (James et al., 2006b). E6 interagit ensuite avec les complexes protéiques régulant la transcription du gène. L'hétérodimère *trans* activateur c-Myc/Max ainsi que le FT Sp1 (Oh et al., 2001) sont recrutés à la place des répresseurs USF1/2 (McMurray and McCance, 2003). A noter que p53 est aussi un répresseur de hTERT (Lin and Elledge, 2003). E6 agit également sur les deux isoformes du FT NFX1. La première NFX1-91, répresseur transcriptionnel de hTERT, est dégradée par E6/E6AP (Xu et al., 2008), la deuxième NFX1-123, activateur transcriptionnel, est recrutée par E6. NFX1-123 lie aussi l'ARNm hTERT pour augmenter sa stabilité et sa traduction (Katzenellenbogen et al., 2007). Enfin, E6 augmente l'activité du complexe de la

téломérase (Liu et al., 2009). Ainsi, la surexpression et l'activation de la téломérase par E6 confèrent la caractéristique clés d'immortalisation aux cellules tumorales.

- Prolifération et invasion

Les protéines à domaines PDZ participent à l'assemblage de complexes protéiques aux membranes et sont impliquées dans la transduction de signaux de survie, dans la polarité cellulaire, les jonctions d'ancrage intercellulaires et l'organisation du cytosquelette (Hung and Sheng, 2002). Le PBM de E6 est indispensable à la réalisation du cycle viral responsable en partie de son pouvoir transformant, au moins autant que sa capacité à dégrader p53 (Yoshimatsu et al., 2017). Avec son PBM de classe I, E6 désorganise ces complexes et dérégule les signaux de survie indépendamment de l'ancrage, les signaux de croissance et la polarité cellulaire, favorisant ainsi la prolifération et l'invasion (Wallace and Galloway, 2015). Le mécanisme d'action par lequel E6 perturbe les voies dont les protéines PDZ font parties n'est pas toujours connu, leur dégradation par le protéasome ou du moins l'implication de E6AP n'étant pas toujours assurée (Lee and Laimins, 2004) (Grm and Banks, 2004). E6 va principalement déréguler les protéines de la superfamille MAGUK (Membrane-Associated Guanylate Kinases). A savoir DLG1 (Kiyono et al., 1997), SCRIB (Nakagawa and Huibregtse, 2000), MAGI-1 (Glaunsinger et al., 2000), 2 et 3 (Thomas et al., 2002), mais aussi MUPP1 (Lee et al., 2000), PATJ (Storrs and Silverstein, 2007), PTPN3 (Jing et al., 2007) et PTPN13 (Spanos et al., 2008). DLG et SCRIB sont des suppresseurs de tumeurs impliqués dans les jonctions serrées et l'attachement des cellules à la lame basale, tout comme MAGI et MUPP1. MAGI-1 est lié à la B-Caténine, MAGI-2 et 3 sont quant à elles nécessaires à l'activation de PTEN (Thomas et al., 2001), un suppresseur de tumeur majeur. La dégradation de PTPN3 et 13 réduit aussi les besoins en facteurs de croissance et favorise la prolifération indépendante de l'ancrage. De nouvelles données d'interactomique obtenues par GPCA (Gaussian Princess Luciferase Protein Complementation Assay) indiquent que E6 d'HPV16 et 18 interagissent avec au moins 50 des 266 protéines à domaine PDZ humaines (Vincentelli et al., 2015) (Duhoo et al., 2019). Le PBM de E6 est soumis à une modification post-traductionnelle par phosphorylation de la tyrosine **ETQL**. Cette phosphorylation inhibe l'interaction de E6 avec les protéines à domaine PDZ. En revanche, elle confère la capacité au PBM d'interagir avec la protéine 14-3-3 ξ qui reconnaît les résidus phosphorylés. Cela pourrait témoigner d'un autre rôle du PBM E6 dans la dérégulation de l'homéostasie cellulaire (Boon et al., 2015).

E6 favorise aussi la dégradation d'autres protéines impliquées dans ces voies mais sans domaine PDZ. La dégradation de E6TP1 (E6 Targeted Protein 1) (Singh et al., 2003) a été montrée comme activant la voie Ras *via* Rap1 (Ras associated protein) favorisant l'invasion des cellules HPV+ (Zhang et al., 2017) (Spanos et al., 2008). E6 favorise également la dégradation d'autres protéines G comme la tubérine, activant la voie mTOR (Zheng et al., 2008). En complément, E6 augmente l'expression de SLUG et TWIST et inhibe l'expression de la cadhérine-E (Jung et al., 2013). Toutes ces caractéristiques sont liées à la transition épithélio-mésenchymateuse et confèrent aux cellules tumorales un pouvoir d'invasion et de progression métastatique. La E-cadhérine est aussi nécessaire à la reconnaissance d'une cellule infectée par les cellules de Langherans, sa répression pourrait donc également intervenir dans l'échappement au système immunitaire (Ghittoni et al., 2010). E6 augmente aussi l'expression du VEGF de manière p53-indépendante et participe ainsi à l'angiogenèse des tumeurs HPV+ (López-Ocejo et al., 2000).

- Dérégulation des ARN non-codants

Au cours des 20 dernières années, l'étude des ARN non-codants a pris une place prépondérante dans la biologie moléculaire en général et dans la physiopathologie des cancers (Slack and Chinnaiyan, 2019). En ce qui concerne les cancers induits par HPV, la recherche s'est d'abord focalisée sur l'étude des micro-ARN (miR). Il est maintenant admis qu'une fraction des miR régule l'inflammation et l'immunité antivirale innée, ou cible directement les transcrits ou ARNg (ARN génomique) viraux (Kumar, 2007) (Umbach and Cullen, 2009) (Nahand et al., 2020). Il a d'ailleurs été montré dans notre laboratoire que le miR-375 cible et déstabilise les transcrits E6/E7 d'HPV16 (Morel et al., 2017). HPV16 dérégule évidemment le métabolisme des miR. Premièrement, en dérégulant l'expression des prépromiR *via* les dérégulations transcriptionnelles engendrées par E6 et E7 (Harden et al., 2017) et en dérégulant la biogenèse des miR *via* la dérégulation de DROSHA et DICER (Harden and Munger, 2017), toutes deux impliquées dans la biogenèse des miR. Encore plus récemment, nous voyons une explosion d'intérêt pour les lncRNA (long non-coding RNA). A noter que leur caractère non codant strict est remis en cause (Wu et al., 2020). La faible conservation de leur séquence au sein du vivant semble indiquer qu'ils n'exercent pas leur fonction principale de manière séquences spécifiques mais plutôt en formant des structures secondaires complexes. Ces structures vont servir de plateforme à l'assemblage de complexes multiprotéiques ou d'interface entre chromatine et protéines régulatrices épigénétiques. Ils servent également « d'éponges à miR » et stabilisent leurs ARNm cibles (Yang et al., 2014).

Les premiers à avoir été identifiés dans les années 90 sont XIST et H19. On estime aujourd'hui leur nombre à plusieurs dizaines de milliers. Plusieurs de ces long ARN non-codants sont dérégulés dans les cancers cervicaux et sont en lien avec la prolifération, la migration et l'invasion, mais aussi avec l'apoptose et la stabilité du génome (pour une revue voir (Sharma and Munger, 2020)).

3.3 L'oncoprotéine E7 d'HPV16

- Structure de l'oncoprotéine E7

E7 d'HPV16 est une protéine sans activité catalytique comme E6. E7 est constituée de 98 aa. Sa séquence est subdivisée en 3 régions conservées CR1, CR2 et CR3 (Conserved Region). CR1 et CR2 présentent respectivement une homologie partielle et complète avec les domaines de la protéine E1a de l'adénovirus et avec l'antigène grand T du SV40, deux autres petits virus à ADN transformants (Vousden and Jat, 1989). La capacité à lier pRb assurée par le motif LxCxE du domaine CR2 est indispensable à la biologie de ces virus. Le domaine CR3 est quant à lui constitué de deux motifs CxxC formant un doigt de zinc. Ce doigt de zinc permet notamment la liaison à la culline 2 impliquée dans la dégradation de pRb (Munger et al., 2001), et permet aussi l'oligomérisation de E7 (McIntyre et al., 1993). L'oligomère de E7 a un rôle de chaperonne et pourrait expliquer pourquoi E7 peut interagir avec un vaste panel de protéines cellulaires (Alonso et al., 2006). E7 existe cependant principalement sous forme d'homodimère (Clements et al., 2000). Le domaine CR3 interagit aussi avec pRb et est nécessaire à la libération des FT E2F normalement séquestrés par pRb (Liu et al., 2006). Il existe un site de phosphorylation en aval du motif LxCxE, nécessaire aux capacités transformantes de E7 mais qui ne gêne pas sa liaison à pRb (Barbosa et al., 1990). Un autre site C-terminal de phosphorylation est aussi impliqué dans la régulation de la protéine et pourrait moduler l'activité de E7 en fonction des étapes du cycle cellulaire (Massimi and Banks, 2000). E7 ne semble pas capable d'interagir avec les acides nucléiques. La localisation subcellulaire de E7 est soumise à controverse en fonction du modèle cellulaire ou de la technique de détection utilisée (Roman and Munger, 2013). La détection par fractionnement cellulaire la premièrement montrée comme cytoplasmique (Smotkin and Wettstein, 1987). Des travaux plus récents l'ont localisé majoritairement dans le noyau (Guccione et al., 2002). E7 possède une NES (Nuclear Export Sequence) et au moins une NLS lui permettant son transport entre noyau et cytoplasme, exerçant dans les deux compartiments son éventail de fonctions transformantes (Knapp et al., 2009). E7 présente une demi-vie d'une à deux heures

(Smotkin and Wettstein, 1987) et est dégradée par le protéasome *via* son ubiquitinylation et adressage au protéasome (Reinstein et al., 2000).

- Dégradation du suppresseur de tumeur pRb

E7 est à l'origine de l'activation du cycle cellulaire dans les cellules infectées. Cette activation repose sur la dégradation de pRb et de ces homologues p107 et p130 (Münger et al., 1989) (Boyer et al., 1996). La dégradation de pRb fait intervenir la culline 2, une ubiquitine ligase recrutée par E7 qui va poly-ubiquitinyler pRb et provoquer son adressage au protéasome (Huh et al., 2007) (figure 10). Il est également possible que pRb soit adressé au protéasome de manière directe car E7 peut interagir avec une de ses sous-unités (Berezutskaya and Bagchi, 1997). En absence de division cellulaire, pRb séquestre les FT de la famille E2F. Lors de l'entrée en cycle, la phosphorylation de pRb par les cycline/CDK4/6 entraîne la libération des facteurs E2F. E2F1, 2 et 3 sont des activateurs transcriptionnels indispensables à la phase S à travers l'activation de l'expression de facteurs limitants comme l'ADN polymérase α , la Déhydrofolate Réductase et les cyclines E et A (Dyson, 1998). A l'inverse, p107 et p130 ont une activité de répresseur transcriptionnel quand ils sont liés à E2F4 et E2F5, respectivement. E7 entraîne donc la libération des FT E2F1, 2 et 3 et la dégradation de p107 et p130 empêche leur activité de répresseur (Ghittoni et al., 2010). E7 interagit aussi directement avec E2F1 pour augmenter son activité transcriptionnelle et avec E2F6 à l'inverse pour bloquer son activité de répresseur (McLaughlin-Drubin and Münger, 2009). Le cycle cellulaire est normalement inhibé par l'expression dépendante de p53 des CKI (CDK Inhibitors) p21^{CIP1} (Funk et al., 1997) et p27^{KIP1} (Zerfass-Thome et al., 1996) qui exercent un rétrocontrôle négatif sur le cycle ou qui l'inhibe définitivement dans les cellules différenciées. E7 interagit avec ces deux inhibiteurs pour bloquer leur activité de répresseur. E7 permet ainsi le découplage entre différenciation et division cellulaire indispensable à l'amplification des épisomes viraux dans les cellules infectées et permet ensuite la prolifération des cellules transformées.

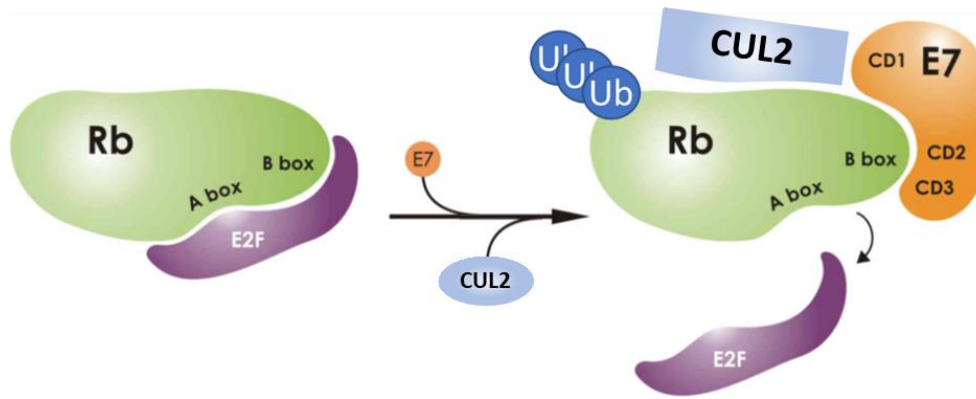


Figure 10 : Dégradation de pRb et libération des facteurs E2F.

Les facteurs E2F se lient à pRb *via* ses domaines A et B box. E7 va interagir avec ces domaines *via* CD2 (CR2) et le motif LxCxE et *via* CD3 (CR3) en C-ter de pRb. L'interaction déplace et libère les facteurs E2F de leur complexe avec pRb. L'interaction de E7 avec la culline 2 *via* CD1 (CR1) va permettre la poly-ubiquitinylation de pRb et sa dégradation par le protéasome. Adapté de (Pappa et al., 2018).

- Inhibition de l'apoptose

E7 joue un rôle limité dans l'inhibition de la mort cellulaire. La prolifération cellulaire engendrée par E7 entre en contradiction avec l'absence de facteurs de croissance et/ou la présence de facteurs inhibiteurs. Cette contradiction déclenche la « réponse trophique sentinelle », une forme de mort cellulaire programmée dépendante de p53 qui fait intervenir l'autophagie mais pas ses gènes cibles ni les caspases (McLaughlin-Drubin and Münger, 2009). Cela entre aussi en contradiction avec l'absence de signaux d'ancrages à la MEC pouvant déclencher une autre forme de mort, l'anoïkis. E7 bloque ce type de mort en interagissant avec l'ubiquitine ligase p600 (Huh et al., 2005). De plus, les caractéristiques décrites pour E6 vont aussi contrecarrer les effets pro-apoptotiques de E7 dans les cellules transformées les coexprimant.

- Dérégulation de l'homéostasie cellulaire

L'expression de E7 dérégule aussi l'homéostasie cellulaire de manière plus générale. En effet, son expression corrèle avec un changement dans le métabolisme glucidique. Aucune dérégulation directe de ce métabolisme par E7 n'a été établie et il se pourrait que ces changements ne soient qu'une adaptation cellulaire aux forts besoins énergétiques nécessaires à l'hyperprolifération induite par E7 (Münger et al., 2001). En revanche, E7 induit directement une alcalinisation du cytosol. Cette caractéristique est souvent retrouvée dans les cellules transformées. Elle est d'autant plus importante que des cellules exprimant E7, mais dans lesquelles cette alcalinisation est inhibée, ne présentent plus les caractéristiques transformées associées à l'expression de E7 (Reshkin et al., 2000).

- Accumulation de mutations génétiques

E7 participe aussi à l'accumulation d'anomalies génétiques et cytogénétiques. E7 va principalement contribuer à l'apparition d'anomalies chromosomiques en perturbant le processus de mitose au niveau de la biogenèse des centrioles. La duplication des centrioles semble faire intervenir l'activation aberrante de CDK2 (Duensing et al., 2006) de manière indépendante de pRb (Duensing and Münger, 2003). Contrairement à E6, l'instabilité génétique causée par E7 ne provient pas d'une dérégulation directe des mécanismes de réparation de l'ADN. Néanmoins, son rôle pourrait être prédominant dans la carcinogenèse car ces anomalies sont constatées au niveau cytologique (mitoses multipolaires) à un stade très précoce d'infection et même avant intégration du génome viral au génome cellulaire (Duensing et al., 2001). On peut émettre l'hypothèse que l'intégration des épisomes viraux se ferait plus facilement dans ces cellules au génome fragilisé. De plus, l'expression de E6 et E7 de l'HPV16 n'entraîne pas ces dérégulations centrosomiques (Münger et al., 2001). Cette instabilité génomique pourrait donc bel et bien être une étape clé précoce de la carcinogenèse induite par HPV16. Si ces anomalies causées par E7 sont précoces, leur nombre augmente aussi grandement avec l'intégration des épisomes et la surexpression des oncoprotéines (Pett et al., 2004).

- Dérégulations épigénétiques

E7 va pour finir avoir un impact global sur les marques épigénétiques et les remodelleurs de la chromatine. L'épigénétique représente un domaine d'étude toujours plus prépondérant dans la biologie des cancers. E7, comme E6, va ainsi permettre de modifier le phénotype cellulaire sans mutation somatique. Ces effets résultent principalement de l'activation par E7 de KDM6A et B (McLaughlin-Drubin et al., 2011) et de l'inactivation de PRC1 (polycomb Repressive Complex) *via* BMI1 et PRC2 *via* EZH2 (Hyland et al., 2011). De plus, E7 est capable d'interagir avec des HDAC et avec des HAT comme CBP/p300, pCAF et SRC1. Ces interactions sont étroitement liées aux FT de la famille E2F (McLaughlin-Drubin and Münger, 2009). Cela aboutit à une dérégulation de la transcription de nombreux gènes cellulaires, activant par exemple l'expression de HIF1 α pour favoriser l'angiogenèse (Bodily et al., 2011), ou inhibant l'expression de TLR9 pour favoriser l'échappement immunitaire (Hasan et al., 2013). Pour finir, E7 entraîne une surexpression des DNMT (DNA MéthylTransférase) qui se traduit par une hyperméthylation des promoteurs de

gènes suppresseurs de tumeurs en particulier (Burgers et al., 2007) (Groves and Coleman, 2015).

- Synergie des oncoprotéines E6 et E7

La coexpression des deux oncoprotéines va donner lieu à des effets croisés. E6 et E7 dégradent p53 et pRb, respectivement. Cependant, E6 affecte directement pRb et E7 affecte p53. Dans un premier temps, E6 entraîne la phosphorylation de pRb et favorise le passage de la phase G1 à S. Pour se faire, E6 est capable d'augmenter l'activité kinase du complexe Cycline A/CDK2 mais aussi d'inhiber p21, répresseur de ce complexe. De manière intéressante, cette capacité est indépendante de la dégradation de p53, la phosphorylation de pRb et l'activation de gènes dépendants des FT E2F étant toujours induite par un mutant E6 incapable d'interagir avec p53 (Malanchi et al., 2004) (Shai et al., 2007) (Parroche et al., 2011). Dans un second temps et de manière tout aussi intrigante, E7 entraîne une stabilisation de p53 en empêchant sa dégradation par son répresseur physiologique, l'Ub ligase MDM2 (Seavey et al., 1999). p53 et pRb étant deux des plus importants suppresseurs de tumeurs, il est envisageable que la perte de fonction de pRb entraîne *de facto* l'activation de p53 dans le but de préserver un minimum d'homéostasie cellulaire. En revanche, cette stabilisation de p53 ne se traduit pas par son activation. Au contraire, l'expression de E7 entraîne une baisse de la transactivation des gènes cibles de p53 (Eichten et al., 2002). La seule explication proposée repose sur l'interaction de E7 avec TBP (TATA Binding Protein) (Phillips and Vousden, 1997), pour former un complexe E7/TBP/p53 qui réprimerait la transcription des gènes cibles de p53 (Massimi and Banks, 1997). Les dégradations de p53 et de pRb induites par E6 et E7 participent aussi à l'immortalisation cellulaire *via* la perte des points de contrôles du cycle en G1/S et G2/M. De manière intéressante, il a été montré que l'abolition de ces deux checkpoints est plus indispensable pour contourner la sénescence répllicative que l'activation de la télomérase (Helt and Galloway, 2001) (McMurray and McCance, 2004). De plus, il semble que E7 joue aussi un rôle dans l'activation de la télomérase (Stöppler et al., 1997).

Les propriétés des deux oncoprotéines qui servent à la base à la réalisation du cycle viral, servent également au développement du phénotype transformé. La surexpression de E6 et E7 va en effet renforcer l'échappement au système immunitaire, la dédifférenciation, l'immortalisation, la prolifération, la survie que ce soit en réprimant l'apoptose ou par les facteurs de croissance, et le métabolisme. Une notion paradoxale que l'on peut souvent lire

dans la littérature HPV est l'incompatibilité entre cycle viral et cancer. En effet, le dernier empêche ou supprime la réalisation du premier. Comme souvent face à ce genre de paradoxe, on réfléchit en personnifiant la problématique : « pourquoi le virus a évolué en génotype HR si ses nouvelles caractéristiques suppriment à terme sa fonction première, c.a.d. se multiplier et perdurer ». Au contraire, E6 et E7 des HPVhr et d'HPV16 notamment, sont optimisées pour réaliser l'infection la plus longue et la plus productive possible. Peu importe que l'infection se termine en cancer dans moins de 1% des cas, les propriétés d'HPV16 auront déjà largement permis sa propagation.

3.4 Epidémiologie des cancers HPV

L'infection par HPV est la première infection sexuellement transmissible avec 300 millions de femmes infectées dans le monde, ou plus exactement positive à l'ADN HPV, dont un tiers par HPV16, ou par 18 (de Sanjosé et al., 2007).

Les HPVhr à eux seuls sont à l'origine de 5% de tous les cancers dans le monde. Les données épidémiologiques mondiales du GLOBOCAN 2012 (Global Cancer Observatory) reprisent en figure 11, indiquent que les HPVhr sont responsables de 100% des 530 000 cas de cancers du col de l'utérus. Il sont aussi responsables de 88% des cancers de l'anus (35 000 cas), de 41% des cancers du vagin et de la vulve (20 000 cas), 50% des cancers du pénis (13 000 cas), 31% des cancers de l'oropharynx (29 000 cas) et 2,4% des cancers de la cavité buccale et du pharynx (9 000 cas) (de Martel et al., 2018).

L'incidence de ces cancers est fortement variable en fonction de l'Index de Développement Humain (IDH). En effet, les cancers associés aux HPV représentent moins de 3% des cancers dans les pays les plus développés. En revanche, ils peuvent causer jusqu'à 20% des cancers chez les femmes des pays Sub-Sahariens par exemple. Le sexe est également un critère d'hétérogénéité car les HPVhr ne causent que 0,8% des cancers masculins contre presque 9% des cancers féminins. Ce qui s'explique évidemment par la surreprésentation des cancers du col de l'utérus (Arbyn et al., 2020).

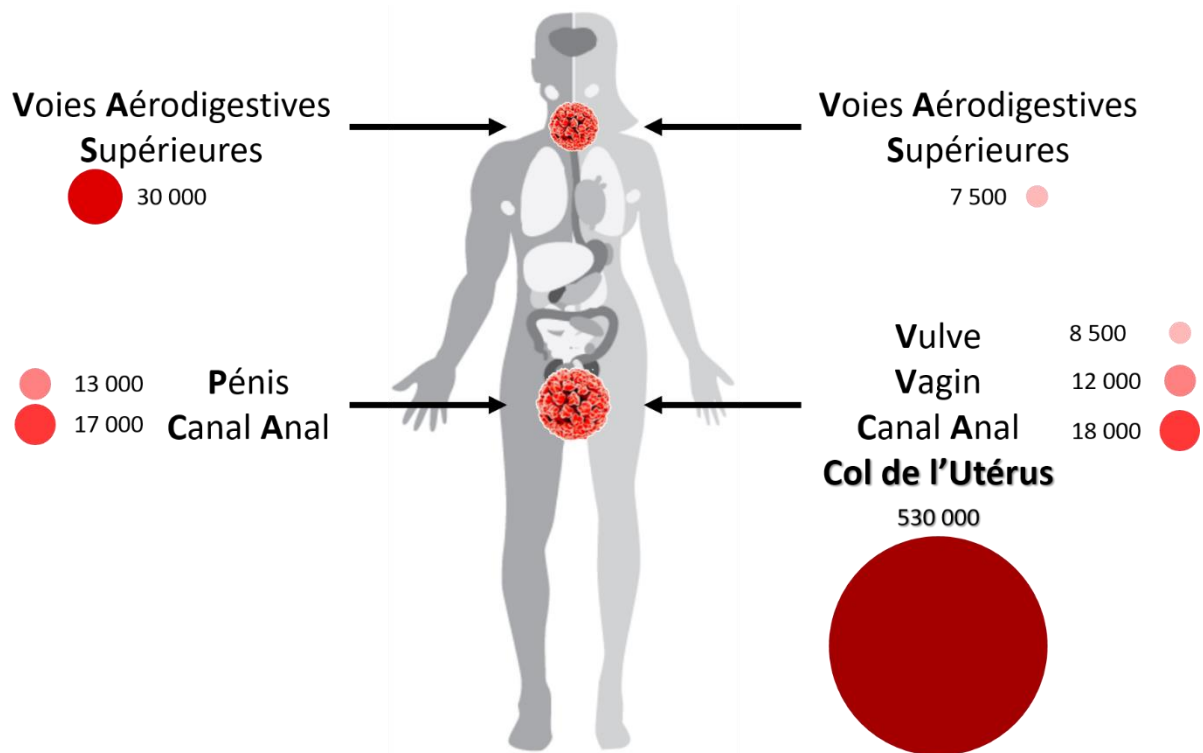


Figure 11 : Incidence des cancers causés par les HPV.

Les données sont tirées du GLOBOCAN 2012 (de Martel et al., 2018) et représentent le nombre de cas de cancers imputables aux HPV par site anatomique et par an. La partie gauche représente les cas de cancers chez l'homme et la partie droite chez la femme.

De 530 000 en 2012, les cancers du col sont passés à 570 000 cas en 2018 dans le monde. Cette augmentation s'explique par le boom démographique toujours en cours dans les pays en voie de développement mais aussi par l'augmentation de leur dépistage. Toujours en 2018, on enregistre 311 000 décès dus au cancer du col, faisant de cette pathologie la 4^{ème} cause de décès chez les femmes dans le monde, avec un taux de mortalité supérieur à 50% (Bray et al., 2018). A noter que certains faits épidémiologiques restent intrigants à ce jour, comme la surreprésentation des cancers des VADS induits par HPV chez les hommes des pays les plus développés par rapport aux femmes et au reste du monde (de Martel et al., 2018).

Le cancer du col de l'utérus est de bon pronostic avec un taux de survie globale à 5 ans de 66%, et les cancers des VADS HPV+ sont de meilleur pronostic que leur contrepartie HPV-. Le traitement des cancers du col consiste en premier lieu en l'ablation de la tumeur ou de la zone cancéreuse par chirurgie. Ces ablations peuvent être plus ou moins extensives en fonction du stade d'avancement du cancer, de la conisation du col (le col uniquement) pour les stades précoces, en passant par une hystérectomie radicale (utérus, col, partie haute du vagin et nœuds lymphatiques pelviens) et, jusqu'à une exentération pelvienne (ablation en

plus des organes périphériques : ovaires, vagin en totalité et possiblement vulve, vessie et rectum). Ces cancers sont également traités par chimiothérapie aux sels de platines et paclitaxel, récemment additionnés de l'anticorps monoclonal anti-VEGF bevacizumab (Koh et al., 2019). A la chimiothérapie s'ajoute un traitement par radiothérapie ou par curiethérapie. La radiothérapie tue les cellules cancéreuses par irradiations externes *via* un faisceau ionisant, là où la curiethérapie consiste à introduire la source radiative directement dans la tumeur, ce qui préserve les organes sains périphériques. Ce dernier traitement est facilité par la localisation du cancer cervical facilement accessible par voie vaginale (Small et al., 2017). Ces traitements pourraient s'étoffer dans les années à venir avec le succès grandissant des thérapies ciblées. D'autres traitements par inhibiteurs chimiques sont aussi envisagés dans la littérature comme la mélatonine qui a montré une certaine efficacité dans de nombreux cancers en complément des chimiothérapies classiques (Shafabakhsh et al., 2019). De nombreuses études cliniques sont également en cours pour développer un vaccin thérapeutique contre les infections par HPVhr de tout stade en combinaison avec des agents immunomodulateurs ou non (Smalley Rumfield et al., 2020). Pour finir, des phases d'essais cliniques sont en cours au CHRU de Besançon (étude VOLATIL) afin de tester l'administration d'un vaccin thérapeutique anti-TERT. Ce vaccin vise à développer une réponse immunitaire contre deux peptides dénommés UCP (Universal Cancer Peptides) dérivés de la télomérase et présentés par les cellules tumorales. L'effet prometteur pour tous les cancers HPV consiste à induire une forte expansion lymphocytaire T anti-TERT et l'activation de ces cellules contre les tumeurs en l'associant à une immunothérapie anti-PD1/PD-L1 (Dosset et al., 2020).

Une vaccination préventive existe également et représente une immense opportunité dans la lutte contre l'apparition des cancers HPV, d'autant plus pour les pays en voie de développement où la mortalité liée à HPV est la plus élevée et où les systèmes de santé n'ont pas les moyens de mettre en place un dépistage systématique et organisé. Il existe à l'heure actuelle trois vaccins prophylactiques permettant de développer une réponse immunitaire dirigée contre la protéine de capsid L1 des génotypes 16 et 18. On estime que le plus récent des vaccins, le nonavalent Gardasil-9™ (Merck) protégera contre 90% des cancers induits par HPV causés par les génotypes 16, 18, 31, 33, 45, 52 et 58, et 90% des verrues génitales causées par HPV 6 et 11 principalement (de Martel et al., 2018) (Riethmuller et al., 2015). Le Gardasil 9 a été ajouté au calendrier vaccinal en 2018 avec une recommandation par l'HAS (Haute Autorité de Santé) d'élargir la vaccination aux garçons de 11 à 14 ans (Blin, 2020). La

prévention secondaire par dépistage des lésions précancéreuses du col reste indispensable pour accumuler un recul nécessaire sur les stratégies vaccinales, pour détecter les cancers induits par d'autres génotypes, et pour suivre les femmes plus âgées ne rentrant pas dans les critères de vaccination.

En conclusion, la compréhension de la biologie des HPV au cours des 50 dernières années a permis de développer dans un premier temps des outils de dépistage, de traitement puis plus récemment de vaccination qui laissent entrevoir la disparition de tous les cancers induits par les HPV, soit 700000 cancers pour 350000 décès annuellement, avant même la fin du XXIème siècle. Malheureusement, la couverture vaccinale est déplorable dans de nombreux pays et notamment en France. La réussite des politiques de santé publique ou un changement de l'opinion publique défavorable à la vaccination ne relèvent pas du domaine scientifique. Ce qui en relève en revanche c'est d'apporter une meilleure compréhension des mécanismes qui sous-tendent la carcinogenèse induite par les HPV, l'expression des oncoprotéines virales et la régulation de leurs ORF.

4 Expression du génome HPV16 : cycle viral et cancer

4.1 Régulation transcriptionnelle des gènes d'HPV16

Certains mécanismes de la régulation transcriptionnelle ont déjà été abordés dans le déroulement du cycle viral et la surexpression de E6 et E7, notamment la transrégulation de E2 sur la LCR.

La transcription à partir du promoteur précoce p97 est régulée de la même façon que les gènes cellulaires de classe II. La transcription est assurée par l'ARN polymérase II dont le complexe de pré initiation contenant TBP (TATA Binding Protein) reconnaît normalement la boîte TATA (-25 nt) en amont du premier codon transcrit. Plus en amont encore on trouve un site de fixation pour le FT Sp1, nécessaire à l'activation du promoteur. C'est entre ces deux séquences (Sp1 et TATA) que se trouvent les E2BS#1 et #2 qui vont réprimer la transcription *via* la fixation de E2 (Tan et al., 1994a). E2 est le plus important régulateur du promoteur précoce (Bernard, 2002). Le produit d'épissage viral E8^E2c a également été montré comme un répresseur de la transcription virale, la partie E8 recrutant des remodeleurs de la chromatine réprimant la transcription, et la partie E2 se fixant à la LCR (Lace et al., 2008).

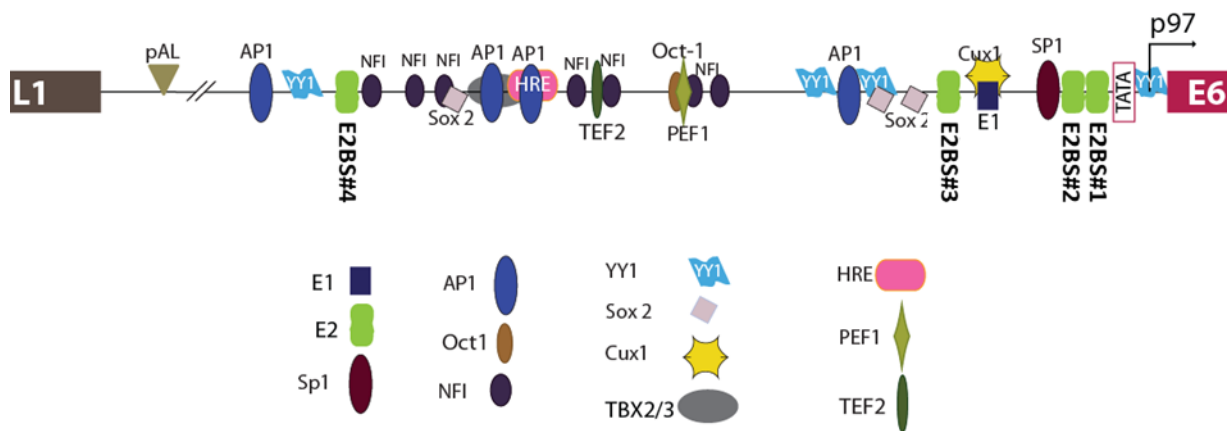


Figure 12 : Les principaux éléments *cis* et *trans* régulateurs de la LCR d'HPV16.

La LCR possède de nombreux sites de fixation pour une grande variété de FT cellulaires activateurs ou répresseurs présentés dans la figure 12, dont AP1, NF1, TEF1, SOX2, YY1 et des éléments de réponse aux hormones dans sa région centrale. Cette région centrale fonctionne en tant qu'enhancer stimulant ou réprimant la transcription virale (Ribeiro et al., 2018). YY1 va par exemple réprimer l'expression de E6 et E7 en induisant un changement de conformation de la chromatine au niveau de cet enhancer (Pentland et al., 2018). Il a d'ailleurs été montré que l'activation du promoteur tardif d'HPV16 dans les kératinocytes différenciés fait intervenir des MPT (Modifications Post-Traductionnelles) des

histones changeant la conformation du génome HPV (Peña and Laimins, 2001). Le changement d'expression entre le promoteur précoce et le promoteur tardif nécessite la différenciation des cellules basales infectées en kératinocytes. Des changements dans l'expression de FT cellulaires suite à la différenciation et leur interaction avec le génome viral favorise l'expression au niveau du promoteur tardif (Wooldridge and Laimins, 2008). La régulation épigénétique de la LCR fait aussi intervenir la méthylation des CpG qui vont moduler la fixation des éléments transrégulateurs (Thain et al., 1996). Cette régulation va de pair avec les MPT des histones compactant le génome HPV au niveau de l'enhancer, du p97 et du promoteur tardif p670 (Stükel and Bernard, 1999).

4.2 Régulation post-transcriptionnelle des transcrits HPV16

Les promoteurs p97 et p670 permettent la production de deux ARN prématurés, l'un possédant les ORF précoces et l'autre les ORF tardives. C'est ensuite leur maturation par divers évènements d'épissage alternatif qui aboutit à la formation des nombreux ARNm d'HPV16. La régulation des évènements d'épissages permet l'expression séquentielle des protéines au cours du cycle viral : E1 et E2, E6 et E7, E4 et E5 et finalement L1 et L2 (figure 5). Un troisième transcrit est produit à partir du promoteur pE8 et ne semble servir qu'à la production de la protéine de fusion E8^{E2}.

4.2.1 Le mécanisme d'épissage alternatif et sa régulation

La présence d'introns a pour la première fois été mis en lumière en 1977 par les travaux de Sharp et Roberts sur la cartographie des ARNm de l'Adénovirus type 2 en réalisant des hybrides ADN-ARN observés par microscopie électronique. Ils se rendirent compte que ces duplex présentent des séquences ou boucles non-hybridées et donc que l'ARNm provient vraisemblablement d'un assemblage de séquences non-continues, avec excision ou « splicing » de ces séquences présentes dans l'ADN mais pas dans l'ARNm mature (Berget et al., 1977; Chow et al., 1977). Sharp and Roberts obtinrent le prix Nobel en 1993 pour leur découverte. Environ 90% des transcrits cellulaires sont soumis à ce mécanisme d'épissage. La séquence de ces gènes est composée d'une alternance de séquences codantes, les exons, et de séquences non codantes, les introns.

L'épissage consiste en l'élimination des introns d'un transcrit dit primaire ou prématuré. L'intron est délimité en 5' par un site donneur d'épissage (SD) et en 3' par un site accepteur d'épissage (SA). L'intron possède également en son sein, une séquence de

branchement ou BPS (Branch Point Sequence), 15 à 40 nt en amont du SD et une séquence riche en pyrimidine ou PPT (PolyPyrimidine Tract) en amont du SD. Ces différents sites sont définis par des séquences consensus reconnues par le complexe d'épissage appelé le spliceosome. Les variations de ses séquences par rapport aux séquences consensus font varier leur reconnaissance et leur utilisation par le spliceosome. Un ARN prématuré possédant plusieurs sites d'épissage sera ainsi plus ou moins efficacement épissé et donnera plusieurs transcrits alternatifs.

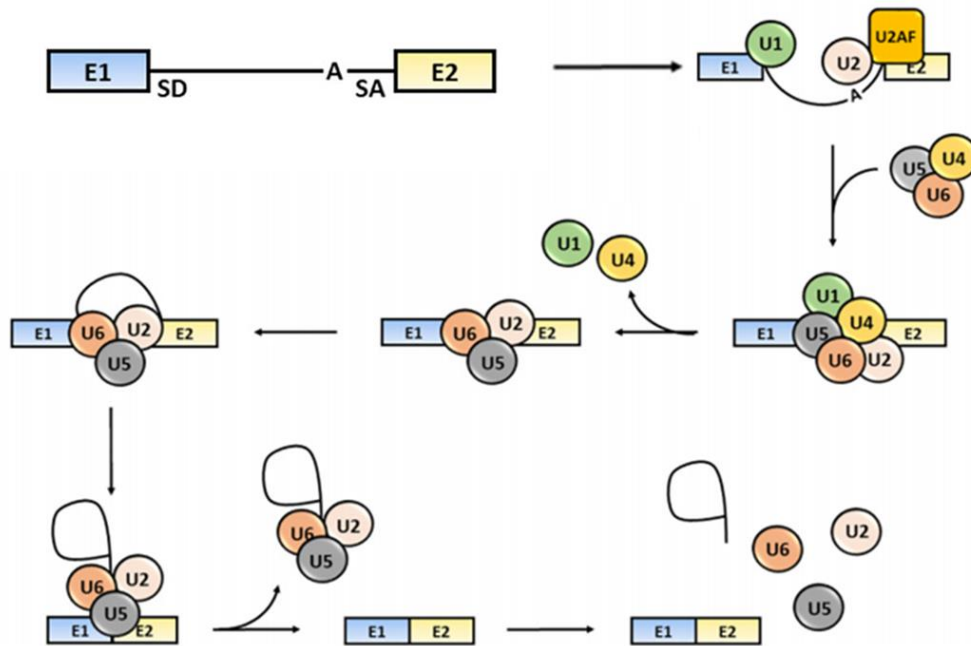


Figure 13 : Le mécanisme d'épissage.

Adapté de (Cerasuolo et al., 2020).

L'excision de l'intron est réalisée par un complexe ribonucléoprotéique composé des snRNA (small nuclear RNA) U1, U2, U4, U5 et U6, riche en uracile. Associés à des dizaines de protéines, ses snRNA forment la machinerie d'épissage : le spliceosome (Will and Lührmann, 2011). Comme montré en figure 13, U1 va reconnaître l'extrémité 5' et U2 le BPS. Avec U2AF qui reconnaît l'extrémité 3', ils forment le pré-spliceosome. Il y a ensuite recrutement de U4, U5, U6 et d'une vingtaine de snRNP qui s'assemblent pour former le spliceosome assemblé. S'en suit le départ de U1 et U4 et le recrutement de RNA hélicases ATP-dépendantes de la famille Prp qui permettent au pré-messager d'adopter la bonne conformation afin de réaliser deux réactions de transestérifications successives pour exciser l'intron. L'adénosine du BPS attaque le **GU** du SD ce qui forme un lariat (boucle) et libère l'exon en amont. Dans la deuxième réaction, c'est l'extrémité du SD qui va attaquer le **AG** du SA et ainsi complètement détacher le lariat intron. Les deux exons sont ligaturés et relâchés

du spliceosome post-catalytique. Le lariat intron est libéré quand le complexe se désassemble (Will and Lührmann, 2011) (Shi, 2017).

La maturation des pré-messagers est régulée par des facteurs d'épissage. Les protéines SR (Serine Arginine Rich, SRSF) et hnRNP (heterogenous nuclear RNP) (Lee and Rio, 2015) modulent le recrutement des composants du spliceosome et la définition des exons et des introns. En effet, ces protéines se fixent sur des séquences consensus *cis* régulatrices de l'ARN pré-messager (figure 14), inhibant ou favorisant la sélection des sites donneur et accepteur d'épissage par le spliceosome. La régulation de l'épissage dépend donc de la disponibilité de ces protéines régulatrices, qui permettent ou non la reconnaissance des sites d'épissage. Les protéines SR se fixent sur les ISE (Intronic Splicing Enhancer) et ESE (Exonic Splicing Enhancer) pour favoriser l'assemblage et l'efficacité du spliceosome, et les hnRNP inhibitrices se fixent soit sur les ISS (Intronic Splicing Silencer) ou ESS (Exonic Splicing Silencer) pour empêcher la fixation des facteurs SR.

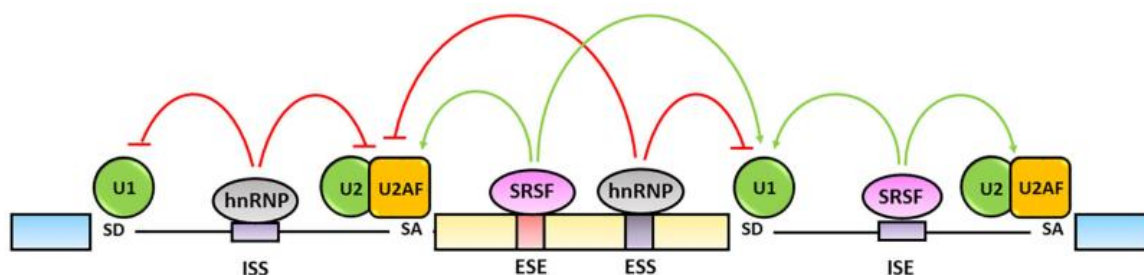


Figure 14 : La régulation de l'épissage.

Sont représentés les différents éléments *cis*-régulateurs dans les introns et exons avec la fixation des facteurs d'épissage de la famille SR et hnRNP. Les flèches représentent l'effet activateur (vert) ou inhibiteur (rouge) sur les facteurs généraux de l'épissage. Adapté de (Cerasuolo et al., 2020).

Au final, plusieurs facteurs entrent en jeu pour réguler la biogenèse des nombreux ARNm d'HPV16. Ces facteurs incluent la conservation des sites SD, BPS et SA, leur agencement et leur nombre sur les transcrits prématurés, ainsi que la présence de facteurs d'épissage et la conservation de leurs séquences *cis* régulatrices. Comme pour les FT, l'expression des facteurs d'épissage évolue avec la différenciation des cellules épithéliales en kératinocytes mais aussi au cours de la carcinogenèse (Mole et al., 2009). Les protéines virales dérèglent également ces facteurs d'épissage, rendant l'expression des ARNm et des protéines virales très dynamique (Johansson et al., 2012).

4.2.2 Le transcriptome d'HPV16



Figure 15 : Le transcriptome d'HPV16.

En haut est représenté le génome linéarisé d'HPV16, les flèches représentent les promoteurs, les triangles les sites de polyadénylation, et les ovales représentent les sites d'épissages, foncés donneur et clairs accepteurs. En dessous les différentes combinaisons d'épissage possible avec les introns excisés représentés en lignes grises infléchies. A droite les protéines potentiellement codées par ces transcrits.

La majeure partie du transcriptome d'HPV16 a été résolue entre les années 1985 et 95 par northern blot (Schwarz et al., 1985) (Baker et al., 1987) (Shirasawa et al., 1988), « nuclease protection assays » (Smotkin and Wettstein, 1986) (Shirasawa et al., 1991) (Rohlf et al., 1991) et RT-PCR (Cornelissen et al., 1990) (Sherman and Alloul, 1992). Il en existe plusieurs représentations dans la littérature, notamment dans (Baker and Calef, 1996), (Zheng and Baker, 2006), (Milligan et al., 2007), (Johansson et Schwartz 2013), (J. Chen et al. 2014) et (Graham and Faizo, 2017). Le génome peut être scindé en trois parties pour décrire les événements d'épissage aboutissant au transcriptome présenté en figure 15 : la partie E6E7, la

partie E1(E8)E2E4E5 et la partie L2L1. L'épissage des transcrits primaires aboutit à la formation d'une vingtaine d'ARNm. Il convient de rappeler que dans une proportion importante des cancers induits par HPV, l'intégration du génome viral au génome cellulaire aboutit à la transcription d'un unique transcrit prématuré bicistronique E6E7.

De manière intéressante, les HPV muqueux du genre α possèdent des sites d'épissage dans l'ORF E6 (figure 16). HPV16 possède notamment un SD en position 226 et un SA en position 409 qui permettent la formation du transcrit E6*IE7 (prononcé « E6 star one E7 ») par épissage d'un intron de 182 pb. A la différence des introns non-codants eucaryotes, cet intron que l'on nomme « intron 1 » code une partie de l'oncoprotéine E6. De fait, ce transcrit épissé peut produire E7 mais pas E6. En revanche, l'épissage de l'intron 1 de E6 entre le SD226 et SA409 provoque un décalage du cadre de lecture sur l'ARNm mature faisant apparaître un codon stop prématuré et créant une nouvelle ORF. L'ORF E6*I code une protéine tronquée de 43 a.a. dont les 41 premiers sont identiques à E6. Les HPVbr ne possèdent pas ces sites dans l'ORF E6 mais possèdent un promoteur pour E6 et un pour E7 (Smotkin et al., 1989). La découverte de cette particularité propre aux HPVhr a immédiatement soulevée la question du rôle de cet épissage dans la régulation de l'expression des oncoprotéines E6 et E7 mais également du rôle de l'isoforme tronquée E6*I dans le processus de carcinogenèse (cf. 5. La protéine tronquée E6*I d'HPV16).

HPV16 possède également un second SA dans E6 et un troisième SA dans E7 dont l'utilisation permettent la production d'ARNm appelés E6*IIIE7 et E6^E7, respectivement (Johansson et Schwartz 2013). Les ARNm codant la protéine E6 entière sont les produits de transcrits prématurés dont les sites d'épissage n'ont pas été reconnus par le spliceosome, donc non-épissés. Vingt-sept HPV du genre alpha possèdent les sites donneur et accepteur d'épissage qui permettent l'excision de l'intron 1 de E6 et la formation d'un transcrit E6*I (base de données PaVE). L'existence d'autres SA chez les autres génotypes est à envisager mais ils ne sont pas autant étudiés qu'HPV16. HPV33 est le seul génotype après HPV16 pour lequel un épissage E6*II a été décrit. De même, HPV18 est capable de l'épissage E6^E7 (Olmedo-Nieva et al., 2018).

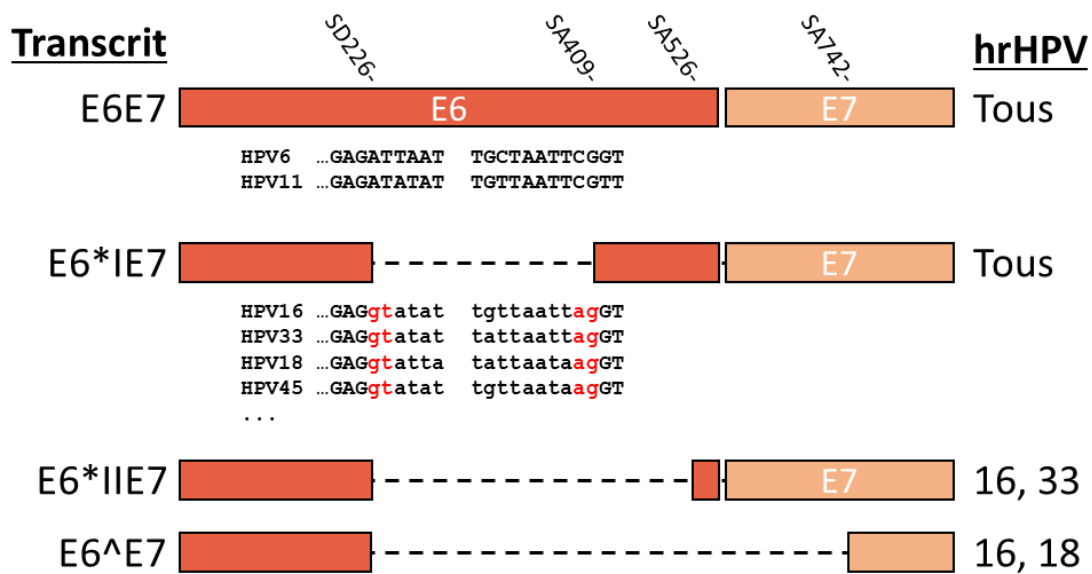


Figure 16 : Conservation des sites d'épissage E6*I chez les HPVhr.

Sont représentés les sites d'épissage E6*I des génotypes 16, 33, 18 et 45. Les sites d'épissage ne sont pas retrouvés chez les HPVbr comme HPV 6 et 11. Les HPVhr 33 et 18 sont les seuls avec 16 pour lesquels a été identifié un épissage E6*II et E6^E7, respectivement. Les bases en majuscules correspondent aux exons, tandis que celles en minuscules correspondent à l'intron 1.

Les sites d'épissage présents dans les ORF E1(E8)E2E4E5 et L2L1 sont bien plus conservés chez les HPV que ceux de E6E7 (Cerasuolo et al., 2020). La plupart des transcrits primaires sont épissés à partir du SD880 dans E1. L'épissage entre le SD880 et SA2582 forme un transcrit mature codant E1c et E2. L'épissage entre le SD880 et SA2709 forme un transcrit mature codant E2. L'épissage entre le SD880 et SA3358 forme un transcrit mature codant E4 et E5. Un dernier épissage potentiel est à mentionner entre le SD226 et le SA3358 formant un transcrit de fusion E6^E4 parfois appelé E6*X dans la littérature. On note également la présence d'ARNm transcrits à partir du p670 mais s'arrêtant au site pA précoce. Ces transcrits sont dépourvus de E6E7 et codent seulement pour E2 ou alors E1^E4 et E5. A noter que E4 ne possède pas de codon d'initiation et qu'il lui est donc fourni par E1. La proximité des sites SD880 et SD1302 fait que le second n'est jamais utilisé par les transcrits primaires qui proviennent du promoteur précoce ou tardif. Ce SD1302 n'intervient que dans l'épissage E8^E2C. En effet, dans ce cas la transcription qui commence au pE8 (p1127) donne un transcrit ne contenant pas le SD880. E8^E2c résulte donc de l'épissage entre le SD1302 et le SA3358 (Straub et al., 2015).

Les transcrits tardifs produits à partir du p670 sont aussi épissés à partir du SD880. L'épissage entre le SD880 et le SA3358 forme un transcrit mature codant E1^E4, E5, L2 et L1. Un second épissage entre le SD3632 et le SA5639 permet d'exciser les ORF E4 et L2, ce transcrit ne produit alors que E1^E4 et L1.

Les événements d'épissage alternatifs qui viennent d'être décrits sont soumis à une régulation par les facteurs d'épissage SR et hnRNP cellulaires (pour une revue, voir (Cerasuolo et al., 2020)). Le profil d'expression des transcrits alternatifs varie au cours du cycle viral. Encore une fois, HPV16 est parfaitement adapté à notre biologie dans l'expression séquentielle de ces protéines afin de réaliser un cycle viral productif. Le contexte en facteurs d'épissage cellulaires va aussi changer au cours de la carcinogenèse. Citons par exemple le rôle de hnRNP H qui se lie à des motifs triple G en aval du pA précoce dans l'ORF L2 (Oberg et al., 2003). Sa fixation favorise la polyadénylation des transcrits primaires au détriment de l'expression des transcrits tardifs. Or les niveaux de hnRNP H dépendent de la différenciation cellulaire. Au cours du cycle viral les kératinocytes se différencient et la diminution de hnRNP H favorise l'expression des protéines tardives dans les couches supérieures de l'épithélium (Oberg et al., 2005). En revanche, la dédifférenciation des cellules pré-tumorales et tumorales augmente l'expression de hnRNP H qui favorise l'expression des protéines précoces (Fay et al., 2009).

4.2.3 Régulation de l'épissage des transcrits prématurés E6E7

Plusieurs facteurs ont été identifiés comme régulant l'épissage du transcrit prématuré E6E7. La régulation de l'épissage de E6E7 est spéciale car une partie des transcrits primaires ne sont pas soumis au mécanisme d'épissage, un phénomène observé chez d'autres virus comme le VIH (Dyhr-Mikkelsen and Kjems, 1995). Ces transcrits non-épissés sont indispensables pour produire l'oncoprotéine E6.

Premièrement, les sites SD226 et SA409 sont suboptimaux dans leur séquence consensus (figure 17). Ensuite, la faible distance entre la coiffe et le SD226 favorise l'excision de l'intron 1 car la formation du spliceosome est dépendante de la coiffe (Zheng et al., 2004) et favorise aussi le choix préférentiel du SA409, puis 526, puis 742. Une étude a proposé qu'un branchpoint suboptimal en position 328 sur une Guanosine non-consensus puisse aussi favoriser la rétention partielle de l'intron 1, comme c'est le cas pour les transcrits rev d'HIV-1 (De la Rosa-Rios et al., 2006). Mais, l'intron 1 a été caractérisé plus encore par (Ajiro et al., 2012) montrant que l'épissage utilise un BPS consensus sur une Adénosine en position 385 et favorise ainsi l'utilisation du SA409 le plus proximal pour exciser le plus petit intron possible. Ces travaux identifient également deux nouveaux sites SD suboptimaux potentiels, SD191 et SD221.

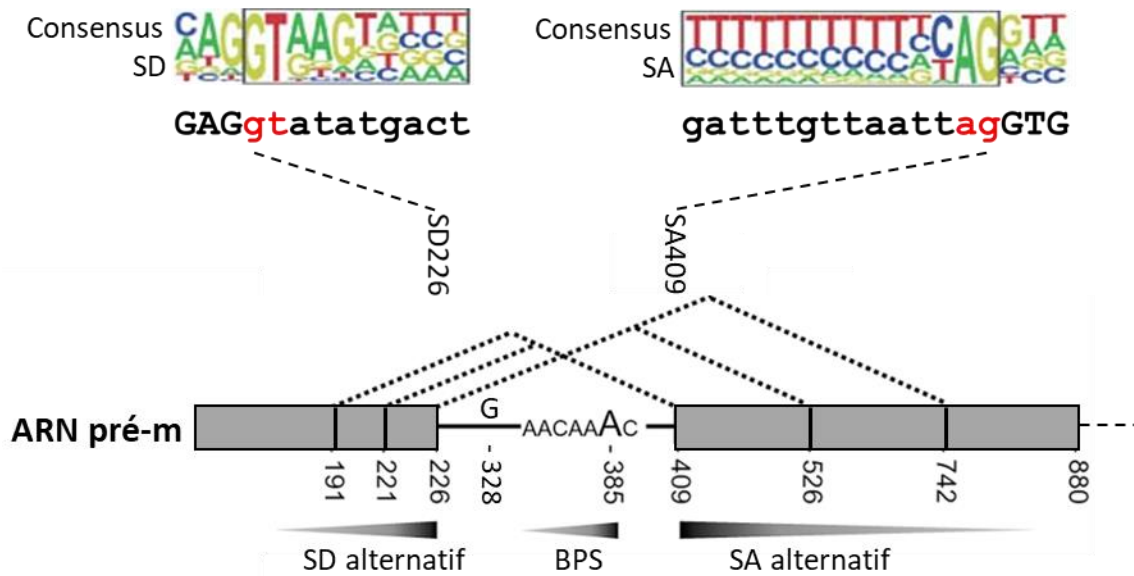


Figure 17 : Éléments *cis* régulateurs de l'ORF E6 d'HPV16.

Sont représentés de haut en bas, la conservation des sites SD226 et SA409, les positions des éléments *cis* régulateurs, et leur probabilité de sélection par le complexe du spliceosome. Adapté de (Ajiro et al., 2012).

Cet épissage dépend aussi des facteurs d'épissage cellulaire (figure 18), ce qui explique les variations du schéma d'épissage E6E7 entre les différents modèles d'études à notre disposition (Paget-Bailly et al., 2019). L'épissage des ORF E6 et E7 est majoritairement contrôlé par hnRNP A1 et A2 qui de manière surprenante ont d'abord été montrées comme stimulant l'épissage en se fixant sur l'intron 1, là où Sam68 et Brm inhibent l'épissage. Ces quatre facteurs sont modulés dans l'épithélium *via* la signalisation par l'EGF. L'EGF, d'autant plus exprimé dans les couches basales de l'épithélium, favoriserait ainsi la rétention de l'intron et la surexpression de E6 (Rosenberger et al., 2010). Il a été montré plus récemment que hnRNP A1 et A2 se fixent sur un ESS *cis* régulateur dans l'ORF E7 et inhibe l'épissage cette fois, ce qui semble plus conventionnel quant au rôle généralement imputable aux hnRNP (Zheng et al., 2020). Plus précisément, hnRNP A1 inhibe tous les épissages de E6 et favorise donc le transcrit non-épissé, là où hnRNP A2 inhibe le SA409 au profit du SA742 sans augmenter le taux de transcrits non-épissés. Ces données sont corroborées par l'étude d'HPV18 qui montre que hnRNP A1 se fixe aussi sur un ESS dans l'ORF E7 et empêche l'épissage 18E6*I entre le SD233 et SA416 (Ajiro et al., 2016). Il est intéressant de noter que les ARNm non épissés devraient être théoriquement retenus dans le noyau du fait qu'ils ne soient pas pris en charge par le spliceosome et donc considérés comme encore prématurés. Or, hnRNP A1 a été montré comme favorisant le transport des ARNm hors du noyau et pourrait donc aussi remplir cette fonction de régulation des transcrits HPV non-épissés (Roy et al.,

2014). Enfin, E6 et E2 d'HPV16 ont été montrées comme se liant au pré-messager E6E7 pour inhiber l'épissage (Bodaghi et al., 2009).

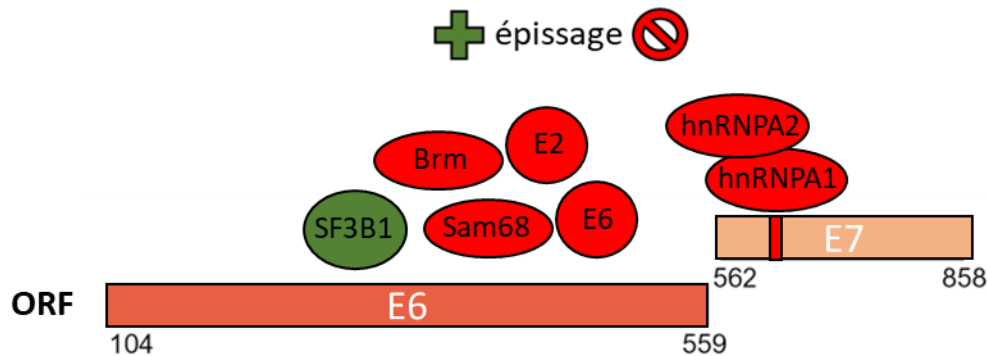


Figure 18 : Les facteurs d'épissage régulant le devenir du préARNm E6E7.

Sont représentés les facteurs d'épissage inhibant (rouge) ou favorisant (vert) l'épissage avec leur position de recrutement probable (la majorité sur l'intron1 de E6 et les hnRNP A1 et A2 sur l'ESS contenu dans E7).

Un seul facteur favorisant l'épissage de E6 a été identifié à l'heure actuelle. Le traitement de trois lignées de cancers des VADS induits par HPV avec un inhibiteur de SF3B1 ou un siRNA contre SF3B1 a montré un arrêt de l'épissage *I, indiquant que ce facteur favorise l'épissage des transcrits E6 (Gao et al., 2014).

Une étude a montré par crosslinking et spectrométrie de masse les facteurs SRSF2, SRSF6, SRp53, hnRNP C1/C2, Sam68 et hnRNP I/PTB comme se liant à l'intron 1 mais sans conclusion sur leur rôle (López-Urrutia et al., 2012). Plus récemment, cette liste d'interacteurs s'est encore étoffée avec de nombreuses protéines hnRNP, du spliceosome et des protéines formant l'EJC (Exon Complex Junction) qui se fixe en amont des jonctions exons-exons après épissage (Martínez-Salazar et al., 2014). D'autres facteurs d'épissage ont également été étudiés mais ne jouent pas de rôle dans l'épissage de E6 et E7. SRSF2 (SC35) semble stabiliser tous les transcrits alternatifs E6E7, il ne serait donc pas ici impliqué dans l'épissage mais dans la stabilité des transcrits en les protégeant possiblement du NMD (Nonsense Mediated mRNA Decay), un mécanisme de surveillance et de dégradation des ARNm. De même, SRSF3 semble avoir une action indirecte favorisant l'expression de E6 et E7 sans modifier l'épissage (McFarlane et al., 2015).

Etonnamment, l'efficacité d'épissage du transcrit prématuré E6E7 ne varie pas significativement entre infection productive, lésions précancéreuses et cancers (Cornelissen et al., 1990) (Kösel et al., 2007) (Häfner et al., 2008). Malgré la surexpression de certains facteurs d'épissage comme hnRNP A1 dans les cancers cervicaux, les transcrits prématurés E6E7 sont en moyenne toujours épissés avec la même efficacité entre le SD226 et le SA409

(Cerasuolo et al., 2020). Des travaux récents par RNA-seq ont permis de quantifier ces transcrits dans divers échantillons de lésions CIN et de cancers du col de l'utérus et n'indiquent pas non plus de changement significatif de l'épissage du transcrit prématuré E6E7 entre CIN et cancer (J. Chen et al. 2014). La variation de l'épissage des transcrits E6E7 ne semble donc pas un facteur clé de la progression tumorale. En revanche, des variabilités interindividuelles sont quand même observées entre les tumeurs induites par HPV et ont permis de stratifier ces tumeurs en fonction de l'épissage E6E7. Ces travaux récents s'intéressant à l'expression de E6*I sont détaillés en fin d'introduction (Zhang et al., 2016) (Qin et al., 2020).

En définitive, les transcrits E6E7 sont massivement pris en charge par la machinerie d'épissage cellulaire. Mais la présence de sites d'épissage suboptimaux et d'éléments *cis* régulateurs pour des facteurs répresseurs de l'épissage assurent une expression minimale mais constitutive de l'ARNm non-épissé E6E7 à partir duquel l'oncoprotéine E6 est produite. La part que joue chaque transcrits E6E7 dans la traduction de E7 est quant à elle plus controversée, de même que la traduction de l'isoforme tronquée E6*I.

4.2.4 Régulation de la polyadénylation, de l'export et de la stabilité des transcrits HPV

La stabilité des ARNm E6E7 va dépendre de l'état physique du génome viral. La stabilité de ces transcrits est régulée par plusieurs séquences ARE (AU-Rich Element) *cis* régulatrices présentes dans sa région 3'. Ces éléments sont reconnus par des protéines modulant la stabilité et/ou la traduction de l'ARNm comme AUF1 ou HuR (Barreau et al., 2005). Il existe plusieurs classes d'ARE. Le premier à avoir été découvert est un ARE de classe 3 riche en AU dans sa 3'UTR, en aval de l'ORF E5 qui favorise la dégradation du transcrit (Jeon and Lambert, 1995). Des travaux plus récents par (Ehrig et al., 2019) ont identifié 3 ARE de classe 1 (AUUUA) dans l'espace intergénique entre E4 et E5 déstabilisant plus fortement le transcrit précoce. Ces éléments rendent le transcrit précoce peu stable (demi-vie de 3h). En revanche, l'intégration du génome viral entraîne d'une part la perte des 3 ARE et de la région AU-rich, et d'autre part entraîne la formation de gènes et de transcrits de fusion. De plus, l'étude de ses intégrations révèle que dans 50% des cas la partie cellulaire du transcrit de fusion ne contient aucun ARE (Schmitz et al., 2012). Il en résulte une stabilisation des transcrits E6E7-ARNcellulaire (demi-vie de 6 à 12h) qui participe à la surexpression des oncogènes viraux au cours de la carcinogénèse (Jeon and Lambert, 1995). Les protéines cellulaires se fixant sur ces ARE et déstabilisant le transcrit primaire n'ont pas été étudiées.

Le facteur d'épissage SRSF2 (SC35) a été montré comme impliqué dans la stabilité des transcrits E6E7 en les protégeant possiblement du NMD (Nonsense Mediated mRNA Decay) (McFarlane et al., 2015). Le NMD est une voie cellulaire de surveillance des ARNm présentant des PTC (Premature Termination Codon). Un codon stop est défini comme prématuré par sa localisation relative par rapport au complexe protéique EJC, déposé après épissage 20 à 24 nt en amont des jonctions exon-exon. Au cours du premier cycle de traduction, les EJC vont être enlevés successivement de l'ARNm au passage du ribosome et la traduction va s'arrêter au codon STOP présent dans le dernier exon. Sur un ARNm contenant un PTC, la traduction s'arrête avant d'avoir enlevé tous les EJC. La conformation du ribosome et de l'EJC va alors servir de plateforme pour la formation du complexe protéique SURF qui aboutit à l'activation par SMG1 de UPF1 par phosphorylation. L'activation du NMD aboutit à la dégradation du transcrit par des enzymes de décapping et des exoribonucléases. Ce mécanisme sert à la fois à réguler des transcrits physiologiques présentant des PTC et à dégrader les transcrits pathologiques produits par des gènes contenant des mutations nonsense. Les ARNg de nombreux virus sont aussi dégradés par le NMD, qui représente un mécanisme de défense antivirale à part entière (Leon and Ott, 2020). A l'inverse, certains virus possèdent des mécanismes de protection contre cette voie de surveillance. Il a récemment été montré que les protéines E6 et E7 d'HPV18 peuvent inhiber le NMD (Rigalli et al., 2018). Les transcrits HPV avec leurs multiples épissages et codons STOP représentent des candidats privilégiés pour cette voie de surveillance. Une partie des travaux de notre laboratoire s'intéresse d'ailleurs à déterminer si les transcrits E6E7 d'HPV16 sont soumis au NMD (cf. annexes).

Pour finir, certains microARN ont été identifiés comme se fixant sur le transcrit précoce pour le déstabiliser et diminuer l'expression de E6 et E7 comme le miR-375 (Jung et al., 2014) ou les miR-875 et -3144 (Lin et al., 2015). Ces deux derniers miR induisent même un changement des ratio de transcrits épissé/non-épissés en ne dégradant que le transcrit non-épissé E6E7 (Li et al., 2018).

4.3 Régulation traductionnelle des isoformes E6 et de E7

Les ARNm cellulaires sont majoritairement monocistroniques. C'est en partie pour s'adapter à cette contrainte que les HPV réduisent le nombre d'ORF sur les transcrits primaires par épissage alternatif. En effet, la traduction dépendante de la coiffe va privilégier la première ORF rencontrée sur un transcrit, en le scannant à la recherche du premier codon

d'initiation AUG. Malgré leurs nombreux événements d'épissage, les ARNm HPV restent en totalité polycistroniques ou au moins bicistroniques quand ils sont transcrits à partir du génome intégré dans les cancers. Certains transcrits cellulaires sont aussi bicistroniques. Sur ces transcrits cellulaires, on distingue une uORF (upstream ORF) et une mORF (main ORF) en aval. La différence entre les ARNm cellulaires présentant des uORF et les ARNm polycistroniques d'HPV réside dans le fait que toutes les ORF des transcrits HPV sont codantes, là où les uORF cellulaires servent généralement à réprimer l'expression de la mORF en aval. La question s'est donc posée de la régulation de la traduction des différentes ORF HPV et notamment de E7 avec l'ORF E6 en amont, et de l'impact de l'épissage de l'intron 1 de E6 sur la traduction de E7.

La littérature présente des données contradictoires concernant la traduction de E7. La première étude *in vitro* a montré que la protéine E7 est produite avec la même efficacité par l'ARNm épissé E6*IE7 ou non épissé E6E7 (Tan et al., 1994b). Il a ensuite été montré que la traduction de E7 dépend de la coiffe et du scanne de l'ARNm E6E7 non-épissé bicistronique (Stacey et al., 1995). Une autre étude incluant les différents transcrits épissés ou non a montré dans des cellules de cancer cervical HPV- C33-A que les trois principaux transcrits E6E7, E6*IE7 et E6*IIIE7 servent à la traduction de E7 mais ici encore, la majorité de la traduction se fait à partir de l'ARNm non-épissé E6E7 (del Moral-Hernández et al. 2010). Dans ces études, l'épissage entre le SD226 et le SA409 ne semble pas réguler la traduction de E7 qui est initiée par leaky scanning (Stacey et al., 2000). Le mécanisme de leaky scanning se produit lorsque le premier codon d'initiation AUG d'un transcrit n'est pas reconnu par le complexe de pré-initiation de la traduction. Comme pour les sites d'épissage, il existe des codons d'initiation optimaux et suboptimaux définis par la séquence KOZAK plus ou moins consensus entourant l'AUG (Kozak, 2002). Le complexe d'initiation choisi alors un codon d'initiation en aval, celui de E7 et pas celui de E6.

D'autres études ont montré au contraire que l'épissage du transcrit E6E7 entre le SD226 et le SA409 est indispensable à la production de l'oncoprotéine E7 (Smotkin and Wettstein, 1986) (Smotkin et al., 1989) (Sedman et al., 1991). Il est proposé que l'excision de l'intron 1 rapproche l'AUG de E7 de la coiffe (Zheng et al., 2004) mais surtout que la forte augmentation de l'espace inter-cistronique, qui passe de 2 entre E6 et E7 à 144 nt entre E6*I et E7, permet la réinitiation du ribosome (figure 19) (Tang et al., 2006). Des études similaires sur HPV18 démontrent également que l'épissage est indispensable à la traduction de E7 *via* la réinitiation du ribosome (Tang et al., 2006) (Brant et al., 2019). Lorsqu'une petite sous-unité

40S ribosomale reste liée à un transcrite après la traduction d'une première ORF, elle continue à se déplacer de 5' en 3' sur le transcrite. La réinitiation du ribosome permet la traduction d'une seconde ORF en aval si une sous-unité 60S se fixe sur la 40S et reforme un ribosome avant le passage du prochain codon d'initiation. Etant donné que la sous-unité 60S ribosomale doit se détacher à la fin de la traduction de la première ORF puis se rattacher avant d'avoir atteint l'AUG de la seconde, la réussite de ce mécanisme est dictée par la distance séparant les 2 ORF, plus elle est grande, meilleure est la réinitiation.

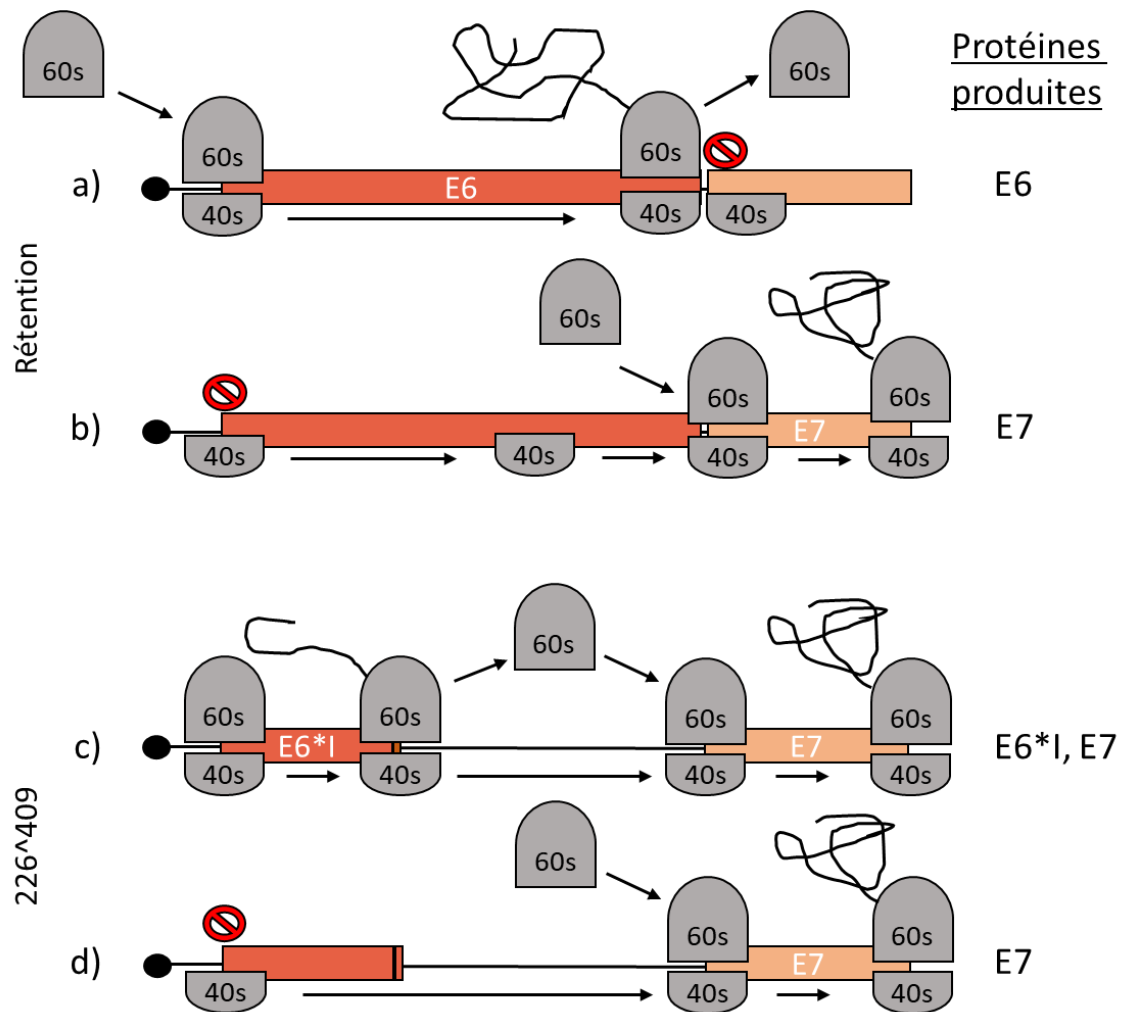


Figure 19 : Régulation et mécanisme de traduction des protéines E6, E7 et E6*I.

a) Le complexe de pré-initiation de la traduction reconnaît l'AUG 104 et traduit E6, le ribosome ne peut pas traduire E7 car l'espace intercistronique de 2 nt ne lui permet pas de se fixer à temps. b) L'AUG 104 est ignoré par le CPI ainsi que les autres AUG présents dans E6 (leaky scanning), le PIC reconnaît l'AUG E7 et traduit E7. c) Le ribosome reconnaît l'AUG et traduit E6*I, le ribosome peut se réinitialiser du fait des 144nt d'espace intercistronique pour traduire E7. d) L'AUG 104 est ignoré par le CPI ainsi que les autres AUG présent dans E6*I (leaky scanning), le PIC reconnaît l'AUG E7 et traduit E7.

Chez les HPVbr, cet aspect de la régulation traductionnelle ne se pose pas du fait que les ARNm E6 et E7 sont produits à partir de deux promoteurs différents (Smotkin et al.,

1989). Il est intéressant de noter que placer l'ORF 6E7 à la place de 16E7 dans un ARNm bicistronique provoque l'abolition de l'expression de E7 (Barbosa et al., 1991). On peut donc supposer qu'outre l'épissage de l'intron 1 de E6, des éléments régulateurs dans l'ORF 16E7 permettent sa traduction à partir de l'ARN non-épissé.

La traduction des différents ARNm E6E7 est aussi régulée par l'utilisation de codons rares qui limitent la vitesse de traduction *via* l'indisponibilité en ARNt spécifiques. Si un codon est peu représenté dans le génome humain, l'ARNt correspondant est lui aussi peu présent dans la cellule. A noter que ce biais d'usage de codons pourrait aussi entrer en jeu dans l'expression sélective des protéines virales au cours du cycle productif de l'infection. Nous l'avons vu, la migration des cellules infectées dans l'épithélium cervical et leur différenciation entraînent un changement d'expression de facteurs de transcription et d'épissage en faveur du promoteur tardif p670 et de l'expression de L1 et L2. Il serait intéressant d'étudier l'impact de la différenciation cellulaire sur l'expression des ARNt et les conséquences sur l'expression des protéine d'HPV (Rak et al., 2018). Il est proposé que ce biais de codons et l'expression réduite des protéines virales soit une stratégie évolutive des HPV pour protéger la viabilité des cellules infectées et favoriser l'échappement au système immunitaire (Cid-Arregui et al., 2003). Concernant nos travaux, on comprend alors facilement que ce biais additionné de l'épissage massif E6*I aboutisse à une expression très faible de l'oncoprotéine E6, difficilement détectable dans nos modèles d'étude. Nous reviendrons en détails sur ce point dans la suite de ce manuscrit.

5 La protéine tronquée E6*I d'HPV16

La séquence de la protéine E6*I est identique aux 41 premiers résidus de E6 et possède donc le même NLS N-terminal ainsi que le premier demi-site CxxC doigt de zinc (figure 20). La localisation de E6*I, étudié grâce à une protéine de fusion GFP-E6*I, a été montrée comme majoritairement nucléaire. La mutation du NLS1 que ce soit dans E6 ou E6*I entraîne d'ailleurs leur accumulation cytoplasmique (Tao et al., 2003) (Vaeteewoottacharn et al., 2005). Sa dégradation, comme celle de E6, se fait par le protéasome (Guccione et al., 2004). Sa nature plus polypeptidique que protéique rend l'étude de sa structure difficile. Une étude physico-chimique s'est penchée sur les différentes conformations qu'E6*I peut prendre. E6*I présente des repliements en hélice α à pH physiologique et forme de larges oligomères en absence de *zinc* via l'oxydation de ses résidus cystéine, tout comme l'isoforme entière E6 (García-Alai et al., 2007). En revanche, l'ajout de *zinc* semble permettre le maintien de dimères de E6*I qui présentent des structures secondaires résiduelles (Heer, Alonso, et de Prat-Gay 2011).

Les premières études sur E6*I se sont intéressées à l'effet que cette isoforme tronquée pouvait avoir sur la cible la plus notoire de E6, à savoir p53. L'étude de (D. Pim, Massimi, et Banks 1997) a montré que 18E6*I peut se lier à E6AP, 16E6 et 18E6 mais pas p53. Là où, (Shally et al., 1996) l'année précédente avaient montré que 16E6*I peut faiblement interagir avec p53. Ces interactions entraînent une hausse des niveaux de p53 dans des cellules HPV+. L'effet sur p53 n'est pas retrouvé dans des cellules HPV- exprimant uniquement 18E6*I ce qui semble indiquer que 18E6*I perturbe la formation du complexe E6/E6AP/p53. Cette hausse des niveaux de p53 entraîne une transactivation de ses gènes cibles et une diminution de la capacité des cellules à former des colonies (D. Pim, Massimi, et Banks 1997). Précédemment, (Sedman et al., 1991) ont montré que la coexpression de 16E6*I avec E6 et E7 dans des kératinocytes diminue également la capacité à former des colonies immortalisées. Une seconde étude par (Pim and Banks, 1999) vient compléter ces résultats en montrant que les quatre résidus de E6*I en amont du demi-site doigt de *zinc* et le demi-site sont indispensables pour l'interaction avec E6AP, 18E6 et 16E6, et l'inhibition de la dégradation de p53. De plus, ils montrent que la diminution de prolifération des cellules HPV+ exprimant 18E6*I provient d'une induction de l'apoptose dépendante de p53. Le fait que 16E6*I colocalise avec E6AP (Kao et al., 2000) et nos propres données semblent indiquer que les isoformes tronquées E6*I d'HPV18 et HPV16 partagent la même capacité à empêcher la dégradation de p53, bien que leurs séquences, représentées en figure 20, ne partagent que 60%

d'identité. Il a ensuite été montré que les niveaux d'expression de E6*I, dans des cellules HPV+ et HPV-, dépendent du cycle cellulaire contrairement à E6. L'expression ectopique de E6*I dans les Ca Ski et les U-2 OS est maximale en G1/S et G2/M, respectivement. Cette expression accrue aboutit à une stabilisation de p53 (Guccione et al., 2004). Trois cas de figure peuvent expliquer ces observations : i) E6*I et E6 interagissent entre elles voire s'hétéro-oligomérisent s'annulant l'une l'autre, ii) E6*I interagit avec E6AP et empêche la formation du complexe tripartite E6/E6AP/p53 et empêche possiblement la stabilisation de E6 par E6AP, iii) E6*I interagit avec p53 ce qui empêche aussi la formation du complexe.

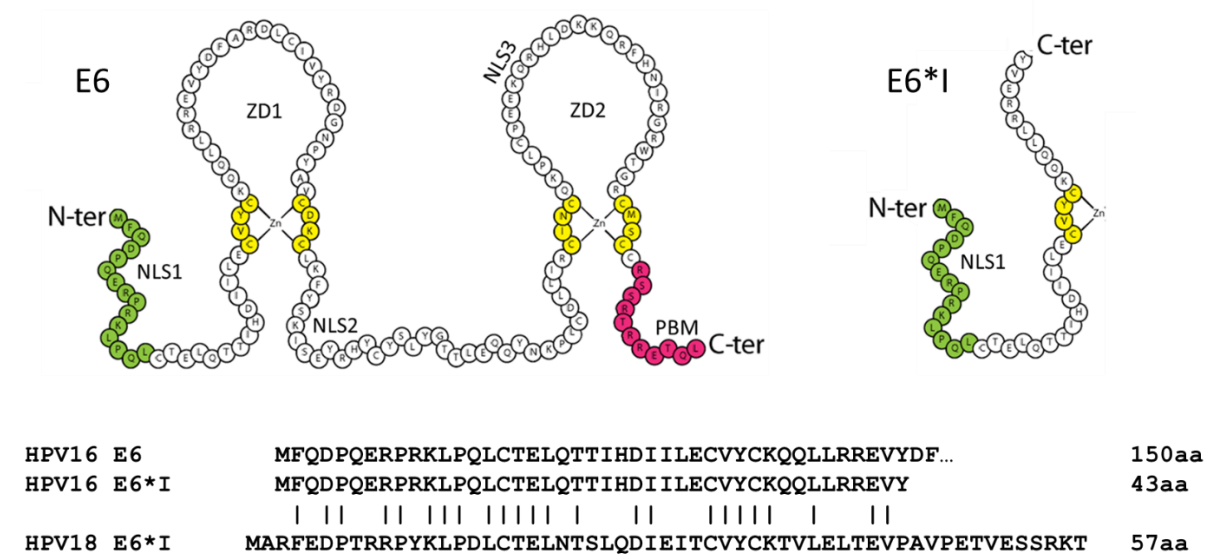


Figure 20 : La protéine tronquée E6*I d'HPV16.

E6*I conserve de E6 sont NLS1 N-terminal et un demi site doigt de zinc. Les protéines tronquées des deux génotypes ne présentent que 56% d'homologie.

Pour finir, des études s'intéressant aux variants d'HPV18 ont montré que les différences d'épissage de E6 se répercutent sur la dégradation de p53. Les variants aboutissant à un plus fort épissage (Africain) entraînent une plus faible dégradation de p53 et une capacité diminuée à former des tumeurs en xénogreffe, un effet qui peut s'expliquer par une plus faible expression de E6, mais aussi par l'expression accrue de E6*I (De la Cruz-Hernandez, 2005). Dans la continuité de ce travail, (Vazquez-Vega et al., 2013) ont montré que la dégradation de p53 par les variants produisant le plus fort taux de E6 s'accompagne de l'expression du suppresseur de tumeur p14^{ARF}. De manière intéressante, l'expression ectopique de 18E6*I seul suffit aussi à augmenter l'expression de p14^{ARF}, mais sans modification les niveaux de p53. L'interaction de E6*I avec p53 pourrait empêcher p53 d'inhiber p14^{ARF}. Toutes ces données semblent indiquer que E6*I joue un rôle d'antagoniste sur E6 et sa capacité à dégrader p53. L'effet de E6*I ici serait donc plutôt anti-oncogénique. La relation dynamique

entre isoforme entière et tronquée de E6 ne se limite pas à la simple inhibition des effets de E6 par E6*I.

E6*I est aussi capable d'agir sur la voie extrinsèque de l'apoptose. Rappelons que E6 est capable d'inhiber la voie extrinsèque très en amont, en s'associant à TNF R1, TRADD, FADD et en entraînant leur dégradation (Filippova et al., 2004). E6 est également capable de réduire les niveaux de procaspase 8 vers laquelle converge l'activation des voies TRAIL, TNF et Fas. De manière intéressante, là où l'effet de E6 est dose dépendent sur l'inactivation de la voie Fas, une trop forte expression de E6 resensibilise les cellules à la voie TNF (Filippova et al., 2005). Cette observation a mené à l'étude de l'impact que E6*I pouvait avoir sur cette voie extrinsèque car les deux isoformes étaient précédemment exprimés ensemble. Et en effet, si l'interaction de E6 avec la procaspase 8 entraîne sa dégradation et l'empêche d'interagir avec elle-même et FADD, l'interaction de E6*I entraîne la stabilisation de la procaspase 8 mais sans gêner son interaction avec elle-même ou avec FADD (Filippova et al., 2007) (Tungteakkhun et al., 2010). La poursuite de ces travaux a montré que E6 et E6*I protègent les cellules du TNF quand elles sont exprimées séparément. Mais cela n'explique pas pourquoi la coexpression des deux isoformes à faible niveau protège du TNF, alors qu'une forte expression sensibilise de nouveau ces cellules. La forte coexpression des deux isoformes va en fait entraîner la formation d'hétérodimères E6/E6*I agissant comme un pseudo complexe DISC, E6 liant TNFR1, et E6*I liant la procaspase-8 menant à son activation et à l'apoptose (Filippova et al., 2009). Plus récemment il a été montré que 18E6*I interagit aussi avec la procaspase 8 et même plus fortement que 16E6*I (Manzo-Merino et al., 2014).

Pour finir, les propriétés anti-oncogénique de 16E6*I ont été confirmées en réalisant des xéno greffes dans des souris de cellules de cancer cervical SiHa HPV+ et C33-A HPV- préalablement transfectées ou non par 16E6*I. Ces expériences montrent que la taille des tumeurs exprimant E6*I est inférieure à la taille des tumeurs contrôles. Le fait que l'effet antiprolifératif de E6*I soit moindre dans les tumeurs C33-A HPV- comparé aux tumeurs SiHa HPV+ indique deux choses. Cela indique premièrement que 16E6*I antagonise certains effets oncogéniques de E6, corroboré par l'interaction de 16E6 avec 16E6*I empêchant la dégradation de p53 démontré une nouvelle fois dans cette étude. Et deuxièmement, que 16E6*I a des effets anti-oncogéniques indépendants de E6 et de p53 qui expliquent la diminution de la formation de tumeurs dans un modèle non seulement HPV- mais surtout p53^{mut} (Filippova et al., 2014a).

En opposition à ses propriétés anti-oncogéniques, il a été montré que l'expression ectopique de 16E6*I est capable de stimuler la production des espèces réactives de l'oxygène appelées ROS (Reactive Oxygen Species) que ce soit dans des modèles HPV+ ou HPV-. E6*I induit cette surproduction de ROS en diminuant les défenses cellulaires antioxydantes *via* la diminution de SOD2 (Super Oxyde Dismutase) et Gpx1/2 (Glutathion peroxydase) qui transforme normalement les ROS en H₂O₂ et en eau, respectivement. Ces ROS à terme provoquent des dommages à l'ADN (Williams et al., 2014). Il a donc été proposé que ces effets de E6*I sur les dommages à l'ADN puissent s'ajouter à l'instabilité génomique provoquée par E6 et E7 pour favoriser l'intégration des épisomes viraux au génome cellulaire (Williams et al., 2010) (Chen et al., 2014b). Effectivement, (Chen Wongworawat et al., 2016) ont confirmé que le stress oxydatif provoqué par E6*I facilite l'intégration d'épisomes HPV16 dans le génome cellulaire, là où E6 a un effet plus modéré. Une autre étude a montré que l'expression des oncoprotéines virales dans des lignées « tête et cou » de carcinomes HPV16+ entraîne une surproduction de ROS *via* l'activation du complexe NOX2 (NADPH Oxydase 2) qui provoque d'importants dommages à l'ADN (Marullo et al., 2015). Cette étude émet l'hypothèse que l'induction des ROS *via* le complexe Nox2 puisse provenir des effets de E6*I. Une étude de protéomique confirme les effets de E6*I sur les ROS montrant dans des lignées HPV+ et – exprimant E6*I une diminution du pool cellulaire en GSH (Glutathion) antioxydante, une augmentation des dommages à l'ADN et un impact de E6*I sur les protéines mitochondriales impliquées dans les oxydations phosphorylantes et la réponse aux dommages à l'ADN (Evans et al., 2016). Il a été montré plus récemment que la coexpression de E6 et E6*I ne fait plus apparaître la dérégulation de SOD2 mais toujours une diminution de la GSH, aboutissant quand même à des dommages à l'ADN. Il est possible que les effets de E6*I soient partiellement masqués par l'isoforme entière, qui est également génératrice de ROS. En revanche, l'ajout de E7 contrebalance les effets pro-ROS de E6 et E6*I en augmentant les niveaux de catalase (Cruz-Gregorio et al., 2018a). Pour finir, la coexpression de E6 et E6*I d'HPV16 ou 18 aboutit à la production de ROS *via* l'activation et le découplage de la respiration mitochondriale (Cruz-Gregorio et al., 2019). Ces études montrent que 16E6*I est capable d'induire une production de ROS ayant des effets biologiques en lien avec la carcinogenèse induite par HPV (Cabeça et al., 2019). Pour finir, des travaux étudiant le transcriptome d'échantillons de cancers de l'oropharynx induits par HPV16 ont montré que l'expression de l'ARNm E6*I corrèle avec une inflammation et un stress oxydant accru, ainsi qu'une intégration et un nombre de mutations génétiques accru (Zhang et al., 2016). Parmi les

DEG surexprimés dans ces tumeurs sont retrouvés les transcrits codant les protéines Duox1, A1, 2 et A2 faisant partie de la famille des Nox. Au niveau clinique, l'expression de E6*I dans cette cohorte corrèle aussi avec l'augmentation du volume tumoral et une survie moindre (Qin et al., 2020).

L'augmentation du niveau de stress cellulaire induite par E6*I module également la survie des cellules cancéreuses soumises à un traitement par chimiothérapie ou radiothérapie (Filippova et al., 2014b). En effet, l'expression de E6 et E6*I (sans E7) dans des lignées de cancers de l'oropharynx positives à HPV16 provoque une sensibilité accrue à la radiothérapie. La sensibilisation induite par E6*I est 8 fois supérieure à celle induite par E6. E6*II est montré comme sans effet (Pang et al., 2011). A priori en contradiction, il a été montré une corrélation entre l'expression de E6*I et une résistance aux UVB et au H₂O₂ dans des kératinocytes (Mouret et al., 2005). On peut supposer que les effets des isoformes de E6 sont au moins partiellement modèle-dépendant et ne sont observés que dans les modèles exprimant les facteurs pro- ou anti-oxydatifs sur lesquels ils peuvent agir, chaque modèle étant aussi plus ou moins sensible au stress oxydant (Olmedo-Nieva et al., 2018). Concernant les agents de chimiothérapies, E6 et E6*I ont un effet protecteur face à un traitement au cisplatine en inhibant l'apoptose, là où E6*II sensibilise grandement à l'apoptose (Vaisman et al., 2018). E6^E7 et E7 n'ont pas d'effet sur l'apoptose dans cette étude. Cette chimiorésistance conférée par E6*I peut être en relation avec sa capacité à transactiver des gènes de la famille des aldo-kéto réductases impliqués dans la résistance aux drogues (Wanichwatanadecha et al., 2012).

Une étude récente a montré que 16E6*I entraîne une forte surexpression d'IL-6 dans des kératinocytes. Se faisant, E6*I promeut un micro-environnement pro-inflammatoire chronique et ses effets viennent s'ajouter à ceux déjà démontrés de E6 et E7. En effet, les protéines virales sont connues pour maintenir cet état pro-inflammatoire car il est pro-prolifératif (Artaza-Irigaray et al., 2019). Ce stress induit (pro-ROS et pro-inflammatoire) par E6*I pourrait être indispensable, dans un premier temps au cycle viral, puis incidemment à la carcinogenèse. Ce même stress maintenu au cours du temps, favoriserait un métabolisme énergétique accru, l'intégration des épisomes viraux et l'accumulation de mutations génétiques nécessaires à la carcinogenèse.

Objectifs

L'expression du génome des HPV est finement régulée au niveau post-transcriptionnel par épissage alternatif. La découverte de sites d'épissage propre aux HPVhr au sein des ORF E6 et E7 a très tôt posé la question de l'impact de cette régulation sur l'expression des oncogènes et sur les propriétés oncogéniques des HPVhr en général. Il n'existe pas à l'heure actuelle de consensus sur cette question. Si l'épissage semble en effet réguler négativement l'expression de E6, les études portant sur la traduction de E7 sont contradictoires. De plus, ces événements d'épissage aboutissent à la formation de transcrits alternatifs codant potentiellement de nouvelles isoformes des protéines E6 et E7. Les fonctions et la place de ces isoformes dans le cycle viral mais surtout la carcinogenèse reste largement débattue.

Les problématiques posées par ce nouveau mode de régulation sont donc multiples. L'objectif principal de ce travail de thèse a été de déterminer si l'expression ectopique de l'isoforme E6*I est capable d'impacter le phénotype cellulaire et si les effets observés peuvent être mis en lien avec la carcinogenèse induite par les HPV. Pour cela, nous avons généré des clones cellulaires exprimant stablement E6*I. Nous avons également produit des clones coexprimant E6*I et E6 afin d'observer si les deux isoformes interagissent entre elles. Les effets des isoformes ont été mesurés à l'aide de tests phénotypiques et un séquençage du transcriptome entier des clones stables. En accord avec la littérature, l'expression de E6*I dans notre modèle U-2 OS aboutit à une surproduction de ROS. De plus, certaines cibles identifiées par RNA-seq comme dérégulées par E6*I sont impliquées dans l'activation des complexes Nox, des protéines cellulaires génératrice de ROS. Nous avons donc ensuite inhibé l'activation des protéines Nox dans les cellules exprimant E6*I afin de déterminer si l'augmentation des ROS résulte d'une activation de ces complexes.

Cet effet quantifiable de E6*I sur le métabolisme des ROS nous a ensuite permis d'étudier si les effets de l'expression de E6*I sont effectivement dus à la traduction d'une protéine E6*I et ce à l'aide de constructions plasmidiques générant des ARNm E6*I incapable de traduction. En parallèle, la traduction de l'ARNm E6*I a été étudiée par profilage polysomique.

Nous avons enfin étudié l'impact de l'épissage alternatif E6*I sur la traduction de l'ORF E7 dans le modèle U-2 OS. Pour ce faire, nous avons générés de nombreuses constructions plasmidiques nous permettant d'isoler les éléments des différents ARNm E6E7 pouvant influencer sur la traduction de E7.

Résultats

1 Impact des isoformes E6 sur le transcriptome cellulaire

1.1 Résumé de la publication

Dans cette publication nous avons étudié l'impact de l'expression de E6*I d'HPV16 sur le transcriptome cellulaire. Pour cela, nous avons établi un modèle cellulaire exprimant les différentes isoformes E6 et/ou E6*I de manière stable dans la lignée U-2 OS. Neuf clones cellulaires ont été sélectionnés : trois clones U-CTL (transfection pXJ40-vide), trois clones U-E6All coexprimant E6 et E6*I (transfection pXJ40-E6All), et trois clones U-E6*I (transfection pXJ40-E6*I). Nous avons soumis ces clones à un séquençage du transcriptome entier et nous avons réalisé une étude d'expression différentielle des gènes cellulaires en comparant leur expression entre les trois conditions. Les ensembles de gènes dérégulés obtenus ont ensuite été regroupés par voie et mécanisme cellulaire, autrement dit par ontologie. Notre attention s'est portée particulièrement sur les cibles de U-E6*I et si ces cibles étaient retrouvées dans U-E6All. On observe un faible recouvrement (11 gènes) entre les 419 DEG (Differentially Expressed Genes) de U-E6All et les 53 de U-E6*I. Ce qui semble indiquer une disparition des effets de E6*I quand elle est coexprimée avec E6. Les 53 DEG U-E6*I s'intègrent significativement dans deux ontologies : « réponse à une diminution des niveaux d'oxygène » (Gene Ontology :0036293) et « matrice extracellulaire » (GO : 0030198). Les DEG U-E6All s'intègrent dans des ontologies déjà décrites dans la littérature mais de nouvelles cibles cellulaires dérégulées ont aussi été identifiées. En accord avec les résultats de transcriptomique, nous avons mesuré une surproduction de ROS dans les clones U-E6*I comparativement aux clones U-E6All et U-CTL. Nous avons également validé la dérégulation des cibles du RNA-seq par RT-qPCR et observé une forte corrélation entre les deux techniques. Enfin, nous avons validé la dérégulation de cibles impliquées dans le métabolisme des ROS par E6*I dans le modèle de W12 naturellement infecté par HPV16, où l'on retrouve l'abrogation des effets de E6*I sur ces cibles en présence de E6.

Nous avons mesuré pour la première fois l'impact de E6*I d'HPV16 sur le transcriptome cellulaire. Nous montrons une corrélation entre la modification du transcriptome et des dérégulations phénotypiques. Ces dérégulations incluent une surproduction de ROS, un effet déjà rapporté pour E6*I dans la littérature, potentiellement causée par des effecteurs encore jusque-là non identifiés. Nous suggérons que la surproduction de ROS induite par E6*I d'HPV16 soit causée, au moins en partie, par l'activation du complexe de la NADPH oxydase 2.

SCIENTIFIC REPORTS

OPEN

Comparative RNA sequencing reveals that HPV16 E6 abrogates the effect of E6*I on ROS metabolism

Philippe Paget-Bailly^{1,2}, Koceila Meznad^{1,2}, Diane Bruyère⁴, Jérôme Perrard^{1,2}, Michael Herfs⁴, Alain C. Jung⁵, Christiane Mougin^{1,2,3}, Jean-Luc Prétet^{1,2,3} & Aurélie Baguet^{1,2}

High-risk Human Papillomavirus infections are responsible for anogenital and oropharyngeal cancers. Alternative splicing is an important mechanism controlling HPV16 gene expression. Modulation in the splice pattern leads to polycistronic HPV16 early transcripts encoding a full length E6 oncoprotein or truncated E6 proteins, commonly named E6*. Spliced E6*I transcripts are the most abundant RNAs produced in HPV-related cancers. To date, the biological function of the E6*I isoform remains controversial. In this study, we identified, by RNA sequencing, cellular targets deregulated by E6*I, among which genes related to ROS metabolism. Concomitantly, E6*I-overexpressing cells display high levels of ROS. However, co-overexpression of both E6 and E6*I has no effect on ROS production. In HPV16-infected cells expressing different E6/E6*I levels, we show that the newly identified targets CCL2 and RAC2 are increased by E6*I but decreased by E6 expression, suggesting that E6 abrogates the effect of E6*I. Taken together, these data support the idea that E6*I acts independently of E6 to increase ROS production and that E6 has the ability to counteract the effects of E6*I. This asks the question of how E6*I can be considered separately of E6 in the natural history of HPV16 infection.

Human papillomaviruses (HPVs) are small non-enveloped viruses that present a tropism for squamous epithelium. More than 200 types of HPVs have been described to infect humans¹. Based on their oncogenic potential, these viruses are classified in high-risk HPV (hrHPV), including HPV16 and HPV18, and low-risk HPV (lrHPV), including HPV6 and HPV11. HPV infections are responsible for cervical intraepithelial lesions that can progress to cancers, but they also cause a large fraction of anal, vulvar, vaginal, penile cancers, and a rising number of oropharyngeal cancers.

HPV16 is the most prevalent type in HPV-associated cancers². Its genome contains a long control region (LCR), 6 open reading frames (ORFs) encoding early (E) proteins under the control of p97 promoter located in the LCR, and 2 ORFs encoding late (L) proteins under the control of p670 promoter located within the E7 ORF³. Viral proteins are produced through the translation of at least 20 polycistronic transcripts obtained by alternative splicing. At least 10 of these transcripts allow the production of the 2 major viral oncoproteins, E6 and E7, but also E6-truncated proteins, E6*I, E6*II and E6^ΔE7^{4,5}. hrHPV E6 and E7 proteins are consistently expressed in HPV-associated cancers^{6,7} and interact with many host cellular proteins. Notably, E6 and E7 proteins target p53 and pRB, respectively, for proteasome-mediated degradation, and thus inactivate these tumor suppressors^{6,8,9}. More than 30 years ago, it was observed that the most abundant HPV16 transcript produced was spliced from the donor site 226 (SD 226) to the acceptor site 409 (SA409), both sites located in E6 ORF^{10,11}. Interestingly, only hrHPVs harbor these splice sites, indicating that E6 ORF splicing events could be relevant for HPV-driven carcinogenesis¹⁰. Several studies and unpublished data from our laboratory also reported increased levels of spliced E6*I mRNA correlating with the severity of cervical lesions^{5,12–14}. Concomitantly, it was proposed that the ratio

¹EA3181, LabEx LipSTIC ANR-11-LABX-0021, UFR Santé, 19 rue Ambroise Paré, Besançon, France. ²Université Bourgogne Franche Comté, Besançon, France. ³Centre Hospitalier Régional Universitaire, CNR HPV, 3 Bvd Alexandre Fleming, Besançon, France. ⁴Laboratory of Experimental Pathology, GIGA-Cancer, University of Liege, Liege, Belgium. ⁵Université de Strasbourg, Inserm, UMR_S1113, Centre de lutte contre le cancer Paul STRAUSS, Strasbourg, France. Correspondence and requests for materials should be addressed to A.B. (email: aurelie.baguet@univ-fcomte.fr)

of E6**I*/E6**II* spliced variants can be used as a predictive marker of clinical outcome in HPV-related cervical lesions¹³ and oropharyngeal cancers¹⁵.

Even if HPV16 early transcripts detection is used as a tool in screening and investigating HPV-related neoplasia, the biological significance of E6 splicing and ensuing E6**I* protein expression remains elusive. It has been proposed that E6 ORF splicing facilitates translation re-initiation of the E7 ORF by increasing the space between E6 and E7 ORF^{16,17}. However, other study showed that E7 is preferentially translated from the unspliced E6/E7 transcript rather than from the E6**I*/E7 one, suggesting that this splicing event regulates E6 expression but not E7^{18,19}. Apart from HPV gene expression regulation, the roles of E6**I* ORF product in HPV life cycle and carcinogenesis also remains unclear, although a variety of functions have been reported for the truncated isoform. E6**I* inhibits E6-mediated degradation of p53²⁰, causes the degradation of some PDZ proteins, such as Dlg, PATJ and MAGI-1^{21,22}, and modulates the expression of a subset of cellular factors involved in stress response, such as aldo-keto reductase genes²³, superoxide dismutase isoform 2 (SOD2), and glutathione peroxidase 1 (GPX1), leading to the accumulation of reactive oxygen species (ROS)²⁴. *In vivo* studies have shown that ectopic E6**I* reduces tumor growth of both HPV-positive and HPV-negative cells in a xenograft nude mouse model²⁵. Depending on the cellular context and especially the presence or absence of the E6 protein, E6**I* seems implicated in different cellular pathways and its functions remain controversial²⁶.

To get a better understanding on the underlying molecular mechanisms driving HPV-related carcinogenesis, the present study used RNA sequencing technology to analyze the impact of HPV16 E6**I* isoform on cellular gene expression.

Results

HPV16 E6 and E6I* expression in U-2 OS HPV-negative cell line.** HPV-negative U-2 OS cell line was transfected with expression vectors encoding either all HPV16 E6 isoforms (pXJ40-E6All or pXJ40b-GloΔint-E6All) or E6 (pXJ40-E6 or pXJ40bGloΔint-E6) or E6**I* (pXJ40-E6**I*) exclusively or with empty vectors (pXJ40 or pXJ40-bGloΔint) (Figs S1 and S2). We compared alternative splicing profiles of E6 transcripts obtained from these different E6-expression vectors. An accumulation of E6 unspliced transcripts produced from the pXJ40bGloΔint-E6All vectors is observed, indicating that the presence of the β-globin intron in the pXJ40 vector favor the splicing between SD226 and SA409 splice sites, and consequently the production of the E6**I* RNAs (Fig. S2). Western blot analysis were performed with two E6-specific antibodies. The 6F4 antibody recognizes the N-terminal part of the E6 protein, while the 3F8 antibody targets the C-terminal part of E6 (Fig. 1A). Consequently, the 3F8 antibody does not recognize the E6**I* protein and the 6F4 antibody recognizes both E6 isoforms. The full length E6 protein was detected with both E6-directed antibodies in cells transiently expressing the pXJ40-E6All and pXJ40-E6 vectors. No E6**I* protein was observed in pXJ40-E6All or pXJ40-E6**I* expressing cells. But consistently with its role as a potent E6 antagonist, we observed an increase of p53 protein level in pXJ40-E6**I* transfected cells. (Figs 1B and S3A).

MG-132 treatment, a proteasome inhibitor, was used on transfected cells to reveal a specific band with the 6F4 antibody that corresponds to the E6**I* isoform (Fig. 1C, lane 8 and Fig. S3B). This indicate that an E6**I* protein can be produced from the pXJ40-E6**I* vector but also suggest that low levels of E6**I* protein can be explain, at least in part, by its rapid turnover in these cells.

Establishment of stable cell lines expressing HPV16 E6 isoforms. Stable cell lines expressing either all E6 isoforms or E6**I* were generated from the U-2 OS cell line. Cellular clones were isolated upon selection of cells resistant to G418: U-CTL (clones #1 and #3, transfected with pXJ40-empty), U-E6All (clones #1, #6 and #7, transfected with pXJ40-E6All), and U-E6**I* (clones #1, #6 and #14, transfected with pXJ40-E6**I*). Then, these cellular clones were tested for E6 and E6**I* expression. As expected, E6 protein was observed in U-E6All stable cell lines compared to U-CTL and U-E6**I*. A decrease of the wild-type p53 protein level is observed in these cells, suggesting that U-E6All cells express of a functional E6 protein (Fig. 2A and S4A). Next, we compared E6 expression from U-E6All clones to naturally-infected HPV16 cell lines such as SiHa, Ca Ski and W12 clones. We observe a significantly higher expression of E6 in U-E6All clones. Interestingly, this high expression does not result in a greater degradation of p53 compared to SiHa and Ca Ski cells. We observe low amount of E6 in the three W12 clones tested but still managing variable levels of p53 between the three clones tested (Figs 2B and S4B). Because E6**I* protein was hardly detectable, E6 and E6**I* RNA levels were simultaneously monitored by RT-PCR. The RNA profile of U-E6All stable cell lines shows a low amount of full-length E6 RNA (unspliced) and E6**II* RNA (SD226/SA556 spliced RNA) in comparison with the high amount of E6**I* RNA (SD226/SA409 spliced RNA). As expected, U-E6**I* stable cell lines express only E6**I* RNA and U-CTL stable cell lines, like parental U-2 OS (U-CTL#par), did not express E6 or E6**I* (Figs 2C and S5A). Similar splicing profiles of E6 transcripts are detected between U-E6All stable cell lines, HPV16-positive cancer cell lines (SCC90, SiHa, Ca Ski) and in most head and neck squamous cell carcinoma (HNSCC) tumor samples tested (Figs 2D and S5B). Also, in the naturally HPV16-infected cervical keratinocyte cells (W12 cells), E6**I* transcript is the most abundant splice product compared to E6 and E6**II* transcripts (Figs 2E and S5C).

E6 and E6**I* expression differentially alters phenotypic characteristics relevant to cancer cells.

To evaluate phenotypic characteristics relevant to cancer cells in our cellular models, functional assays were performed with U-CTL, U-E6All and U-E6**I* cell lines. First, cell proliferation was assessed by measuring the number of cells once a day over 4 days. Results show that all cell lines had the same growth rate (Fig. 3A). In contrast, colony formation assays, performed to monitor the ability of a single cell to grow into a colony, show a significant decrease in the number of colonies for U-E6All compared with U-CTL. We observe that expression of E6**I* alone has a stronger inhibitory effect on colony formation (Fig. 3B,C). E6 and E6**I* from HPV16 have both been shown to induce oxidative stress in different cellular models^{24,27}. Relative production and release of H₂O₂ in medium

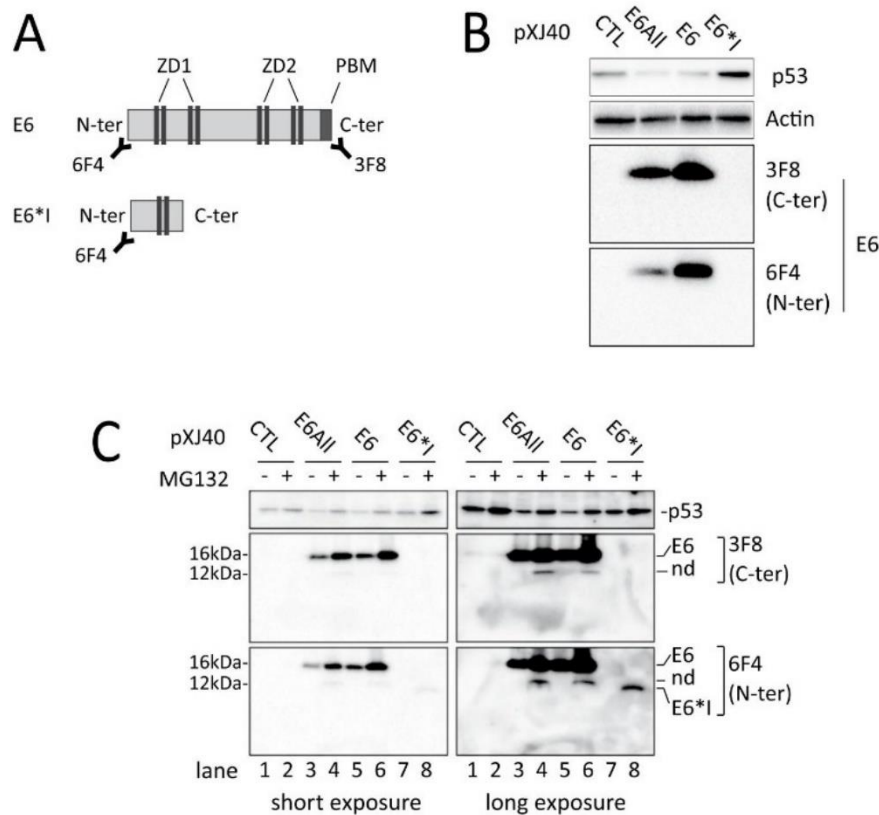


Figure 1. Transient transfection of E6 expression vectors in U-2 OS HPV-negative cell line. **(A)** Schematic representation of the E6 and E6*I protein isoforms with indicated regions recognized by the two E6-specific antibodies. The 6F4 antibody can recognize the N-terminal part of the E6 and E6*I proteins and the 3F8 antibody can target the C-terminal part of E6 protein. **(B)** Western blotting analysis showing p53, E6 and E6*I protein levels in U-2 OS cells transiently transfected with the pXJ40 empty vector (CTL), pXJ40-E6All vector (E6All), pXJ40-E6 vector (E6) or pXJ40-E6*I vector (E6*I). After 48-hour incubation, cells were treated either with DMSO or 10 μM MG132 proteasome inhibitor for 16 hours. The membrane was then incubated with the two E6 antibodies. For the detection of E6*I (corresponding to the first 43 amino acids of E6 protein), only the 6F4 antibody can be used. “nd” referred to a band at 11 kDa corresponding to a cleavage product observed in U-E6All cells treated with MG132. Full-length blots are presented in Figure S3.

from cells, reported by the oxidation of amplex red reagent to fluorescent resorufin, was used as indicator of ROS production. We observe a 20% significant increase in H₂O₂ released from U-E6*I clones relative to control. However, no change in H₂O₂ release from U-E6All clones is observed (Fig. 3D). It was previously observed that HPV16 E6 increases the radio sensitivity of HPV-negative cancer cells²⁸. To confirm this in our cellular models, the radiation sensitivity was measured *via* clonogenicity assays following irradiation. In all cell lines, survival decreases with increasing irradiation dose (Fig. 3E). We observed that only U-E6All cell lines show a significant decrease of survival after irradiation compared to U-CTL and U-E6*I cell lines (Fig. 3F). These data suggest that E6 and E6*I can have isoform-specific functions in the cells.

Transcriptome profiles reveal that E6*I deregulates cellular targets independently of E6. To assess cellular targets potentially deregulated by E6*I, the whole transcriptome of stable cell lines – U-E6All (clones #1, #6 and #7), U-E6*I (clones #1, #6 and #14) and U-CTL (clones #1, #3 and #par) – were analyzed by RNA sequencing. First, unsupervised hierarchical clustering analysis based on RNA sequencing datasets reveals 2 clusters. One of these clusters contains the 3 clones of U-E6All cells (Fig. 4A and S6). Venn diagram representation of deregulated genes in U-E6All and U-E6*I compared with U-CTL show that 419 transcripts were significantly deregulated (FDR < 0.05) in U-E6All (Fig. 4B), of which 264 were up-regulated and 155 were down-regulated (Fig. 4C and Table S1). The expression of 53 transcripts were deregulated in U-E6*I cell lines (Fig. 4B), of which 45 were up-regulated and 8 were down-regulated (Fig. 4C and Table S2). An overlap of 11 deregulated transcripts was observed between U-E6All and U-E6*I cell lines. These data suggest that E6*I modulates cellular gene expression independently of E6 but also indicate that the effect of E6*I on transcriptome is disrupted when co-expressed with E6.

GO terms and KEGG pathway analysis of deregulated transcripts in U-E6All and U-E6*I cells. Classification using GO (Gene Ontology) enrichment and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis showed an over-representation of U-E6All deregulated transcripts in enriched terms,

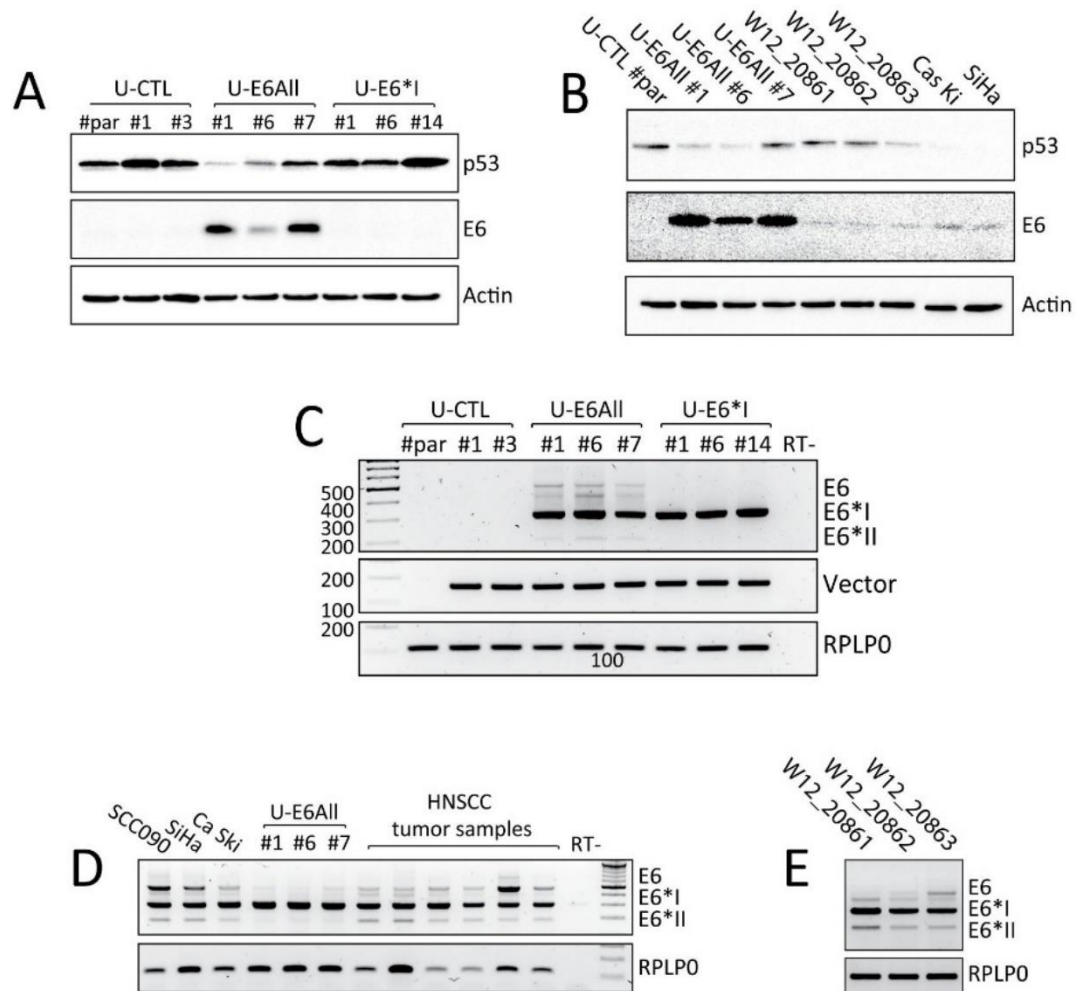


Figure 2. Expression profiles of HPV16 E6 and E6*I in stable cell lines. **(A)** Western blotting analysis showing p53, E6 and β -actin protein levels in U-2 OS parental cell line (U-CTL#par) and in stably transfected cells with pXJ40 empty vector (U-CTL #1 and #3) or pXJ40-E6All vector (U-E6All #1, #6 and #7) or pXJ40-E6*I vector (U-E6*I #1, #6 and #14). β -actin is used as loading control. Full-length blots are presented in Figure S4. **(B)** Western blotting analysis comparing protein levels of E6 and p53 between U-2 OS cellular clones (U-CTL#par, U-E6All #1, #6 and #7) and naturally HPV16-infected cells (W12_20861, W12_20862, W12_20863, SiHa and Ca Ski cell lines). β -actin is used as loading control. **(C)** RT-PCR analysis of alternatively spliced HPV16 E6 transcripts in U-E6All and U-E6*I stably transfected cells. As expected no E6 transcripts were observed in U-2 OS parental cell line (U-CTL#par) and in U-CTL #1 and #3 stably transfected cells. The detection of the vector in stably transfected cells was done by pXJ40 vector specific primers and RPLP0 was used as loading control. Full-length gels are presented in Figure S5. **(D)** RT-PCR analysis of alternatively spliced E6 transcripts in HPV16+ cell lines SCC090, SiHa, Ca Ski, in stable cell lines U-E6All #1, #6 and #7 and in HPV16+ HNSCC tumor samples. RPLP0 was used as loading control. Note that a band can be observed at 400 bp and could correspond to heteroduplexes between the different amplicons generated during PCR reaction. Full-length gels are presented in Figure S6. **(E)** RT-PCR showing alternatively spliced E6 transcripts in naturally HPV16-infected cells (W12_20861, W12_20862, W12_20863).

such as single-organism cellular process (GO: 0044763), protein-metabolic process (GO: 0019538), metabolic pathways (KEGG: hsa01100), ribosome (KEGG: hsa03010) and pathways in cancer (KEGG: hsa05200) (Fig. 5A). Based on protein-coding transcripts deregulated in U-E6All cells, a protein interaction network was constructed using the STRING database (Fig. 5B). Many of the deregulated transcripts identified in U-E6All cells have been previously identified in HPV-infected cells and tumor samples such as transcripts encoding ribosomal subunits, SYCP2, NUP210 and DICER1^{29–33}.

Gene ontology and KEGG pathway analysis of U-E6*I cell lines revealed 6 gene classes deregulated by ectopic expression of E6*I: Extracellular Matrix Organization (GO:0030198), Response to decreased oxygen levels (GO: 0036293), MAPK signaling pathway (KEGG: hsa04010), Focal adhesion (KEGG: hsa04510), Cytokine-cytokine receptor interaction (KEGG: hsa04060) and Pathway in cancer (KEGG: hsa05200) (Fig. 6A). In addition, the STRING database reported interactions between 10 protein-coding transcripts based on text mining, protein-protein interaction and co-expression data. Furthermore, these data show that the core interactors belong

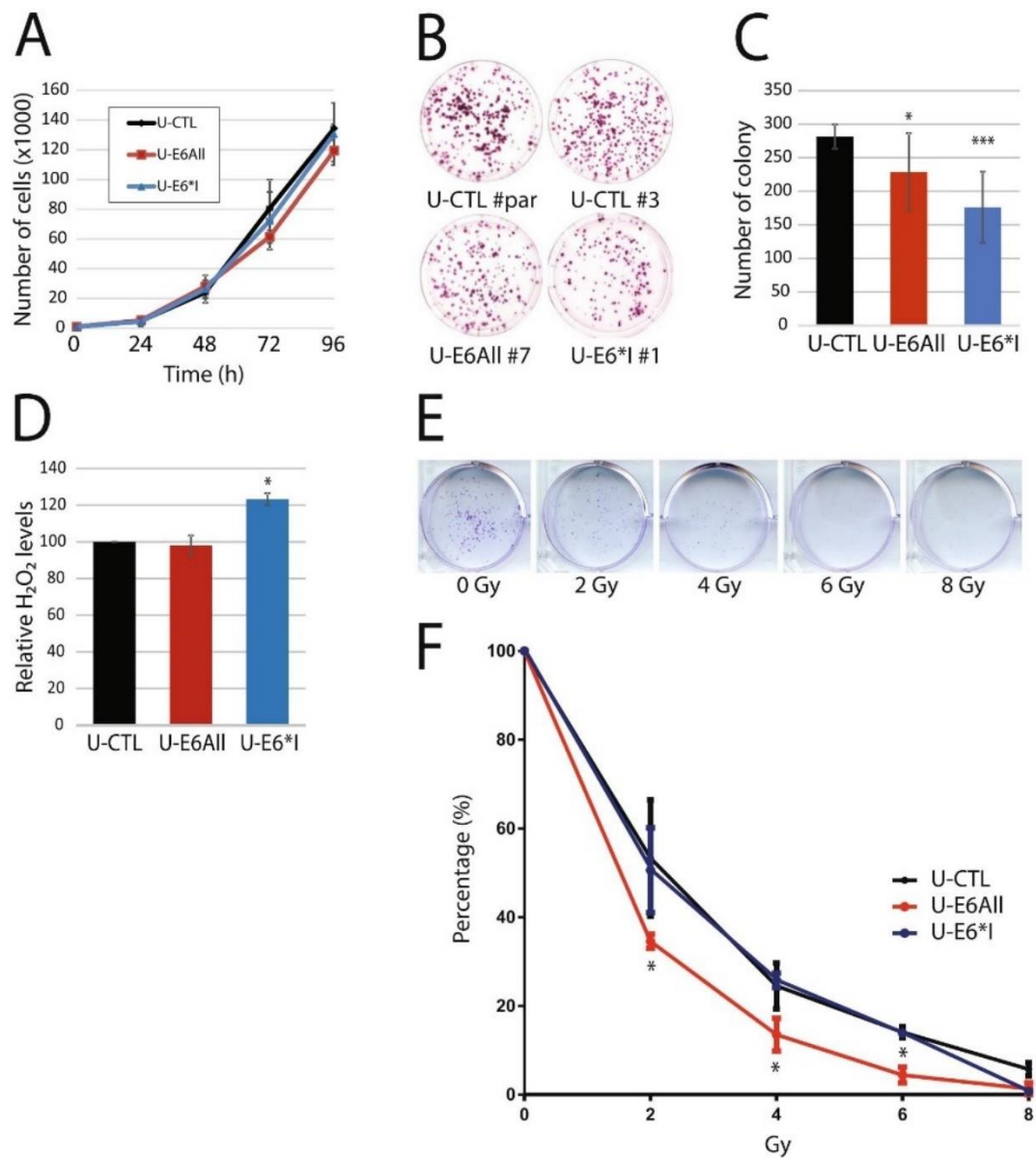


Figure 3. HPV16 E6 and E6*I expression differentially alters phenotypic characteristics relevant to cancer cells. **(A)** Cell proliferation assay with stably transfected cells U-CTL (black), U-E6All (red) and U-E6*I (blue) was performed to monitor daily live cells for 4 days. Cell counts are represented as mean \pm S.D. of 3 independent experiments. **(B)** Representative images of the colony formation assay realized with U-CTL#par, U-CTL#3, U-E6All#7 and U-E6*I#1. **(C)** Quantification of the colony formation assay realized with U-CTL (#par, #1, #3), U-E6All (#1, #6 and #7) and U-E6*I (#1, #6 and #14) cell lines. Foci counts are represented as mean \pm S.D. of 3 independent experiments. **(D)** Detection of H₂O₂ released from U-CTL (#par, #1, #3), U-E6All (#1, #6 and #7) and U-E6*I (#1, #6 and #14) cell lines in 1 h using the Amplex Red assay. H₂O₂ relative levels are represented as mean \pm S.D. of 3 independent experiments. **(E/F)** Clonogenic radio sensitivity assay performed with U-CTL, U-E6All and U-E6*I cell lines. Cell colonies (>50 cells) were counted 10 days following irradiation (0 to 8 Gy). Mann-Whitney test were used to assess statistical significance (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$).

to the extracellular matrix organization and/or oxidative stress response pathways (Fig. 6B). These data reveal specific pathways deregulated by E6*I in an E6-independent manner.

RNA sequencing transcripts validation by RT-qPCR. To validate RNA-seq data, expression of several transcripts was confirmed by RT-qPCR. We selected SYCP2, NUP210 and DICER1 transcripts deregulated in U-E6All cells and already known to be modulated by E6 expression (Fig. 7A), the 11 transcripts deregulated in both U-E6All and U-E6*I cells (Fig. 7B) and 13 transcripts only deregulated in U-E6*I cells including the 6 transcripts assigned to “response to decreased oxygen levels” (GO: 0036293) (Fig. 7C). Except for the lincRNA (long intergenic non-coding RNA) named CH507-210p18.1, fold changes estimated by RT-qPCR and RNA-sequencing are consistent for all transcripts tested. Pearson correlation coefficients (R) were calculated. We found high

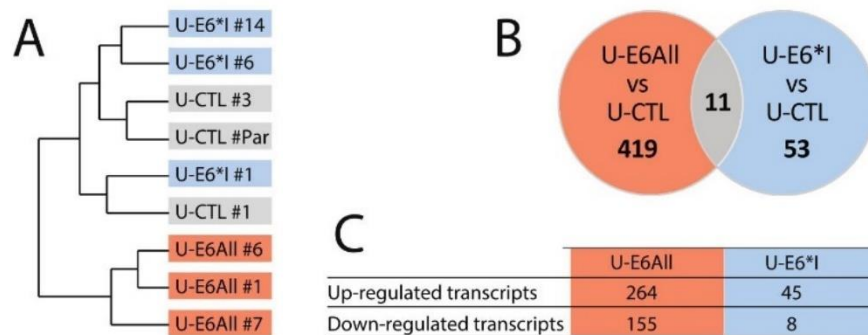


Figure 4. Transcriptome profiles reveal that E6*I deregulates cellular targets independently of E6. (A) Unsupervised hierarchical clustering of RNA-seq data from the U-CTL, U-E6All and U-E6*I cell lines. (B) Venn diagram showing the number of deregulated genes in the 2 pairwise comparisons U-E6All vs. U-CTL and U-E6*I vs. U-CTL. (C) Number of deregulated transcripts in U-E6All and U-E6*I cell lines compared to U-CTL.

correlations ($R = 0.996$ for U-E6All-deregulated transcripts, $R = 0.853$ for both U-E6All- and U-E6*I-deregulated transcripts and $R = 0.993$ for U-E6*I-deregulated transcripts) indicating a strong concordance between RNA-seq and RT-qPCR data.

Expression pattern of E6 or E6*I deregulated transcripts in W12 cell line. To validate the biological significance of the cellular transcripts found deregulated by E6*I, we performed experiments on W12 cells, a naturally HPV16-infected cervical keratinocyte cell line. First, absolute quantification was done to measure the copy number of E6 and E6*I mRNAs in three undifferentiated W12 clones (W12_20861, W12_20862, W12_20863). The W12_20861 clone expresses the highest level of E6*I transcript (approximately 4.10^6 copies/ μL) compared to W12_20862 and W12_20863 clones (approximately 500 000 copies/ μL) (Fig. 8A). Unspliced E6 transcripts are highly expressed in the W12_20863 clone (approximately 1.10^6 copies/ μL) compared to W12_20861 and W12_20862 clones (approximately 230 000 and 14 000 copies/ μL , respectively). In each W12 clone tested, we observed a strong correlation between RNA and protein levels for E6 with high level of E6 protein and low level of p53 protein in W12_20863 clone compared to W12_20861 and W12_20862. The high levels of E6*I mRNA and E7 protein in W12_20861 clone is also consistent with the lowest level of pRB protein (Figs 8B and S7). Overall, each clone possesses a unique expression profile of HPV transcripts and proteins allowing us to correlate those profiles to the expression of cellular targets identified by RNA-seq. So, the expression levels of transcripts modulated by E6 such as SYCP2, p21 and TNFAIP6 was measured by RT-qPCR in each W12 clone. As expected, SYCP2 and TNFAIP6 transcripts are upregulated while p21, a p53 target gene, is downregulated in the W12_20863 clone expressing the highest level of E6 (Fig. 8C).

Next, we monitored the RNA levels of CCL2, RAC2 and PDGFB transcripts modulated by E6*I. These transcripts encode proteins implicated in the activation of the NADPH Oxidase complex-2 (Nox2), required for Nox-dependent ROS production. There were higher levels of CCL2 and RAC2 in W12 clone expressing high level of E6*I (W12_20861). Nevertheless, a decreased expression of these transcripts is observed in cells expressing high level of E6 (W12_20863) suggesting that E6 could have the ability to counteract the effects of E6*I. Altogether, these results show that E6*I expression could drive the expression of genes implicated in oxidative stress metabolism in the W12 cell lines but also indicate that E6 can abrogate the effect of E6*I on ROS metabolism.

Discussion

To our knowledge, this study is the first report assessing the impact of HPV16 E6*I on cellular gene expression using RNA sequencing. The aim of this study was to uncover target genes specifically modulated by E6*I independently of the full length E6 protein. To this end, stable cell lines expressing either all E6 protein isoforms (U-E6All) or E6*I (U-E6*I) exclusively, were generated from U-2 OS. The U-2 OS cell line represents an adequate model for studying HPV gene expression and HPV-associated carcinogenesis because it expresses a wild-type p53, which can be targeted by E6. Furthermore, it was shown that viral gene expression of both alpha and beta papillomaviruses in U-2 OS is very similar to that described in keratinocytes³⁴. Immunoblotting experiments allowed to confirm the expression of a functional E6 by loss of p53 in U-E6All cell lines³⁵. E6 early transcripts were also efficiently spliced in U-E6All cell lines, yielding high amount of E6*I transcripts as observed in HPV16-positive cell lines and HNSCC clinical samples³⁶. However, although the E6*I protein was detected in transient transfection, no protein was observed in stable cell lines, probably due to the extremely short half-life of the E6*I isoform. This was already documented in a previous study, which reported the detection of an HPV16 E6*I fusion protein following transient transfection but failed to detect E6*I in stable cell lines³⁷. In the present study, we observe an accumulation of endogenous p53 protein in the HPV16 E6*I-expressing U-2 OS cells, suggesting that E6*I could have opposite effect of E6 on p53 protein level. But, we also show that E6*I expression causes an accumulation of ROS, indicating that the increasing level of p53 observed can be mediated by the ROS production. Also, emerging roles of non-coding RNAs in gene expression regulation and the noticeable difficulties to detect the E6*I protein by western blotting encouraged us to consider that the transcriptome deregulation

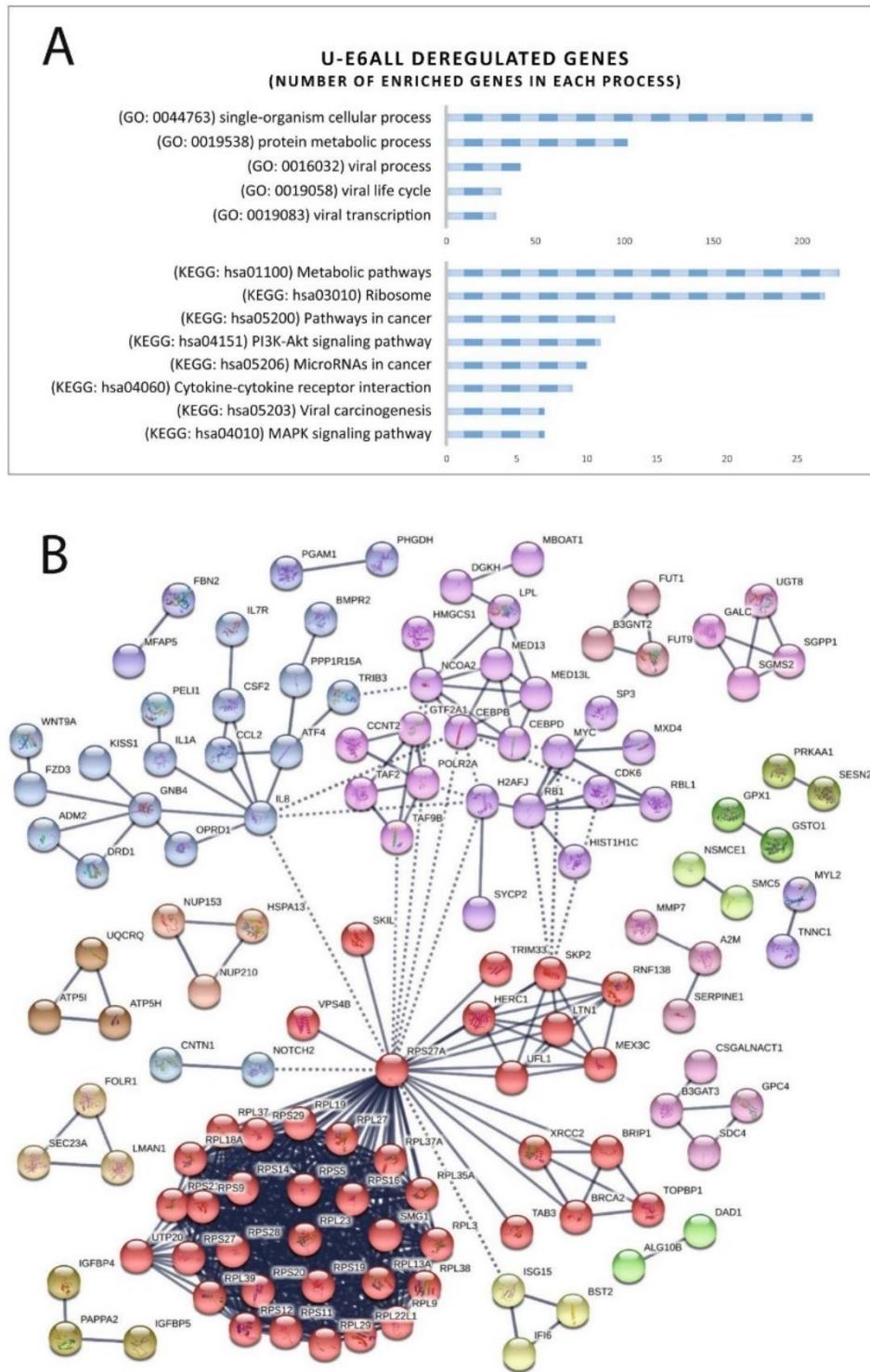


Figure 5. Enrichment and pathway analysis of deregulated genes in U-E6All cell lines compared with U-CTL. **(A)** Gene Ontology and KEGG pathway analysis of differentially expressed transcripts in U-E6All cell lines. **(B)** Protein-protein interaction network, corresponding to the protein-encoding genes deregulated in U-E6All cell lines, generated using STRING database⁷⁵. Line thickness indicates the strength of data support.

observed in the present study could be mostly mediated by HPV16 E6*I RNA rather than the E6*I protein³⁸. This information will raise exciting and challenging questions for understanding HPV-mediated carcinogenesis.

To assess the biological characteristics of the stable cell lines, we carried out several functional assays. We confirm that E6 isoforms or E6*I expression in HPV-negative cells had no impact on cell proliferation^{20,39}. A previous study⁴⁰ has observed that E6 expression inhibited colony formation. Here, we show that E6*I had a stronger inhibitory effect on colony formation than E6. Congruent with a previous study using U-2 OS cells transiently

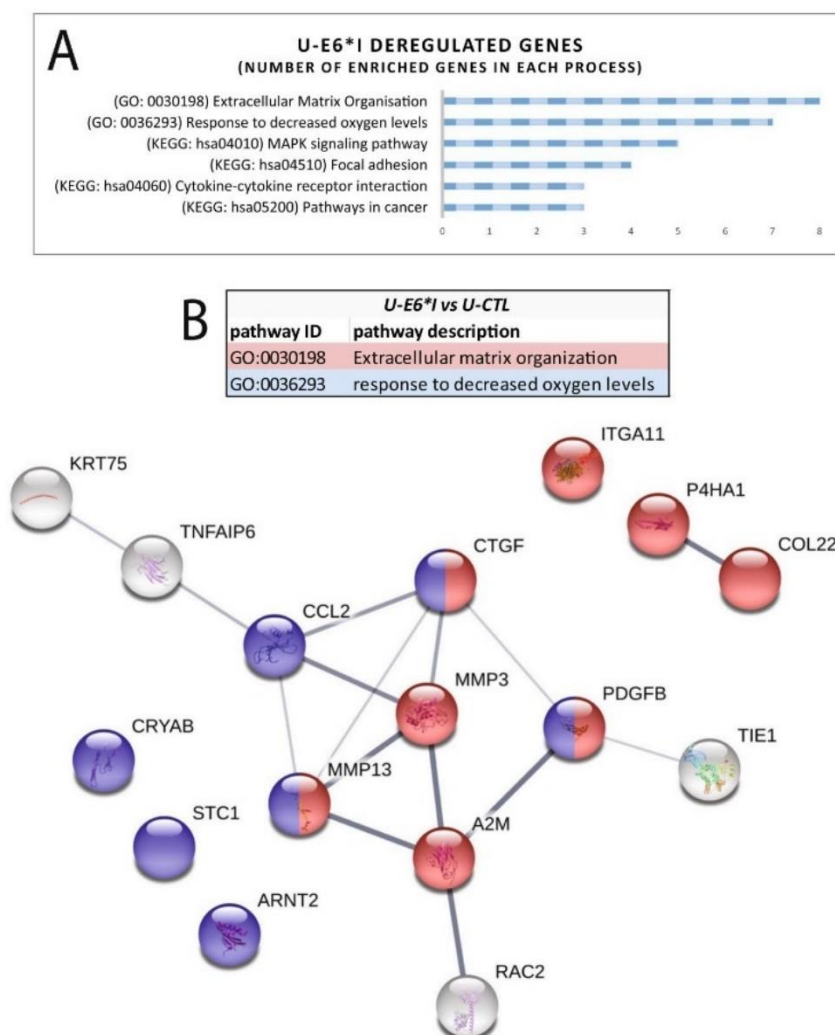


Figure 6. Enrichment and pathway analysis of deregulated genes in U-E6*I cell lines compared with U-CTL. **(A)** Gene Ontology and KEGG pathway analysis of differentially expressed transcripts in U-E6*I cell lines. **(B)** Protein-protein interaction network, corresponding to the protein-encoding genes deregulated in U-E6*I cell lines, generated using STRING database⁷⁵. Line thickness indicates the strength of data support. In blue and red are represented proteins involved in extracellular matrix organization and response to decreased oxygen levels, respectively.

expressing E6 isoforms or only E6*I²⁴, we report that stable expression of E6*I increases ROS production but no change is observed in cells co-expressing E6. Finally, radio sensitivity assays showed that U-E6All cells are more sensitive to radiation than both E6*I and control cells. These latter results could be partially explained by the *TP53* wild-type status of our cellular model given the fact that E6 was shown to disrupt the p53-mediated cellular responses to DNA damage⁴¹. Altogether, these data suggest that E6 isoforms have a specific impact on cellular phenotype and that the impact of E6*I is modulated when co-expressed with E6.

Recently, a whole transcriptome study using differentiated keratinocytes showed that HPV16 infection modulates genes involved in the regulation of cellular matrix adhesion, ROS biosynthesis and inflammatory response⁴². In our study, RNA sequencing identified 419 transcripts significantly deregulated in HPV-negative cellular model expressing E6 isoforms. Among them, a substantial number was already described in the literature as deregulated in HPV16 cellular models or in HPV16-positive clinical samples such as the down-regulation of a large set of transcripts encoding ribosomal subunits²⁹. Hence, KEGG pathway and GO enrichment analysis showed viral gene expression and protein synthesis related terms enriched. Our HPV-negative U-2 OS cellular model highlights the implication of E6 in the downregulation of ribosomal subunits transcripts. In addition, we confirm the upregulation of *SYCP2*, *Nup210* and *DICER1* transcripts, which were previously shown to be increased in HPV16-positive clinical samples from cervical²⁹ and head and neck cancers, and HPV-negative cell lines expressing E6 and E7 proteins^{30,33}.

Comparative RNA sequencing analysis also revealed that E6*I expression was sufficient to modulate the cellular transcriptome. KEGG pathway and GO enrichment annotations showed that E6*I-deregulated transcripts were assigned to extracellular matrix organization, response to decreased oxygen levels, mitogen activated protein

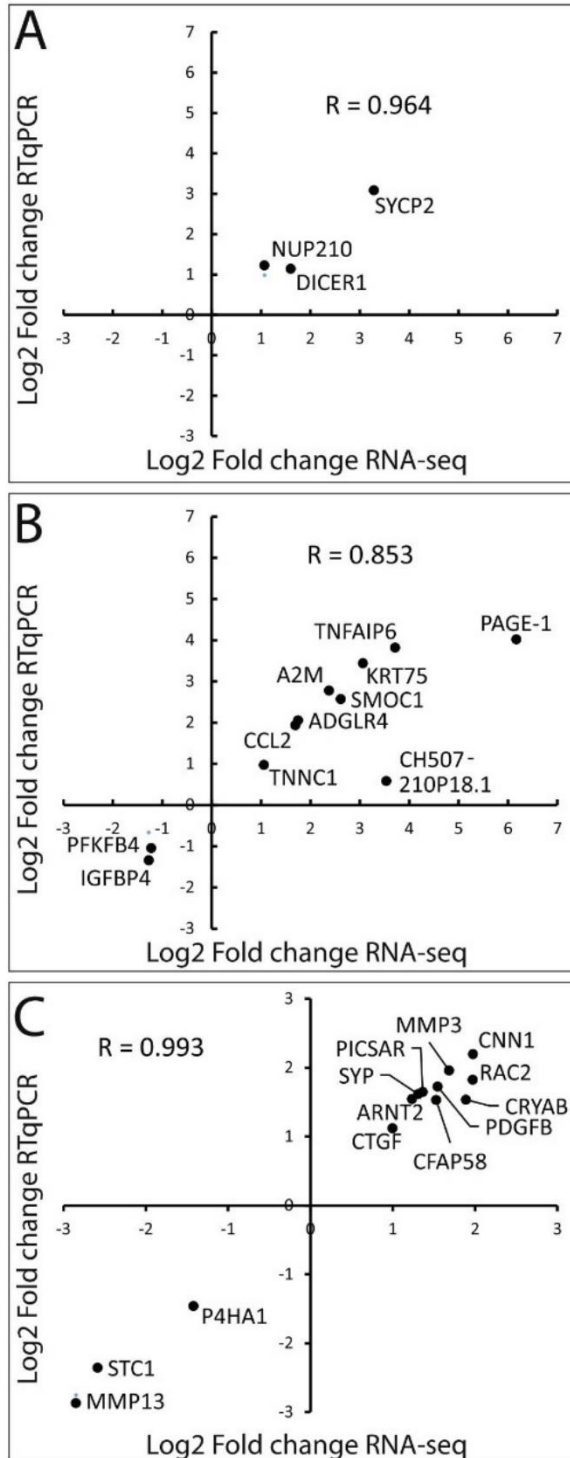


Figure 7. Correlation between RNA-sequencing and RT-qPCR results. RT-qPCR analyses of (A) 3 transcripts deregulated in U-E6All cell lines (U-CTL vs U-E6All), (B) the 11 transcripts deregulated in both U-E6All and U-E6*I cell lines (U-CTL vs U-E6*I) and (C) 13 transcripts deregulated in U-E6*I cell lines (U-CTL vs U-E6*I). Fold changes are represented in x- and y-axes for RNA-seq and RT-qPCR, respectively. For each condition, the data represent mean of the 3 clones. For RT-qPCR, RPLP0 was used as housekeeping gene.

kinase (MAPK), focal adhesion, cytokine-cytokine receptor interaction and cancer pathways. STRING analysis indicates that most transcripts implicated in extracellular matrix and response to decreased oxygen levels are predicted to form a significantly enriched core of interactors. A previous study, as well as ROS measurement in our model, indicate that E6*I increases levels of ROS²⁴. Thus, we focused downstream analyses on E6*I-deregulated

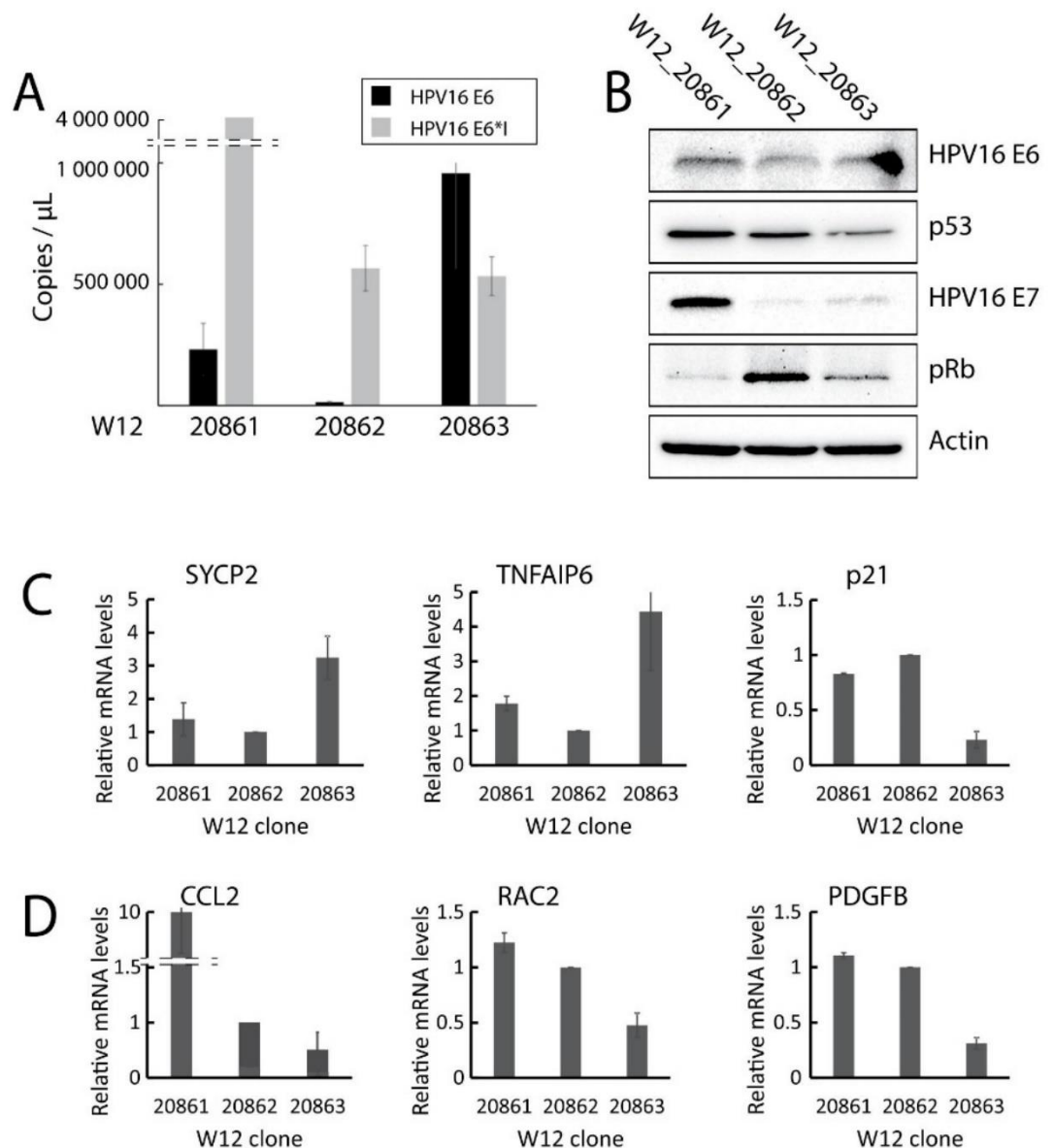


Figure 8. Expression levels of E6- and E6*I-deregulated transcripts in W12 clones. **(A)** RT-qPCR analysis showing E6 and E6*I absolute mRNA levels (copies per μL of pure cDNA) in W12_20861, W12_20862 and W12_20863 clones. pXJ40-E6 and pXJ40-E6*I diluted in salmon sperm DNA (50 ng/ μL) were used to generate standard curves. **(B)** Western blot analysis showing HPV16 E6 and E7, p53 and pRb protein levels in W12 clones. β -actin is used as loading control. RT-qPCR analyses showing. Full-length gels are presented in Figure S7. **(C)** SYCP2, TNFAIP6 and p21, and **(D)** CCL2, RAC2 and PDGFB relative mRNA expression in W12 clones. RPLP0 was used as housekeeping gene. RT-qPCR are represented as means \pm S.D. of 3 independent experiments.

transcripts implicated in the latter ontology, namely CRYAB⁴³, CCL2⁴⁴, CTGF⁴⁵, PDGFB^{46,47}, STC1⁴⁸, ARNT2⁴⁹, MMP13^{50,51}, MMP3^{52,53}, RAC2^{54,55} and TNFAIP6^{56,57} were also found associated with ROS metabolism. These E6*I-deregulated genes have been validated using RT-qPCR. The high correlations found demonstrate the robustness of fold changes estimated from RNA-seq.

To further investigate the effects of E6*I expression observed in the U-2 OS model, we validated the cellular target genes involved in ROS metabolism in the W12 HPV16-infected keratinocyte cell line. These cells, derived from a low-grade cervical lesion, has been extensively used for the study of HPV infection, life cycle and carcinogenesis. Many clones have been established harboring episomal or integrated type I and II HPV16 genomes to study all those aspects⁵⁸. Interestingly, each clone expresses different levels of E6 and E6*I mRNAs allowing us to get a better understanding on how E6*I potentially modulates the cellular targets identified by RNA-seq and on the existing interplay between E6 and E6*I. Quantification of HPV16 transcripts and proteins in W12 clones revealed a close correlation between mRNA levels and subsequent protein levels as it has been previously reported⁵⁹. TNFAIP6 was found upregulated in both U-E6All and U-E6*I cells by RNA-seq. Nevertheless,

TNFAIP6 RNA levels correlate with E6 expression in the W12. Finally, we confirm that RAC2, CCL2 and PDGFB mRNA levels correlate with E6*I expression in the W12 cellular model. RAC2 is a Rho-like small GTPase, implicated in the activation of the NADPH Oxidase complex-2 (Nox2), catalyzing the production of ROS. It has been shown that ROS production induced by Rac2 leads to DNA damage in leukemia stem cells causing genome instability suspected to drive the rapid accumulation of mutations in progenitor cells⁵⁴. Low levels of RAC2 have also been linked to lower ROS production and DNA damage in quiescent cells exposed to radiation⁵⁵. Interestingly, PDGF growth factor is also closely related to the Nox complex. PDGF interaction with its receptor was shown to activate ROS production by the Nox complex in hepatic⁴⁷ and epithelial model⁴⁶. Finally, CCL2 increases production of ROS by regulating gp91phox, a subunit of the Nox complex⁶⁰. CCL2 expression has also been shown to be reduced when the Nox complex is impaired and produce low levels of ROS⁶¹. The expression of RAC2, CCL2 and PDGFB induced by E6*I could drive Nox2 complex activation, thus representing a new pathway by which E6*I is able to increase cellular ROS production.

Oxidative stress is crucial in carcinogenesis because of its role in inflammation, mutagenesis, initiation and neoplastic progression⁶². Cellular antioxidant defenses exist to maintain cellular homeostasis and viruses are well known to induce oxidative stress by interfering with those defenses^{63,64}. Studies have linked HPV16 E6*I expression to changes in the activity of antioxidant enzymes, increased level of ROS and accumulation of DNA damage^{24,65,66}. In our study, RNA sequencing did not reveal any change in SOD and GPX mRNA levels when E6*I is expressed, probably due to a post-translational regulation of these enzymes by E6*I. In the context of hrHPV-driven carcinogenesis, it has been proposed that oxidative stress induced by E6*I could cause genome instability thus facilitating the integration of HPV genomes into the host cell genome^{67,68}. This hypothesis is consistent with the correlation between the severity of cervical lesions and the increasing levels of spliced E6*I mRNA^{5,12–14,26}.

In conclusion, our study reports that HPV16 E6*I induces modulation of the U-2 OS cellular transcriptome using RNA sequencing. We identified cellular genes involved in ROS metabolism that were deregulated by the ectopic expression of E6*I in our model and that are also found deregulated by endogenous E6*I expression in the W12 keratinocyte model. In agreement with previous studies, E6*I can promote viral genome integration into host genome⁶⁹ by inducing ROS and subsequently increasing genome instability²⁴. Here, we highlight the NADPH oxidase pathway as a new potential target of HPV16 E6*I that could explain the increased ROS production. Further studies will be needed to fully understand how oxidative stress related genes modulated by E6*I could affect HPV-driven carcinogenesis and life cycle. This study also demonstrate that E6*I-cellular transcriptome modulation is altered in the presence of E6 suggesting, among others things, a rivalry between the two protein isoforms. The overlap of only 11 transcripts between U-E6All and U-E6*I conditions suggests that E6 is able to partially negate the effects of E6*I. The modulation of ROS production in U-2 OS, and ROS-related genes expression in U-2 OS and W12 clones depending on E6/E6*I levels is in agreement with this interplay. Similarly, it has been observed that the effects of ectopic E6*I on antioxidant enzymes²⁴ seems to be lost when E6*I is co-expressed with E6⁷⁰. Other paradoxical observations were made concerning the expression of E6 and E6*I. The first aspect of this interplay was reported on the opposite effects of both isoforms on p53 degradation and subsequent apoptosis of E6-expressing cells²⁰. More recently, the protection of E6-expressing cells from TNF was only seen when E6 or E6*I are expressed alone but not when both isoforms are expressed³⁷. These observations ask the question of how E6*I can be considered separately of E6 in the natural history of HPV16 infection.

Materials and Methods

Cell culture. Human osteosarcoma derived cell line U-2 OS, HPV16-positive carcinoma derived cell lines SiHa (ATCC[®] HTB-35[™]), SCC090 (ATCC[®] CRL-3239[™]) and U-2 OS stable cell lines were cultured in DMEM medium (Lonza, Verviers, Belgium). U-2 OS stably transfected clones were routinely cultured in the presence of 625 µg/mL G418 to maintain selection. HPV16-positive carcinoma derived cell line Ca Ski (ATCC[®] CRL-1550[™]) was cultured in RPMI-1640 medium (Lonza, Verviers, Belgium). All media were supplemented with 10% Fetal Calf Serum (Eurobio, Courtaboeuf, France). Undifferentiated W12 20861, 20862 and 20863 clones were cultured in monolayer on Mitomycin C treated mouse 3T3 cells as previously described⁵⁸. Cell lines were incubated at 37 °C in a humidified atmosphere with 5% of CO₂.

Plasmids, transient and stable transfections. pXJ40 (empty) and pXJ40-E6All vectors, containing the HPV16 E6 coding sequence, have been generously provided by Murielle Masson (ESBS, Illkirch, France). The pXJ40βGloΔint vector was generated from the pXJ40 vector, using Q5Site-Directed Mutagenesis Kit (New England Biolabs) and the following primers 5'-CTCCTGGGCAACGTG-3' and 5'-CCTGAAGTTCTCAGGATCG-3'. The pXJ40-E6*I vector, lacking the intron 1 located in the E6 ORF corresponding to the nucleotides from 227 to 408 was generated from the pXJ40-E6All vector (Primers: 5'-GTGTATTAAGTGTCAAAGCC-3' and 5'-CTCACGTCGCAGTAACTG-3'). Transient transfections of the U-2 OS cell line were performed using JetPEI reagent (Polyplus-transfection, Illkirch, France) according to manufacturer's instructions. Cells were incubated in the presence or absence of 10 µM of proteasome inhibitor MG132 for 16 h and harvested after 48 h incubation post-transfection. U-2 OS stable cell lines were obtained using JetPEI reagent by co-transfecting with a ratio of 1:9 respectively, the pcDNA3.1-2Flag vector (neomycin resistance) and corresponding pXJ40 vectors (pXJ-CTL, pXJ40-E6All and pXJ40-E6*I). G418 resistant stable clones were then selected after 3 weeks under selection with 800 µg/mL G418 (Euromedex).

Cell proliferation and clonogenicity assays. Cell proliferation assays were done with 10,000 cells in 12-well plates by counting cells each day for 4 days using the EVE[™] automatic cell counter (Nano EnTek). Clonogenicity assays were performed by seeding 500 cells of each U-CTL, U-E6All and U-E6*I cell lines in 10 cm dishes. After 14 days of culture, colony were counted after crystal violet staining. Clonogenicity radio sensitivity

assays were performed by seeding 300 cells in 6-well plates 24 h before gamma irradiation (0 to 8 Gy) using Gammacell[®] 40 Exactor (Best Theratronics, Ltd). After treatment, cells were cultured for 10 days then stained with crystal violet. Cell colonies presenting >50 cells were finally counted. For each biological replicate (n = 3), 2 wells were analyzed for colonies counting.

Measurement of Hydrogen peroxide (H₂O₂). Measurement of H₂O₂ released from stably transfected U-2 OS cells was carried using the Amplex[®] Red Hydrogen Peroxide/Peroxidase Assay Kit (Invitrogen) following manufacturer instructions. Four hours after seeding 800,000 cells of each clone in 35 mm dishes, cells were incubated for 1 h in 500 μ L PBS1X. Then, 50 μ L was transferred in an opaque 96 well plate in triplicate and mixed with an equal amount of Amplex Red reagent (50 μ M Amplex Red and 0,1 U/mL HRP final concentrations). After 1 h incubation at room temperature in the dark, fluorescence was measured (560_{ex}/590_{em} nm) on a Synergy[™] H1 microplate reader (Biotek).

Immunoblot assays. U-2 OS, SiHa, Ca Ski or W12 cells were lysed using RIPA buffer (50 mM Tris-HCl pH7.4, 150 mM NaCl, 1% NP40, 0.5% Na deoxycholate, 1 mM EDTA) supplemented with 30 μ g/mL of anti-protease mixture (Roche Diagnostics) followed by sonication. After centrifugation at 10,000 \times g for 10 min at 4 °C, protein concentrations were determined in supernatant using Bio-Rad Protein Assay (Bio-Rad) according to manufacturer's instructions. Forty μ g of protein in Laemmli lysis buffer were then loaded onto a sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gel. After the transfer of proteins onto 0.2 μ m Amersham[™]Hybond[™]-PVDF membrane (GE Healthcare) and blocking 1 hour with 5% nonfat milk in Tris-buffered Saline-Tween20, membranes were probed with primary antibodies for 2 hours (except cell signaling antibodies on night) then with HRP-conjugated secondary antibodies (see below) for 1 h. Chemiluminescent signal was detected using PierceECL 2 Western Blotting Substrate (ThermoScientific).

Antibodies. Anti-HPV16E6 antibodies 1E-6F4 (E6 N-ter) and 2E-3F8 (E6 C-ter) were obtained from Euromedex. Anti-p53 (DO-1) and anti-pRb (4H1) antibodies were obtained from Cell Signaling. Anti- β -actin (AC-15) antibody was obtained from Sigma-Aldrich. Anti-HPV16E7 (NM2) antibody was obtained from Santa Cruz Biotechnology. HRP-conjugated goat anti-mouse secondary antibody was purchased from BD Pharmingen. Primary and secondary antibodies have been diluted respectively to 1:1000 and 1:5000 in blocking buffer (except for β -actin at 1:20,000 and 1:30,000).

RT-PCR and RT-qPCR. Total RNAs were extracted using Ribozol reagent (Amresco) and then 500 μ g were retro-transcribed into cDNA using Maxima first strand cDNA synthesis kit (ThermoScientific) according to manufacturer instructions. For PCR, cDNAs were amplified using DreamTaq polymerase (Thermo Scientific) according to manufacturer's instructions. PCR products were analyzed on a 2% (w/v) agarose gel prepared with Tris-Borate-EDTA buffer. Quantitative PCR was performed using Power SYBR Green PCR mix (Life Technologies) according to manufacturer instructions and analyzed on a StepOnePlus Real-Time PCR system (Applied Biosystems). Serial dilutions of pXJ40-E6 and pXJ40-E6*I from 10⁷ to 10¹ copies/ μ L diluted in salmon sperm DNA (50 ng/ μ L) (Invitrogen) were used as template to generate standard curve for absolute quantification of E6 and E6*I mRNA copies respectively in W12 clones. Primer sequences used are listed in Supplementary Table 3. RPLP0 expression levels was used as internal normalization standards. The 2^{- $\Delta\Delta$ Ct} method was used for relative mRNA quantification.

RNA sequencing. RNA-seq was carried out to identify differentially expressed transcripts in U-E6All or U-E6*I cell lines compared to U-CTL cell lines. For each clone, total RNAs were extracted using Ribozol reagent (Amresco) according to manufacturer's instructions. cDNA libraries were produced with 1 μ g of total RNA using TrueSeq Stranded mRNA Sample preparation kit (Illumina, Inc.). Sequencing was carried on 75 bp in single read on the NextSeq500 platform using the NextSeq500 High Output v2 kit (Illumina, Inc.).

Reads alignment. Raw reads quality was assessed using the fastQC software v0.11.3 (Babraham Institute, UK). Reads from each sample were then aligned to the reference human genome (GRCh38, release 87) using TopHat v2.1.1 (John Hopkins University, USA)⁷¹ with a minimum mapping quality score of 30. For each sample, an average of 45 million reads were uniquely mapped.

Differential expression analysis of cellular transcripts deregulated by E6 and/or E6*I. Cufflinks (Trapnell Laboratory, University of Washington, USA) was used for transcripts assembly and to assess transcripts abundance, which was normalized between all 9 samples using the fragments per kilobase per million mapped reads (FPKM) method to correct transcript length and library size bias⁷². Differentially expressed transcripts between control condition U-CTL and U-E6All or U-E6*I conditions were finally identified using Cuffdiff (Trapnell Laboratory). Transcripts with a false discovery rate (FDR) < 0,05 were considered significantly deregulated. No fold change cutoff was applied to the data. Nevertheless, more than 95% of transcripts with FDR < 0,05 had a fold change >2.

Accession numbers. RNA sequencing data in the form of BAM files have been submitted to SRA@ncbi.nih.gov under the BioProject PRJNA421890 and BioSamples SAMN08159315 (U-CTL #par); SAMN08159316 (U-CTL #1); SAMN08159317 (U-CTL #3); SAMN08159318 (U-E6All #1); SAMN08159319 (U-E6All #6); SAMN08159320 (U-E6All #7); SAMN08159321 (U-E6*I #1); SAMN08159323 (U-E6*I #6); SAMN08159324 (U-E6*I #14).

Functional analysis of deregulated genes. Enrichment analyses were performed using Gene Ontology (GO) database v1.2 on PANTHER v13.0⁷³, Kyoto Encyclopedia of Genes and Genomes (KEGG) release 81.0⁷⁴ and STRING v10.5⁷⁵.

Statistics. Statistical analyses were carried out using XLSTAT v2017.4 (Addinsoft, Paris, France). Pearson's correlation coefficients (r) were used to investigate correlation between RNA sequencing and RT-qPCR analyses. Mann-Whitney tests were used to analyze statistical significance of proliferation, clonogenicity and H₂O₂ assays, and RT-qPCR experiments (*p < 0.05; **p < 0.01; ***p < 0.001). Proliferation, clonogenicity, H₂O₂ and RT-qPCR graphs represent the mean of 3 experiments and the error bars represent the standard deviation.

References

- Schiffman, M. *et al.* The role of human papillomavirus (HPV) genotyping in cervical cancer screening: A large-scale evaluation of the cobas HPV test. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* **24**, 1304–1310 (2015).
- de Martel, C., Plummer, M., Vignat, J. & Franceschi, S. Worldwide burden of cancer attributable to HPV by site, country and HPV type. *Int. J. Cancer* **141**, 664–670 (2017).
- Graham, S. V. Keratinocyte Differentiation-Dependent Human Papillomavirus Gene Regulation. *Viruses* **9**, 245 (2017).
- Johansson, C. & Schwartz, S. Regulation of human papillomavirus gene expression by splicing and polyadenylation. *Nat. Rev. Microbiol.* **11**, 239–251 (2013).
- Chen, J. *et al.* Mapping of HPV transcripts in four human cervical lesions using RNAseq suggests quantitative rearrangements during carcinogenic progression. *Virology* **462**, 14–24 (2014).
- Moody, C. A. & Laimins, L. A. Human papillomavirus oncoproteins: pathways to transformation. *Nat. Rev. Cancer* **10**, 550 (2010).
- McLaughlin-Drubin, M. E. & Münger, K. The human papillomavirus E7 oncoprotein. *Virology* **384**, 335–344 (2009).
- Dyson, N., Howley, P. M., Münger, K. & Harlow, E. The Human Papilloma Virus–16 E7 Oncoprotein is Able to Bind to the Retinoblastoma Gene Product. *Science* **243**, 934–937 (1989).
- Werness, B. A., Levine, A. J. & Howley, P. M. Association of human papillomavirus types 16 and 18 E6 proteins with p53. *Science* **248**, 76–80 (1990).
- Smotkin, D., Prokoph, H. & Wettstein, F. O. Oncogenic and nononcogenic human genital papillomaviruses generate the E7 mRNA by different mechanisms. *J. Virol.* **63**, 1441–1447 (1989).
- Cornelissen, M. T. E. *et al.* Uniformity of the Splicing Pattern of the E6/E7 Transcripts in Human Papillomavirus Type 16-transformed Human Fibroblasts, Human Cervical Premalignant Lesions and Carcinomas. *J. Gen. Virol.* **71**, 1243–1246 (1990).
- Kösel, S., Burggraf, S., Engelhardt, W. & Olgemöller, B. Increased levels of HPV16 E6* I transcripts in high-grade cervical cytology and histology (CIN II+) detected by rapid real-time RT-PCR amplification. *Cytopathology* **18**, 290–299 (2007).
- Cricca, M. *et al.* Molecular analysis of HPV 16 E6I/E6II spliced mRNAs and correlation with the viral physical state and the grade of the cervical lesion. *J. Med. Virol.* **81**, 1276–1282 (2009).
- Lin, K. *et al.* E6-associated transcription patterns in human papilloma virus 16-positive cervical tissues. *Oncol. Lett.* **9**, 478–482 (2015).
- Hong, A. *et al.* E6 viral protein ratio correlates with outcomes in human papillomavirus related oropharyngeal cancer. *Cancer Biol. Ther.* **17**, 181–187 (2016).
- Tang, S., Tao, M., McCoy, J. P. & Zheng, Z.-M. The E7 Oncoprotein Is Translated from Spliced E6* I Transcripts in High-Risk Human Papillomavirus Type 16- or Type 18-Positive Cervical Cancer Cell Lines via Translation Reinitiation. *J. Virol.* **80**, 4249–4263 (2006).
- Zheng, Z.-M. Regulation of alternative RNA splicing by exon definition and exon sequences in viral and mammalian gene expression. *J. Biomed. Sci.* **11**, 278–294 (2004).
- 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. *J. Virol.* **74**, 7284–7297 (2000).
- del Moral-Hernández, O. *et al.* 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 (2010).
- Pim, D., Massimi, P. & Banks, L. Alternatively spliced HPV-18 E6* protein inhibits E6 mediated degradation of p53 and suppresses transformed cell growth. *Oncogene* **15**, 257–264 (1997).
- Storrs, C. H. & Silverstein, S. J. PAT1, a Tight Junction-Associated PDZ Protein, Is a Novel Degradation Target of High-Risk Human Papillomavirus E6 and the Alternatively Spliced Isoform 18 E6*. *J. Virol.* **81**, 4080–4090 (2007).
- Pim, D., Tomaić, V. & Banks, L. The Human Papillomavirus (HPV) E6* Proteins from High-Risk, Mucosal HPVs Can Direct Degradation of Cellular Proteins in the Absence of Full-Length E6 Protein. *J. Virol.* **83**, 9863–9874 (2009).
- Wanichwatanadecha, P., Sirisrimangkorn, S., Kaewprag, J. & Ponglikitmongkol, M. Transactivation activity of human papillomavirus type 16 E6* I on aldo-keto reductase genes enhances chemoresistance in cervical cancer cells. *J. Gen. Virol.* **93**, 1081–1092 (2012).
- Williams, V. M., Filippova, M., Filippov, V., Payne, K. J. & Duerksen-Hughes, P. Human Papillomavirus Type 16 E6* Induces Oxidative Stress and DNA Damage. *J. Virol.* **88**, 6751–6761 (2014).
- Filippova, M. *et al.* The small splice variant of HPV16 E6, E6N, reduces tumor formation in cervical carcinoma xenografts. *Virology* **450–451**, 153–164 (2014).
- Olmedo-Nieva, L., Muñoz-Bello, J. O., Contreras-Paredes, A. & Lizano, M. The Role of E6 Spliced Isoforms (E6*) in Human Papillomavirus-Induced Carcinogenesis. *Viruses* **10** (2018).
- Liu, Y., Tergaonkar, V., Krishna, S. & Androphy, E. J. Human Papillomavirus Type 16 E6-enhanced Susceptibility of I929 Cells to Tumor Necrosis Factor α Correlates with Increased Accumulation of Reactive Oxygen Species. *J. Biol. Chem.* **274**, 24819–24827 (1999).
- Shin, H.-J. *et al.* Human papillomavirus 16 E6 increases the radiosensitivity of p53-mutated cervical cancer cells, associated with up-regulation of aurora A. *Int. J. Radiat. Biol.* **86**, 769–779 (2010).
- den Boon, J. A. *et al.* Molecular transitions from papillomavirus infection to cervical precancer and cancer: Role of stromal estrogen receptor signaling. *Proc. Natl. Acad. Sci. USA* **112**, E3255–E3264 (2015).
- Pyeon, D. *et al.* Fundamental Differences in Cell Cycle Deregulation in Human Papillomavirus-Positive and Human Papillomavirus-Negative Head/Neck and Cervical Cancers. *Cancer Res.* **67**, 4605–4619 (2007).
- Masterson, L. *et al.* Deregulation of SYCP2 predicts early stage human papillomavirus-positive oropharyngeal carcinoma: A prospective whole transcriptome analysis. *Cancer Sci.* **106**, 1568–1575 (2015).
- Rajkumar, T. *et al.* Identification and validation of genes involved in cervical tumorigenesis. *BMC Cancer* **11**, 80 (2011).
- Harden, M. & Munger, K. Perturbation of DROSHA and DICER expression by Human Papillomavirus 16 Oncoproteins. *Virology* **507**, 192–198 (2017).
- Isok-Paas, H., Männik, A., Ustav, E. & Ustav, M. The transcription map of HPV11 in U2OS cells adequately reflects the initial and stable replication phases of the viral genome. *Virol. J.* **12**, 59 (2015).
- Scheffner, M., Werness, B. A., Huibregtse, J. M., Levine, A. J. & Howley, P. M. The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell* **63**, 1129–1136 (1990).

36. Jung, A. C. *et al.* Biological and clinical relevance of transcriptionally active human papillomavirus (HPV) infection in oropharynx squamous cell carcinoma. *Int. J. Cancer* **126**, 1882–1894 (2010).
37. Filippova, M. *et al.* Complexes of Human Papillomavirus Type 16 E6 Proteins Form Pseudo-Death-Inducing Signaling Complex Structures during Tumor Necrosis Factor-Mediated Apoptosis. *J. Virol.* **83**, 210–227 (2009).
38. Czech, B. & Hannon, G. J. Small RNA sorting: matchmaking for Argonautes. *Nat. Rev. Genet.* **12**, 19 (2011).
39. Sedman, S. A. *et al.* The full-length E6 protein of human papillomavirus type 16 has transforming and trans-activating activities and cooperates with E7 to immortalize keratinocytes in culture. *J. Virol.* **65**, 4860–4866 (1991).
40. Ristriani, T., Fournane, S., Orfanoudakis, G., Travé, G. & Masson, M. A single-codon mutation converts HPV16 E6 oncoprotein into a potential tumor suppressor, which induces p53-dependent senescence of HPV-positive HeLa cervical cancer cells. *Oncogene* **28**, 762–772 (2009).
41. Kessis, T. D. *et al.* Human papillomavirus 16 E6 expression disrupts the p53-mediated cellular response to DNA damage. *Proc. Natl. Acad. Sci.* **90**, 3988–3992 (1993).
42. Klymenko, T. *et al.* RNA-Seq Analysis of Differentiated Keratinocytes Reveals a Massive Response to Late Events during Human Papillomavirus 16 Infection, Including Loss of Epithelial Barrier Function. *J. Virol.* **91**, e01001–17 (2017).
43. van de Schootbrugge, C. *et al.* Effect of hypoxia on the expression of α B-crystallin in head and neck squamous cell carcinoma. *BMC Cancer* **14**, 252 (2014).
44. Li, X. *et al.* A CCL2/ROS autoregulation loop is critical for cancer-associated fibroblasts-enhanced tumor growth of oral squamous cell carcinoma. *Carcinogenesis* **35**, 1362–1370 (2014).
45. Murphy-Marshman, H. *et al.* Antioxidants and NOX1/NOX4 inhibition blocks TGF β 1-induced CCN2 and α -SMA expression in dermal and gingival fibroblasts. *PLoS ONE* **12** (2017).
46. Chen, K. C.-W., Zhou, Y., Zhang, W. & Lou, M. F. Control of PDGF-induced reactive oxygen species (ROS) generation and signal transduction in human lens epithelial cells. *Mol. Vis.* **13**, 374–387 (2007).
47. Adachi, T. *et al.* NAD(P)H oxidase plays a crucial role in PDGF-induced proliferation of hepatic stellate cells. *Hepatology* **41**, 1272–1281 (2005).
48. Ohkouchi, S., Kikuchi, T., Ebina, M. & Nukiwa, T. Novel functions of stanniocalcin-1 (STC1) through uncoupling protein 2 (UCP2) up-regulation; promoting survival of cancer cells under oxidative stress and inducing the uncoupling respiration (Warburg effect). *Eur. Respir. J.* **38**, p1490 (2011).
49. Keith, B., Adelman, D. M. & Simon, M. C. Targeted mutation of the murine arylhydrocarbon receptor nuclear translocator 2 (Arnt2) gene reveals partial redundancy with Arnt. *Proc. Natl. Acad. Sci. USA* **98**, 6692–6697 (2001).
50. Waldron, A. L. *et al.* Oxidative stress-dependent MMP-13 activity underlies glucose neurotoxicity. *J. Diabetes Complications* **32**, 249–257 (2018).
51. Del Carlo, M., Schwartz, D., Erickson, E. A. & Loeser, R. F. The Endogenous Production of Reactive Oxygen Species Is Required For Stimulation of Human Articular Chondrocyte MMP Production by Fibronectin Fragments. *Free Radic. Biol. Med.* **42**, 1350–1358 (2007).
52. Radisky, D. C. *et al.* Rac1b and reactive oxygen species mediate MMP-3-induced EMT and genomic instability. *Nature* **436**, 123–127 (2005).
53. Choi, D.-H. *et al.* Matrix Metalloproteinase-3 Causes Dopaminergic Neuronal Death through Nox1-Regenerated Oxidative Stress. *PLoS ONE* **9** (2014).
54. Nieborowska-Skorska, M. *et al.* Rac2-MRC-cIII-generated ROS cause genomic instability in chronic myeloid leukemia stem cells and primitive progenitors. *Blood* **119**, 4253–4263 (2012).
55. Pei, H. *et al.* RAC2-P38 MAPK-dependent NADPH oxidase activity is associated with the resistance of quiescent cells to ionizing radiation. *Cell Cycle* **16**, 113–122 (2017).
56. Watanabe, R. *et al.* Emerging Roles of Tumor Necrosis Factor-Stimulated Gene-6 in the Pathophysiology and Treatment of Atherosclerosis. *Int. J. Mol. Sci.* **19** (2018).
57. Shkolnik, K. *et al.* Reactive oxygen species are indispensable in ovulation. *Proc. Natl. Acad. Sci. USA* **108**, 1462–1467 (2011).
58. Jeon, S., Allen-Hoffmann, B. L. & Lambert, P. F. Integration of human papillomavirus type 16 into the human genome correlates with a selective growth advantage of cells. *J. Virol.* **69**, 2989–2997 (1995).
59. Scarpini, C. G., Groves, I. J., Pett, M. R., Ward, D. & Coleman, N. Virus transcript levels and cell growth rates after naturally occurring HPV16 integration events in basal cervical keratinocytes. *J. Pathol.* **233**, 281–293 (2014).
60. Chun, E. *et al.* CCL2 Promotes Colorectal Carcinogenesis by Enhancing Polymorphonuclear Myeloid-Derived Suppressor Cell Population and Function. *Cell Rep.* **12**, 244–257 (2015).
61. Chen, X.-L., Zhang, Q., Zhao, R. & Medford, R. M. Superoxide, H₂O₂, and iron are required for TNF- α -induced MCP-1 gene expression in endothelial cells: role of Rac1 and NADPH oxidase. *Am. J. Physiol.-Heart Circ. Physiol.* **286**, H1001–H1007 (2004).
62. Reuter, S., Gupta, S. C., Chaturvedi, M. M. & Aggarwal, B. B. Oxidative stress, inflammation, and cancer: how are they linked? *Free Radic. Biol. Med.* **49**, 1603–1616 (2010).
63. Paracha, U. Z. *et al.* Oxidative stress and hepatitis C virus. *Virol. J.* **10**, 251 (2013).
64. Reshi, M. L., Su, Y.-C. & Hong, J.-R. RNA Viruses: ROS-Mediated Cell Death. *International Journal of Cell Biology*, <https://doi.org/10.1155/2014/467452> (2014).
65. Mouret, S. *et al.* E6* oncoprotein expression of human papillomavirus type-16 determines different ultraviolet sensitivity related to glutathione and glutathione peroxidase antioxidant defence. *Exp. Dermatol.* **14**, 401–410 (2005).
66. Evans, W. *et al.* Overexpression of HPV16 E6* Alters β -Integrin and Mitochondrial Dysfunction Pathways in Cervical Cancer Cells. *Cancer Genomics. Proteomics* **13**, 259–273 (2016).
67. Williams, V. M., Filippova, M., Soto, U. & Duerksen-Hughes, P. J. HPV-DNA integration and carcinogenesis: putative roles for inflammation and oxidative stress. *Future Virol.* **6**, 45–57 (2010).
68. Camini, F. C., da Silva Caetano, C. C., Almeida, L. T. & de Brito Magalhães, C. L. Implications of oxidative stress on viral pathogenesis. *Arch. Virol.* **162**, 907–917 (2017).
69. Chen Wongworawat, Y., Filippova, M., Williams, V. M., Filippov, V. & Duerksen-Hughes, P. J. Chronic oxidative stress increases the integration frequency of foreign DNA and human papillomavirus 16 in human keratinocytes. *Am. J. Cancer Res.* **6**, 764–780 (2016).
70. Cruz-Gregorio, A. *et al.* Human Papillomavirus Types 16 and 18 Early-expressed Proteins Differentially Modulate the Cellular Redox State and DNA Damage. *Int. J. Biol. Sci.* **14**, 21–35 (2018).
71. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
72. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5**, 621–628 (2008).
73. Mi, H. *et al.* PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res.* **45**, D183–D189 (2017).
74. Ogata, H. *et al.* KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **27**, 29–34 (1999).
75. Szklarczyk, D. *et al.* The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* **45**, D362–D368 (2017).

Acknowledgements

We thank Muriel Masson for kindly providing the pXJ40-E6All vector, Julie Frejaville and Anne Peigney for technical assistance, Gilles Travé for helpful advice and discussions and Nathalie Dourdin for carefully reading the manuscript. We also wish to thank Paul F. Lambert for kindly providing W12 cell clones. K.M. and P.P. had fellowships from the French ministry of research and technology. J.P. had fellowships from the Conseil Régional de Franche-Comté. This work was supported by Cancéropôle Est (CE), Ligue Contre le Cancer (Comité du Doubs and CCIR-GE), Université de Franche-Comté (UFC) and Région Franche-Comté.

Author Contributions

A.B. conceived the project. P.P.B. performed the bioinformatics analysis, the biochemical assays and the functional assays, K.M. realized the cloning and established the stable cell lines. D.B. performed radio sensitivity tests, A.J. provided the clinical samples, P.P.B. K.M., D.B., J.P., M.H., C.M., J.L.P. and A.B. analyzed the data. P.P.B. and A.B. wrote the paper.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-019-42393-6>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019

Supplementary Information

Comparative RNA sequencing reveals that HPV16 E6 abrogates the effect of E6*I on ROS metabolism

Philippe Paget-Bailly^{1,2}, Koceila Meznad^{1,2}, Diane Bruyère⁴, Jérôme Perrard^{1,2}, Michael Herfs⁴, Alain Jung⁵, Christiane Mougin^{1,2}, Jean-Luc Prétet^{1,2,3}, Aurélie Baguet^{1,2*}

¹ EA3181, LabEx LipSTIC ANR-11-LABX-0021, UFR Santé, 19 rue Ambroise Paré, Besançon.

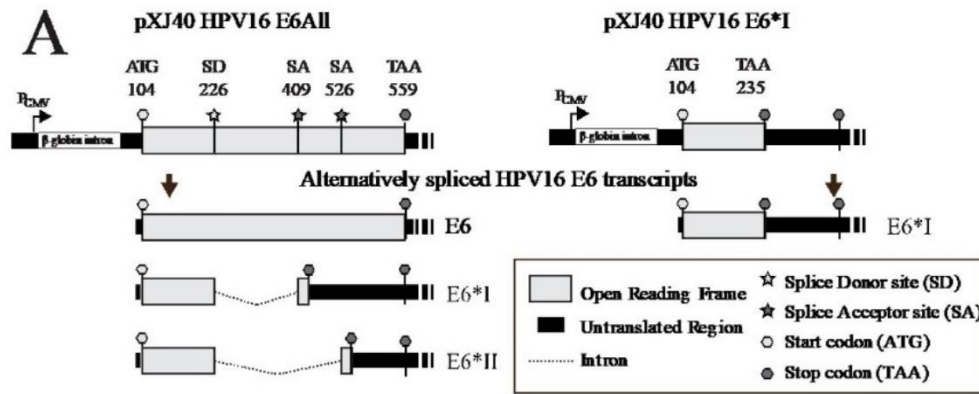
² Université Bourgogne Franche Comté, France.

³ Centre Hospitalier Régional Universitaire, CNR HPV, 3 Bvd Alexandre Fleming, Besançon, France.

⁴ Laboratory of Experimental Pathology, GIGA-Cancer, University of Liege, Liege, Belgium.

⁵ Université de Strasbourg, Inserm, UMR_S1113, Centre de lutte contre le cancer Paul STRAUSS, Strasbourg, France.

*Corresponding author : Aurélie Baguet, UFR Santé, EA 3181 - Carcinogenèse associée aux HPV, 19 rue Ambroise Paré, 25000 Besançon, France. Tel : +33 3 63 08 22 65. E-mail : aurelie.baguet@univ-fcomte.fr



B

E6 104 **ATG**TTTCAGGACCCACAGGAGCGACCCAGAAAGTTACCACAGTTATGCACAGAGCTGCAA
 1 M F Q D P Q E R P R K L P Q L C T E L Q

ACAAC TATACATGATATAATATTAGAATGTGTGTACTGCAAGCAACAGTTACTGCGAGCT
 T T I H D I I L E C V Y C K Q Q L L R R

GAGGTATATGACTTTGCTTTTCGGGATTTATGCATAGTATATAGAGATGGGAATCCATAT
 E V Y D F A F R D L C I V Y R D G N P Y

GCTGTATGTGATAAATGTTTAAAGTTTATTCTAAAATTAGTGAGTATAGACATTATTGT
 A V C D K C L K F Y S K I S E Y R H Y C

TATAGTTTGTATGGAACAACATTAGAACAGCAATACAACAACCCGTTGTGTGATTTGTTA
 Y S L Y G T T L E Q Q Y N K P L C D L L

ATTAGTGTATTAACTGTCAAAGCCACTGTGTCTGAAGAAAAGCAAAGACATCTGGAC
 I R C I N C Q K P L C P E E K Q R H L D

AAAAAGCAAAGATTCCATAATATAAGGGGTCGGTGGACCGGTCGATGTATGTCTTGTTC
 K K Q R F H N I R G R W T G R C M S C C

AGATCATCAAGAACACGTAGAGAAACCCAGCTGTAA 559
 R S S R T R R E T Q L * 151

C

E6*I 104 ATGTTTCAGGACCCACAGGAGCGACCCAGAAAGTTACCACAGTTATGCACAGAGCTGCAA
 1 M F Q D P Q E R P R K L P Q L C T E L Q

ACAAC TATACATGATATAATATTAGAATGTGTGTACTGCAAGCAACAGTTACTGCGAGCT
 T T I H D I I L E C V Y C K Q Q L L R R

GAGGTGAT**TAA**ctgtcaaaagccactgtgtcctgaagaaaagcaaaagacatctggacaa
 E V Y * 43

aaagcaaaagattccataatataaggggctcggaggaccggtcgatgtatgtcttgttgacag

atcatcaagaacacgtagagaaacccagctgtaa

Figure S1. HPV16 E6 cDNA sequences cloned into pXJ40 vector and their encoding capacity. (A) Schematic representation of HPV16 E6 and E6*I constructs used in our study. The expression of HPV16 E6 vector can generate by alternative splicing E6, E6*I and E6*II transcripts (left panel). The expression of HPV16 E6*I vector can only produce E6*I transcripts (right panel). (B) E6 coding sequence cloned into pXJ40. (C) E6*I coding sequence cloned into pXJ40.

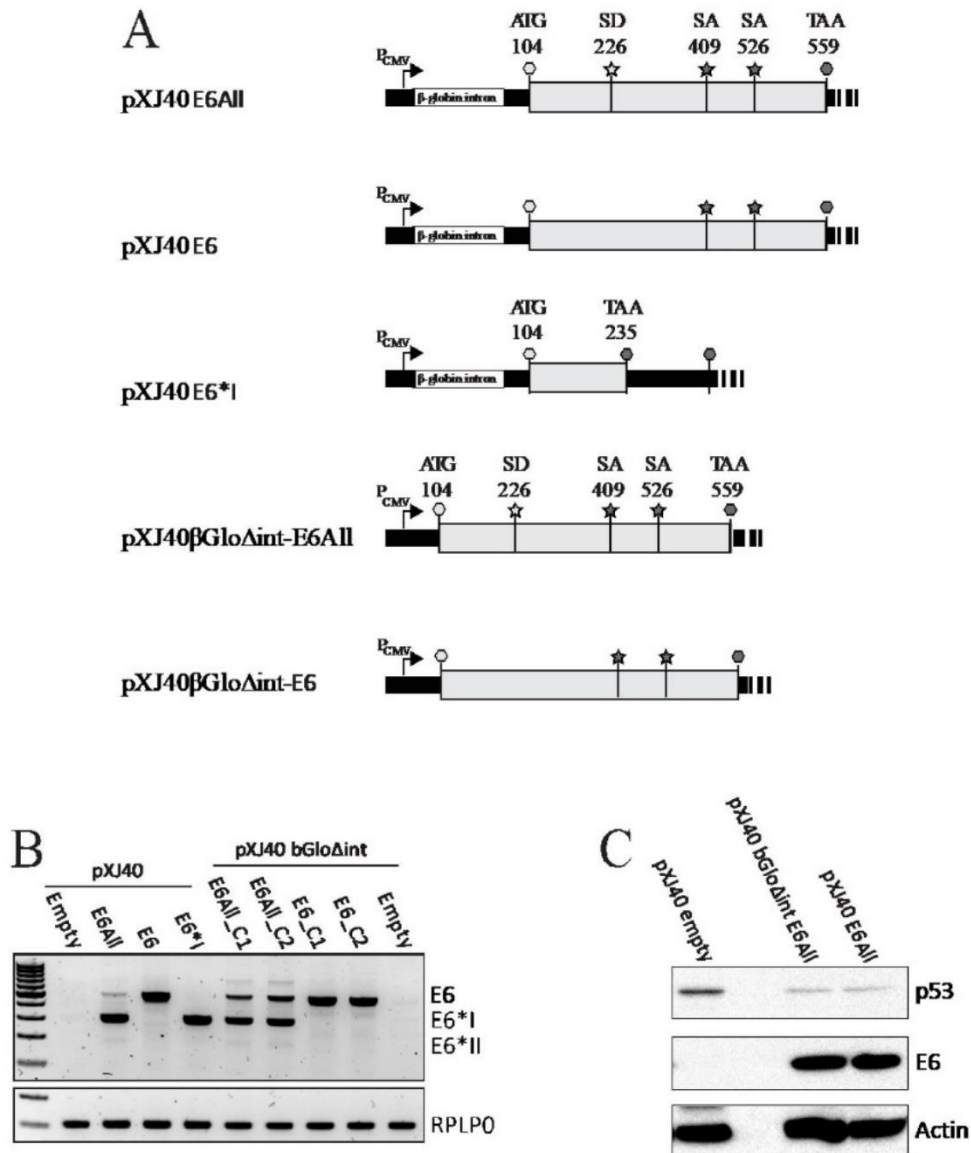


Figure S2. HPV16 E6 and E6*I sequences cloned into pXJ40 or pXJ40 β Glo Δ int vectors. (A) Schematic representation of HPV16 E6 and E6*I constructs used in this study. (B) RT-PCR analysis of alternatively spliced E6 transcripts generated from the pXJ40E6All, pXJ40E6, pXJ40E6*I, pXJ40 β Glo Δ intE6All (clones C1 and C2) and pXJ40 β Glo Δ intE6 (clones C1 and C2) vectors after transient transfection in U-2 OS cell line. pXJ40 β Glo Δ int transfections display a less efficient splicing of E6 transcripts but without altering the overall transcription level of E6 orf. RPLP0 was used as loading control. (C) Western blot analysis showing no change of E6 protein expression in U-2 OS cells transiently transfected with the pXJ40-E6All and pXJ40 β Glo Δ int-E6All vectors.

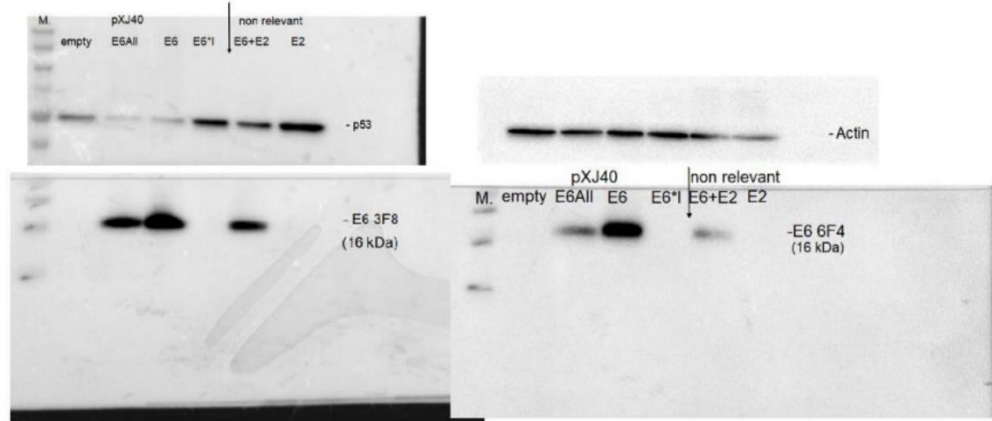
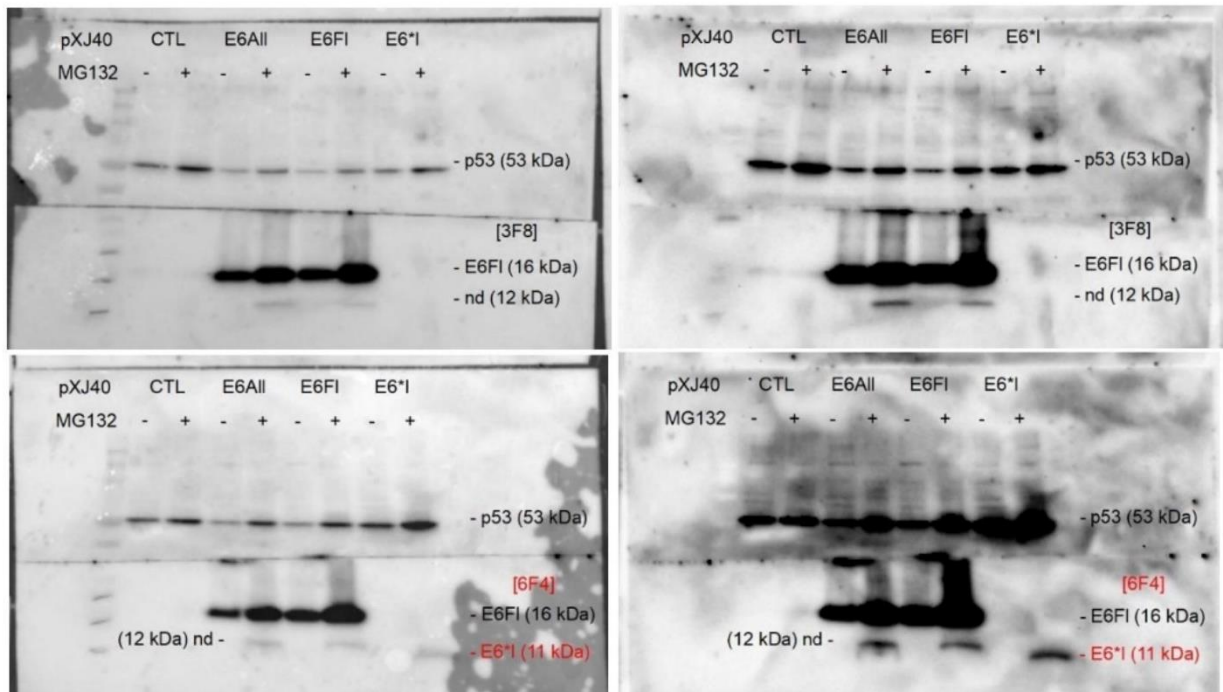
A**B**

Figure S3. Uncropped Western blot used in figure 1. (A) Figure 1B. E6 (3F8 or 6F4) and p53 hybridizations were performed on the same membrane. (B) Figure 1C. E6 (3F8 or 6F4) and p53 hybridizations were performed on the same membrane. The right panel represent a longer time of exposure compared to the left one. The upper blot of E6 was performed with the 3F8 antibody which does not recognize E6*I, where the lower one was performed with the 6F4 which does ¹. The Precision Plus Protein dual Xtra Standards (Bio-rad) was used as marker for all protein gel migration.

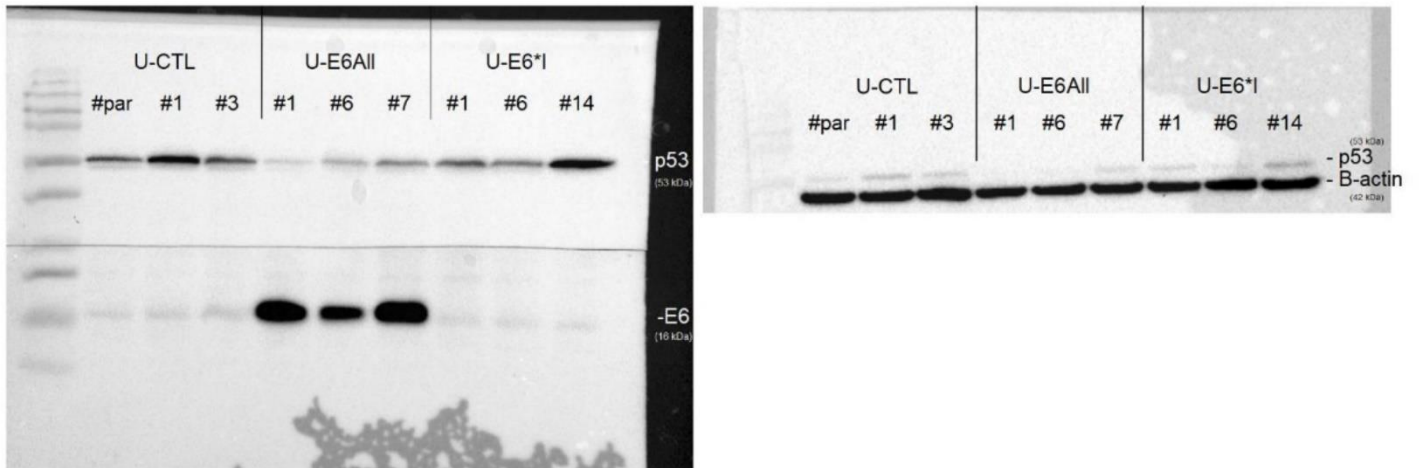
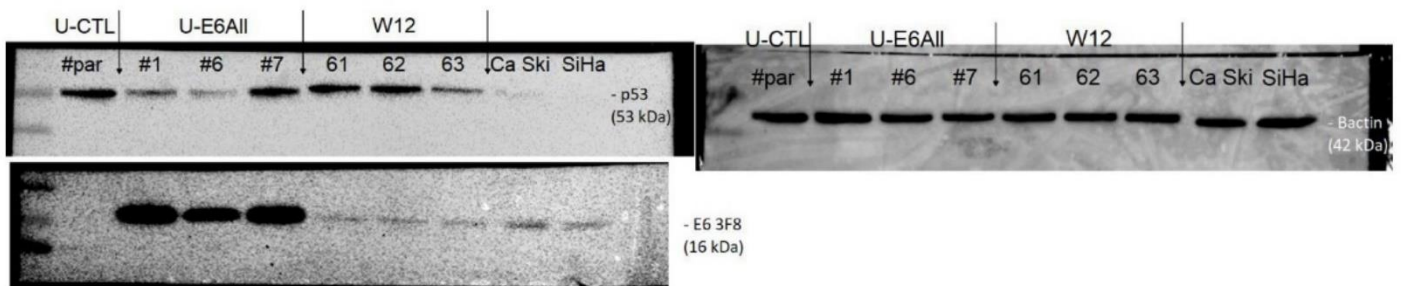
A**B**

Figure S4. Uncropped Western blot used in figure 2. (A) Figure 2A. B-actin antibody² hybridization and revelation were performed after p53³ and E6 (3F8), on the same membrane. (B) Figure 2B. All revelations were also done on the same membrane but with a longer time of exposure for E6, and p53 membrane piece was stripped before actin hybridization.

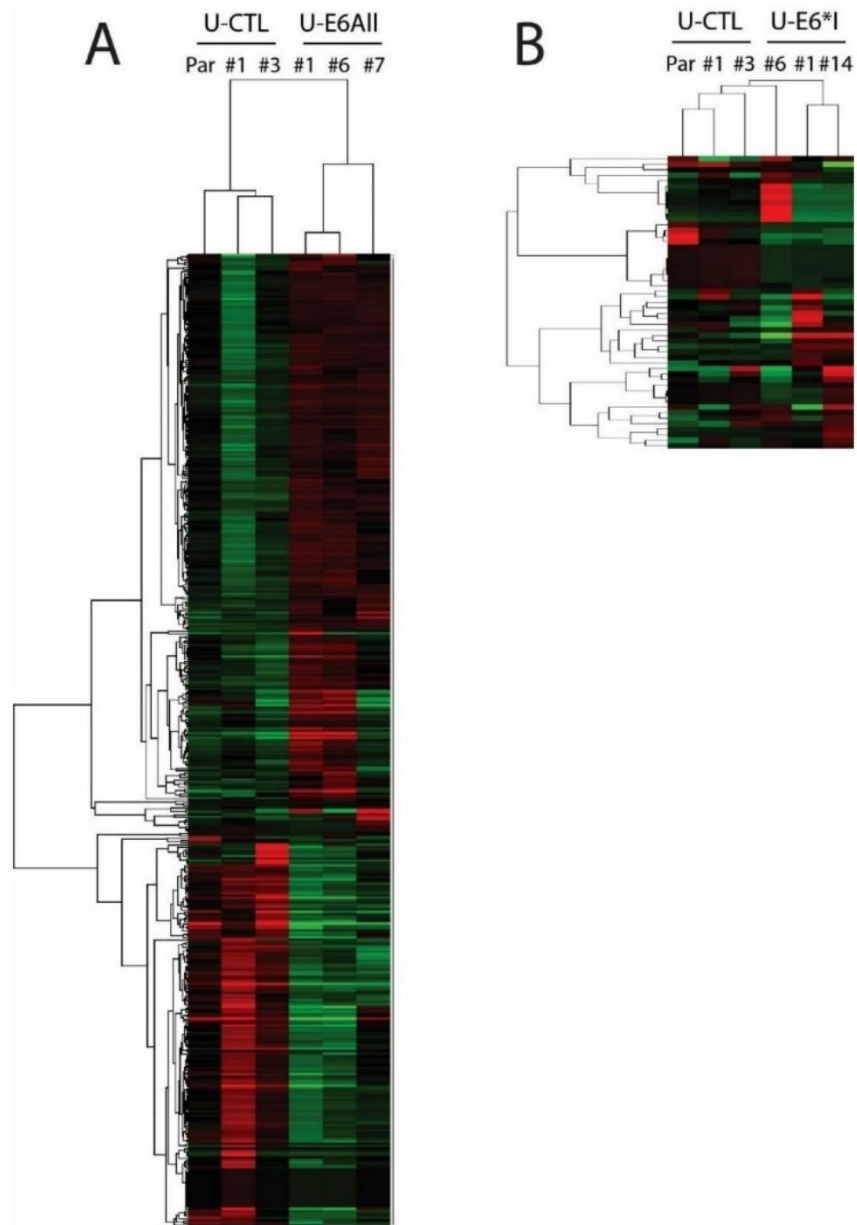


Figure S6. Heatmaps showing gene expression signature for U-E6All vs. U-CTL and U-E6*I vs. U-CTL. (A) The heatmap presentation of the 419 genes significantly deregulated in U-E6All cell lines compared to U-CTL. High (red) and low (green) expression, corresponding to a +7.6 to -5.8 log₂ scale of fold change relative to the average of each transcript across all 6 clones. (B) The heatmap presentation of the 53 genes significantly deregulated in U-E6*I cell lines compared to U-CTL. High (red) and low (green) expression, corresponding to a +6.2 to -2.8 log₂ scale of fold change relative to the average of each transcript across all 6 clones.

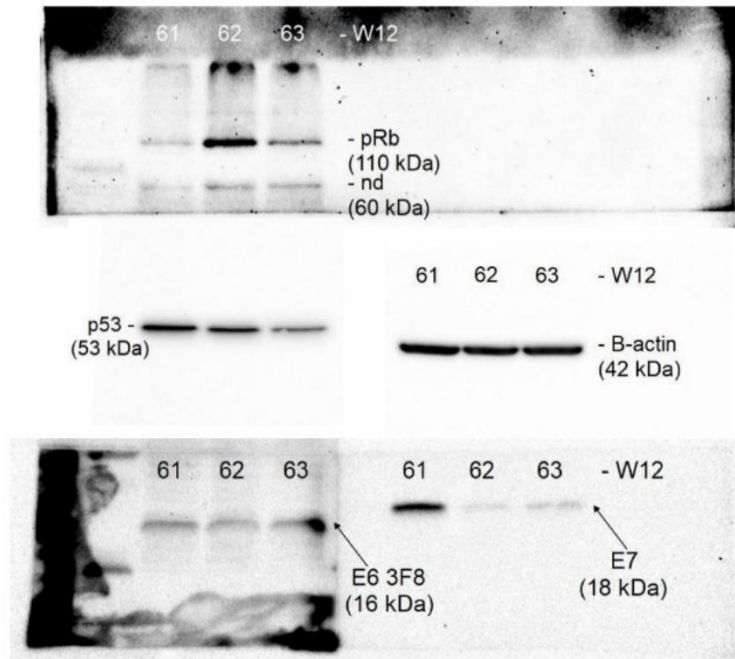


Figure S7. Uncropped Western blot used in figure 8B. The three samples were run in duplicate on the same gel and membrane. Hybridizations and revelation of E6, E7⁴, p53, pRb⁵ and B-actin were done on the same membrane.

Supplemental references

1. Lagrange, M. *et al.* Binding of human papillomavirus 16 E6 to p53 and E6AP is impaired by monoclonal antibodies directed against the second zinc-binding domain of E6. *J. Gen. Virol.* **86**, 1001–1007 (2005).
2. Li, M., Brooks, C. L., Kon, N. & Gu, W. A Dynamic Role of HAUSP in the p53-Mdm2 Pathway. *Mol. Cell* **13**, 879–886 (2004).
3. Kojima, Y. *et al.* Adenovirus-mediated transfer of HPV 16 E6/E7 antisense RNA combined with cisplatin inhibits cellular growth and induces apoptosis in HPV-positive head and neck cancer cells. *Cancer Gene Ther.* **25**, 274 (2018).
4. Hoppe-Seyler, K. *et al.* Viral E6/E7 oncogene and cellular hexokinase 2 expression in HPV-positive cancer cell lines. *Oncotarget* **8**, 106342–106351 (2017).
5. Guo, C. *et al.* Potent Anti-Tumor Effect Generated by a Novel Human Papillomavirus (HPV) Antagonist Peptide Reactivating the pRb/E2F Pathway. *PLOS ONE* **6**, e17734 (2011).

1.2 Travaux complémentaires

1.2.1 Obtention des clones stables

Nous avons généré au laboratoire plusieurs vecteurs à partir du pXJ40-*vide* nous permettant d'étudier l'impact des différentes isoformes de E6. Le plasmide pXJ40-E6All contient l'ADNc correspondant à l'ORF E6 sauvage d'HPV16 et permet d'exprimer les transcrits et protéines E6, E6*I et E6*II. Le plasmide pXJ40-E6 contient l'ORF E6 mutée au niveau du SD226 ce qui empêche l'épissage de l'intron 1 de l'ORF E6 et ne permet que la production de E6. Enfin, le plasmide pXJ40-E6*I contient l'ORF E6 sans l'intron 1 de 182 pb situé entre le SD226 et le SA409 et ne permet que l'expression de E6*I. Nous avons ensuite réalisé des clones stables à partir de la lignée U-2 OS à l'aide de ces plasmides. Grâce aux trois conditions de clones stables générées U-CTL, U-E6All et U-E6*I, nous avons été capables d'étudier l'impact de l'expression des différentes isoformes E6, et l'impact de l'expression de E6*I seule sur le transcriptome cellulaire par RNAseq. Ces clones stables ont été obtenus à partir des mêmes cellules parentales U-2 OS. La sélection des trois clones U-E6All s'est basée sur l'expression des ARNm E6 et E6*I, l'efficacité de l'épissage entre le SD226 et le SA409 et, sur l'expression d'une protéine E6 fonctionnelle. La sélection des clones U-E6*I s'est basée sur l'expression de l'ARNm E6*I à des niveaux comparables à l'expression de l'ARNm E6*I dans les clones U-E6All.

Nos résultats indiquent que la majorité des transcrits cellulaires dérégulés dans la condition U-E6*I, ainsi que l'augmentation des ROS, ne sont pas retrouvés dans la condition U-E6All. Nous avons initialement tenté d'obtenir une quatrième condition exprimant le plasmide pXJ40-E6 de manière stable. Cependant, aucun clone exprimant uniquement E6 de manière fonctionnelle n'a pu être obtenu. L'épissage du transcrit primaire E6 entre le SD226 et le SA409 est efficace à 90% dans le modèle U-2 OS, et seuls 10% des transcrits matures codent la protéine E6. Ce qui n'est pas le cas des transcrits mutés produits à partir du pXJ40-E6, non reconnus par le spliceosome et codant donc tous E6. Cette quantité accrue d'ARNm et de protéine E6 est vraisemblablement létale pour notre modèle cellulaire. L'obtention de la condition U-E6 n'était pas indispensable pour répondre à l'objectif principal de notre étude qui était de définir si l'expression ectopique de E6*I d'HPV16 a un impact sur le transcriptome cellulaire. En revanche, elle aurait pu nous éclairer sur la dynamique entre E6 et E6*I que nous avons constatée et étudiée dans ce travail.

1.2.2 Caractérisation phénotypique complémentaire des clones stables

Nous avons réalisé un ensemble de tests fonctionnels sur les clones stables afin d'identifier des différences phénotypiques significatives imputables aux effets de E6 et E6*I. Nous avons montré que les clones U-E6All sont plus sensibles aux radiations ionisantes que les clones U-CTL et U-E6*I. Et que les clones U-E6*I sont diminués dans leur capacité à former des colonies comparées aux clones U-E6All et U-CTL. Nous avons aussi montré une augmentation du niveau d' H_2O_2 produit par les clones U-E6*I, comparativement aux clones U-E6All et U-CTL.

Nous avons également réalisé d'autres tests phénotypiques sur les clones stables. Afin de mesurer l'impact de l'expression des isoformes E6 sur la prolifération cellulaire, les clones stables ont été ensemencés et leur prolifération a été mesurée pendant 17 jours (figure 21).

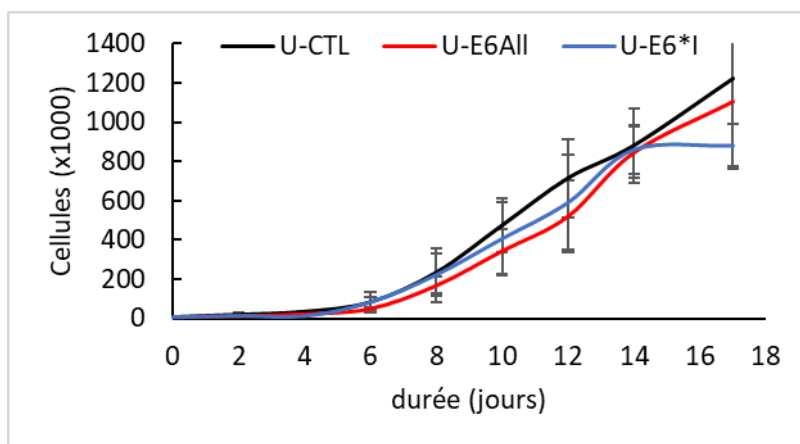


Figure 21 : Etude de la prolifération des clones stables.

Les cellules ont été décollées à la trypsine et comptées tous les deux jours. Chaque courbe représente la moyenne du nombre de cellules de 3 clones (+ ou - S.D.).

L'étude de la prolifération cellulaire n'a pas révélé de différence significative entre les trois conditions. L'expression des deux isoformes E6 et E6*I, ou uniquement de E6*I ne modifie pas la prolifération cellulaire (figure 21).

Afin d'étudier si l'expression des isoformes E6 module la résistance des clones stables à un stress cellulaire, la viabilité des clones traités au sorbitol (stress osmotique) a été mesurée par quantification de la production d'ATP par luminescence (CTG : CellTiter-Glo® assay) au cours du temps (figure 22).

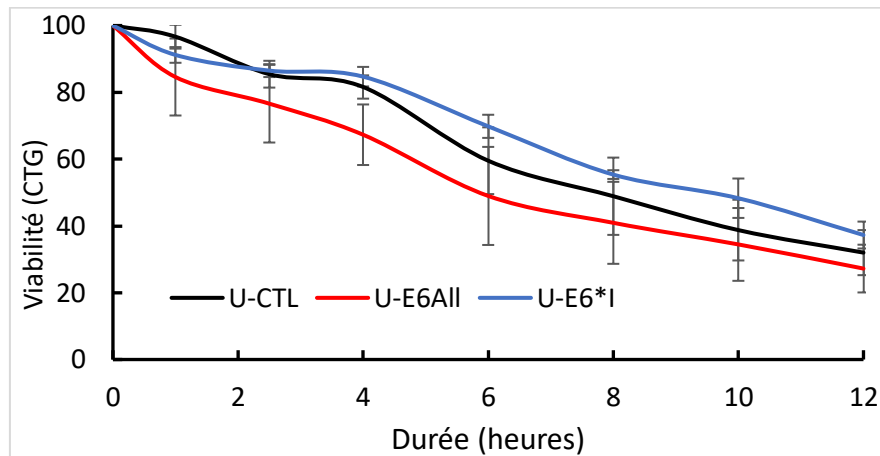


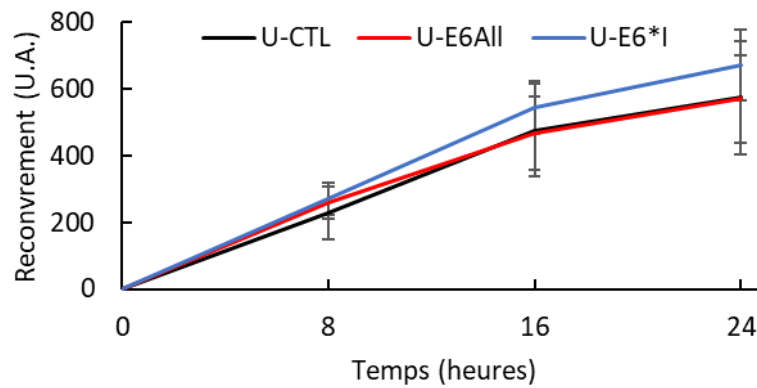
Figure 22 : Etude de la viabilité des clones stables soumis à un stress osmotique.

Après 24h d'ensemencement, 0,3M de sorbitol a été ajouté au milieu de culture des clones stables. Leur viabilité a ensuite été mesurée toutes les deux heures pendant 12h à l'aide du kit CTG. Chaque courbe représente la moyenne des valeurs de luminescence de 3 clones (+ ou - S.D.).

Aucune différence significative n'a été observée dans la diminution de la viabilité des clones indiquant que l'expression des deux isoformes ou uniquement d'E6*I n'a pas d'effet sur la viabilité cellulaire en condition de stress.

L'analyse fonctionnelle des 53 DEG de la condition U-E6*I par la base de données Gene Ontology a montré un enrichissement de gènes impliqués dans l'organisation de la matrice extracellulaire. Afin de déterminer si la dérégulation de ces gènes par E6*I a un impact sur la capacité d'invasion des clones stable, nous avons étudié leur capacité de migration après les avoir cultivés en monocouche à confluence et réalisation d'une blessure (figure 23) (Jonkman et al., 2014).

A



B

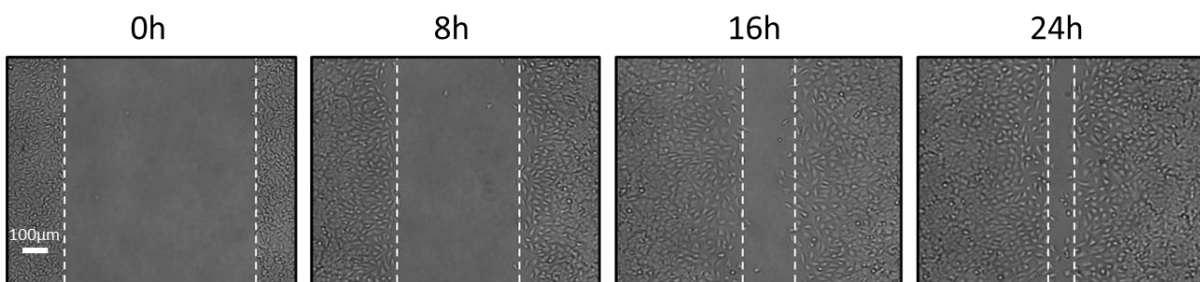


Figure 23 : Etude de la capacité d'invasion des clones stables par blessure/cicatrisation.

La distance de recouvrement/cicatrisation a été mesurée toutes les 8h par microscopie à contraste de phase au même grossissement (x400). (A) Chaque courbe représente la moyenne de la distance de recouvrement de la blessure de 3 clones d'une même condition (+ ou - S.D.) en unité arbitraire. (B) Représentation des images obtenues pour le clone U-E6*I#1, les lignes pointillées représentent les bords de la blessure.

Aucune différence significative n'a été observée dans la capacité des clones stables à combler les blessures, indiquant que ni l'expression de E6*I, ni de E6 ne semblent favoriser l'invasion cellulaire.

Ces tests phénotypiques nous ont permis d'identifier des effets propres à l'isoforme E6 et à l'isoforme E6*I sur le phénotype cellulaire. Nous avons notamment montré que les clones U-E6*I produisent plus de ROS que les clones U-CTL et U-E6All, et que cette production accrue s'accompagne de la dérégulation de l'expression de gènes impliqués dans le métabolisme des ROS dans les clones U-E6*I. Nous nous sommes donc concentrés sur l'étude de la dérégulation du métabolisme des ROS.

2 La protéine E6*I induit la production de ROS

2.1 E6*I induit la production de ROS *via* les complexes Nox

Suite à l'analyse dans la littérature des 53 DEG dérégulées dans les clones U-E6*I, nous avons émis l'hypothèse que la surexpression de CCL2 (Chun et al., 2015), RAC2 (Raad et al., 2009) et PDGFB (Chen et al., 2007) induite par E6*I puisse entraîner l'activation du complexe de la NADPH Oxydase 2. Ce complexe est composé des protéines p67phox, p22phox, p47phox, p40phox et NOX2. La protéine NOX2 fait partie d'une famille de 7 flavoprotéines à la capacité de transfert d'électrons à travers les membranes. NOX2 est enchâssée dans la membrane plasmique et son activation par les GTPases RAC1/2 entraîne l'assemblage du complexe Nox2. Une fois activé, le complexe Nox2 génère des anions superoxydes O_2^- . Ces anions sont ensuite réduits sous forme d' H_2O_2 se retrouvant dans le milieu intra- et extracellulaire pour participer à l'homéostasie redox (Schröder, 2020).

Afin de déterminer si la famille des complexes Nox est impliquée dans la surproduction d' H_2O_2 induite par l'expression de E6*I, nous avons traité les clones stables U-CTL, U-E6All et U-E6*I au DPI (Diphenyleneiodonium), un inhibiteur de l'ensemble des complexes Nox à la dose de 10 μ M (figure 24).

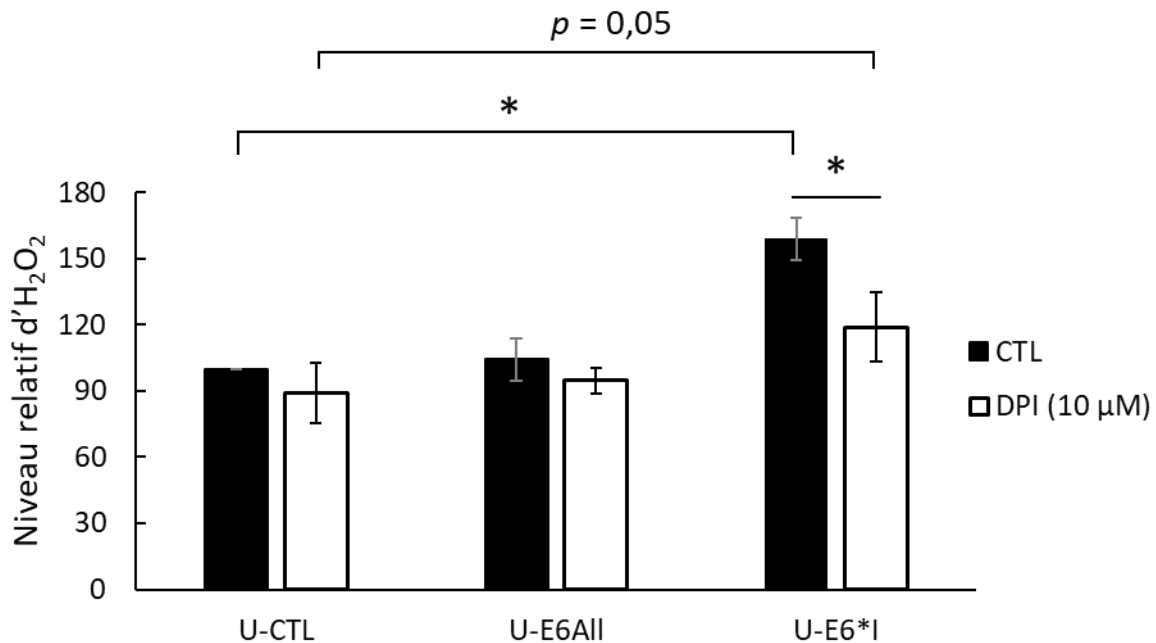


Figure 24 : Détection de l' H_2O_2 produit par les clones stables traités au DPI.

Les données représentent la quantité relative d' H_2O_2 produite par les clones stables en 1h avec ou sans traitement au DPI 10 μ M. Les niveaux relatifs d' H_2O_2 sont représentés en tant que pourcentage moyen +/- S.D. de trois expériences indépendantes par rapport au contrôle. Test statistique Mann-Whitney unilatéral * $p < 0,05$.

Nous confirmons la surproduction significative d'H₂O₂ par les clones U-E6*I par rapport aux U-CTL et U-E6All (figure 28). Le traitement des clones contrôle et des clones U-E6All par 10 µM de DPI n'entraîne pas de diminution significative de la production d'H₂O₂. En revanche, le traitement des clones U-E6*I entraîne une diminution significative de la production d'H₂O₂ de l'ordre de 25% par rapport aux clones U-E6*I non traités. L'activation des complexes Nox est donc au moins en partie à l'origine de la surproduction d'H₂O₂ induite par l'expression de E6*I. En effet, le traitement au DPI des clones U-E6*I n'entraîne pas le retour au niveau basal de la production d'H₂O₂ observée dans les clones U-CTL. Cette différence indique que d'autres effecteurs du métabolisme des ROS sont dérégulés par l'expression de E6*I.

Afin de déterminer si l'activation du complexe Nox2 est impliqué dans l'augmentation du taux d'H₂O₂ induite par l'expression de E6*I, les clones stables U-E6*I ont été traités par 1µM de DPI (figure 25). A cette dose le DPI inhibe spécifiquement le complexe Nox2 (Reis et al., 2020).

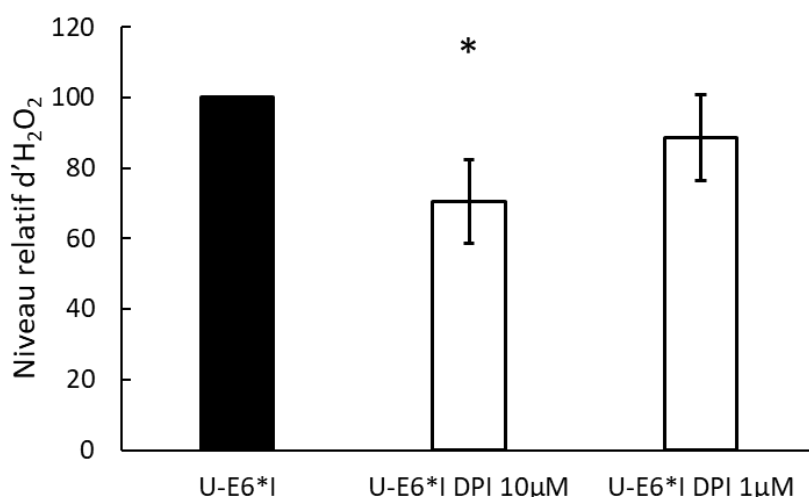


Figure 25 : Détection de l'H₂O₂ produit par les clones U-E6*I traités au DPI.

Les données représentent la quantité relative d'H₂O₂ produite par les clones U-E6*I en 1h avec ou sans traitement au DPI 10 µM ou 1 µM. Les niveaux relatifs d'H₂O₂ sont représentés en tant que pourcentage moyen +/- S.D. de trois expériences indépendantes par rapport au contrôle. Test statistique Mann-Whitney unilatéral *p<0,05.

Nous retrouvons la diminution significative du niveau d'H₂O₂ lorsque les clones U-E6*I sont traités par 10 µM de DPI. Nous observons également une tendance non-significative à la diminution du taux d'H₂O₂ produit par les clones U-E6*I traités par 1 µM de DPI. Cette expérience ne nous permet pas de conclure sur l'activation du complexe Nox2. Des expériences complémentaires, d'interférence ARN anti-NOX2 notamment, seront nécessaires pour démontrer si le complexe Nox2 est activé par l'expression de E6*I.

2.2 Détection de la protéine E6*I par western blotting

La littérature et nos travaux ont montré avec consistance que la transfection de manière stable ou transitoire d'un vecteur codant E6*I d'HPV16 entraîne une dérégulation du métabolisme des ROS (Williams et al., 2014) (Paget-Bailly et al., 2019). Bien que les transcrits E6*I soient quantitativement majoritaires par rapport aux transcrits E6, et ce dans tous les modèles d'étude, la protéine E6*I est difficile à détecter par western blotting. En transfection transitoire, la protéine E6*I n'est détectable qu'après un traitement de 16h au MG132 10 μ M, inhibant sa dégradation par le protéasome. De manière intéressante, la protéine E6*I n'est pas détectable dans les cellules U-2 OS transfectées avec le pXJ40-E6All permettant l'expression de E6 et E6*I et traitées au MG132. Pourtant, les taux de transcrits E6*I sont similaires lorsque l'on transfecte transitoirement le pXJ40-E6All ou le pXJ40-E6*I (Paget-Bailly et al., 2019).

Afin de déterminer si la protéine E6*I est détectable dans les clones stables U-E6All et U-E6*I, nous avons traité les clones au MG132 10 μ M pendant 16h (figure 26).

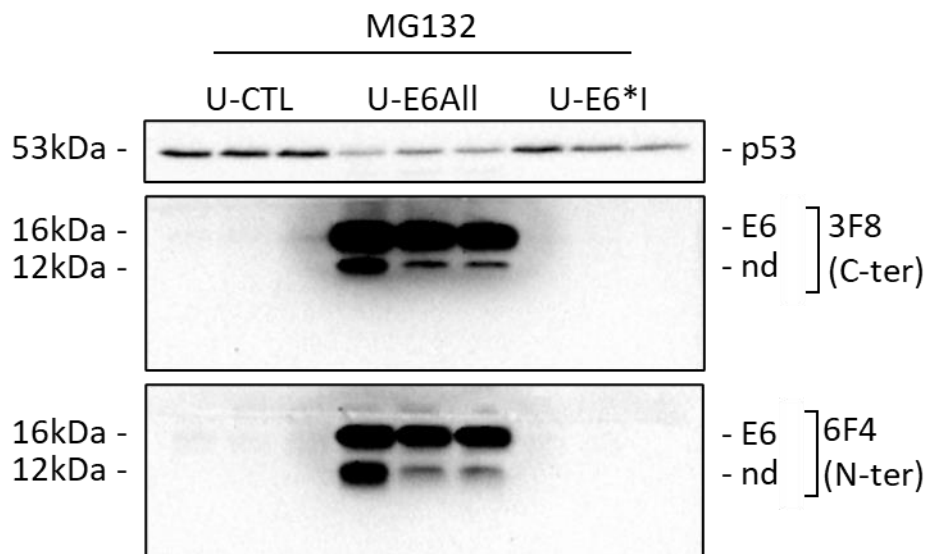


Figure 26 : Détection des isoformes E6 dans les clones stables après inhibition du protéasome.

Les isoformes E6 ont été détectées dans les clones stables par western blotting après 16h de traitement au MG132 10 μ M. L'anticorps 6F4 reconnaît la partie N-terminale de E6 et peut donc détecter E6 et E6*I, là où l'anticorps 3F8 ne reconnaît que E6 *via* sa partie C-terminale. La bande nd retrouvée avec les deux anticorps correspond vraisemblablement à un produit de clivage s'accumulant suite à l'inhibition du protéasome.

Après traitement au MG132, nous observons une accumulation massive de la protéine E6, détectée avec les anticorps reconnaissant sa partie C-terminal ou N-terminal dans les clones U-E6All. Nous observons aussi une bande à 12 kDa correspondant vraisemblablement à un produit de clivage de E6, qui s'accumule en absence de dégradation par le protéasome.

De par sa taille et sa détection avec les deux anticorps anti-E6, cette bande ne peut pas correspondre à E6*I. La protéine E6*I est indétectable par western blotting que ce soit dans les clones stables U-E6All ou U-E6*I. L'absence de détection de la protéine E6*I après traitement au MG132 peut provenir de l'expression plus faible des transcrits E6*I dans les clones stables par rapport aux transfections transitoires.

Il se pourrait également que les transcrits E6*I ne soient pas traduits en protéine dans les clones stables. Dans ce cas, nous pouvons émettre l'hypothèse que l'expression de l'ARNm E6*I suffit à provoquer les dérégulations du transcriptome et l'augmentation de la production des ROS observées dans les clones stables U-E6*I. Afin d'étudier les effets de l'expression des ARNm E6 et E6*I, nous avons muté les vecteurs pXJ40-E6All, -E6 et -E6*I, au niveau des codons d'initiation de la traduction ATG83 et 104, les rendant incapable de produire les isoformes E6. Les séquences codantes ont été mutées sur ces deux ATG car il existe encore à l'heure actuelle une incertitude concernant le choix de l'un ou l'autre pour l'initiation de la traduction de E6 et de ses isoformes. Les vecteurs pXJ40- Δ ATGE6All, Δ ATGE6 et Δ ATGE6*I ont ensuite été transfectés transitoirement dans les cellules U-2 OS (figure 27).

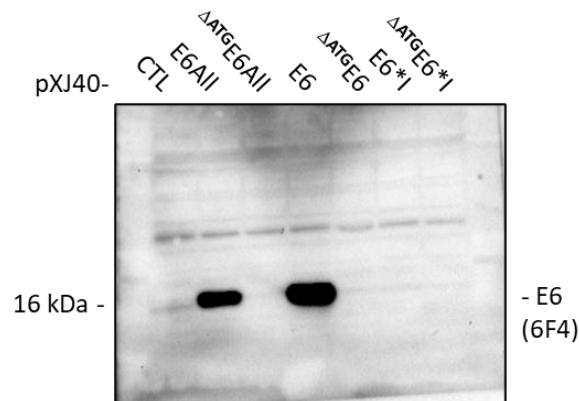


Figure 27 : Détection des isoformes E6 exprimés à partir des vecteurs mutés ou non sur les codons d'initiation de la traduction.

Analyse par western blotting de l'expression des isoformes E6 après transfection des U-2 OS pendant 24h avec les vecteurs pXJ40-CTL, -E6All, Δ ATGE6All, -E6, Δ ATGE6, -E6*I et Δ ATGE6*I.

Nous observons l'expression de l'isoforme entière E6 par western blotting dans les cellules transfectées par les vecteurs pXJ40-E6All et pXJ40-E6. La protéine E6 est en revanche indétectable dans les cellules transfectées par les vecteurs pXJ40- Δ ATGE6All et pXJ40- Δ ATGE6. Cependant, la protéine E6*I reste indétectable dans les cellules transfectées avec le pXJ40-E6*I ou avec le pXJ40- Δ ATGE6*I.

2.3 L'induction des ROS est causée par la protéine E6*I d'HPV16

Afin de déterminer si la surproduction d'H₂O₂ observée dans les clones U-E6*I est dû à un effet de l'ARNm E6*I ou de la protéine E6*I, nous avons mesuré le taux d'H₂O₂ produit par les cellules U-2 OS transfectées transitoirement avec les vecteurs pXJ40-E6All, - Δ ATG E6All, -E6*I et - Δ ATG E6*I. Les vecteurs pXJ40-E6 et - Δ ATG E6 ont aussi été transfectés afin d'étudier si l'expression de l'ARNm ou de la protéine E6 module les niveaux d'H₂O₂ produit par les U-2 OS (figure 28).

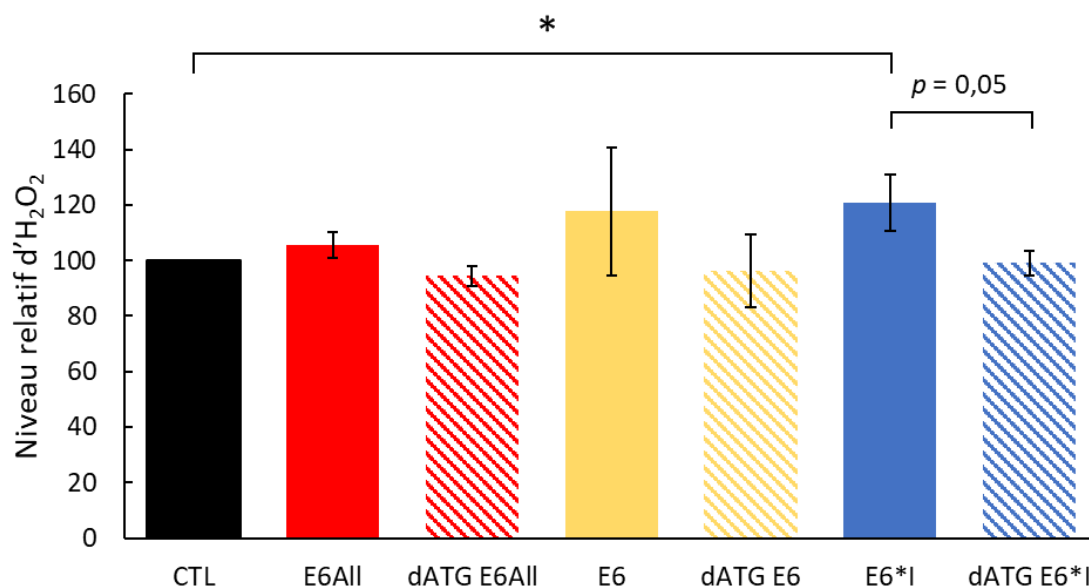


Figure 28 : Détection de l'H₂O₂ produit par les cellules U-2 OS exprimant les différents vecteurs E6.

Les données représentent la quantité relative d'H₂O₂ produite en 1h après 24h de transfection. Les niveaux relatifs d'H₂O₂ sont représentés en tant que pourcentage moyen +/- S.D. de trois expériences indépendantes par rapport au contrôle. Test statistique Mann-Whitney unilatéral *p<0,05.

Nous observons un taux d'H₂O₂ similaire dans la condition contrôle ou transfectées avec le pXJ40-E6All (figure 28), en accord avec le taux d'H₂O₂ mesuré pour les clones stables U-CTL et U-E6All (Paget-Bailly et al., 2019). Également en accord avec les clones stables U-E6*I, nous retrouvons une augmentation significative d'H₂O₂ dans la condition exprimant transitoirement le pXJ40-E6*I. L'expression de E6 montre une tendance non-significative à l'augmentation du taux d'H₂O₂ par rapport aux cellules contrôles. Enfin, l'expression des trois vecteurs Δ ATG n'entraîne pas d'augmentation du taux d'H₂O₂ indiquant que l'expression des ARNm E6 et E6*I n'a pas d'impact sur le métabolisme des ROS. L'augmentation d'H₂O₂ suite à l'expression du vecteur pXJ40-E6*I mais pas du vecteur - Δ ATG E6*I démontre bien que les effets sur le métabolisme des ROS sont médiés par l'expression protéique de E6*I.

2.4 Etude de la traduction de l'ARNm E6*I par profilage polysomique

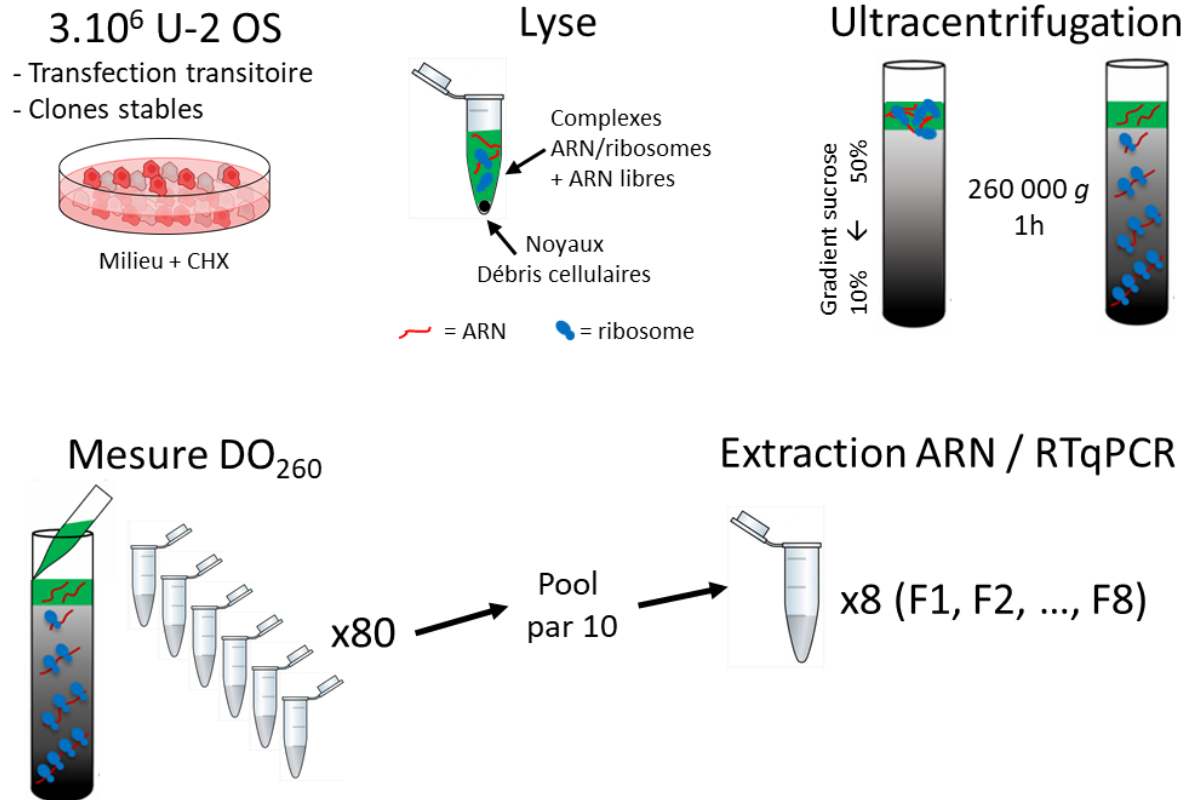


Figure 29 : Technique de profilage polysomique.

Pour le protocole détaillé voir 4. Matériel et méthodes. Les cellules U-2 OS sont d'abord traitées à la CHX afin de fixer les ribosomes sur les ARNm en cours de traduction. Après récolte, les cellules sont lysées et le cytosol contenant les polysomes et les ARNm libre est purifié. Le lysat est déposé sur un gradient de sucrose et est ultracentrifugé à 260000g pendant 1h. Ensuite, le gradient est fractionné en 80 fractions. Chaque fraction est dosée afin de mesurer la présence des ARN et établir un graphique correspondant au profil de polysome. Les fractions sont regroupées par 10 pour en extraire les ARN. Les ARN sont enfin analysés par RT-qPCR.

La difficulté à détecter la protéine E6*I peut avoir plusieurs explications comme une faible efficacité de traduction, une dégradation très efficace ou encore un problème d'ordre technique relatif à la détection par western blotting. A ce jour, aucune protéine cellulaire capable de dégrader la protéine E6*I n'a été identifiée dans la littérature. Nous avons donc décidé d'étudier sa production, en absence de traitement au MG132. Afin de déterminer si les transcrits E6*I sont efficacement traduits en protéine, nous avons étudié la prise en charge de ces transcrits par les ribosomes par profilage polysomique. Cette technique permet d'étudier semi-quantitativement la traduction d'un transcrit donné. Pour se faire, les cellules exprimant les différents transcrits E6 de manière stable ou transitoire ont été traitées à la CHX (cycloheximide) qui fixe les ribosomes en cours de traduction sur les ARNm. Plus un ARNm est traduit, plus le nombre de ribosomes fixés est grand. Les complexes ARN/ribosome(s) appelés polysomes sont donc plus ou moins lourds et peuvent ainsi être séparés par

ultracentrifugation sur un gradient de sucrose. Le fractionnement de ce gradient et la quantification par RT-qPCR des transcrits E6 dans les différentes fractions permet d'estimer l'efficacité de traduction des transcrits. Les différentes étapes de cette technique sont résumées en figure 29.

Afin de déterminer si l'ARNm E6*I est traduit, nous avons dans un premier temps transfecté transitoirement les plasmides pXJ40-E6*I et pXJ40-E6 dans la lignée U-2 OS, réalisé un fractionnement et quantifié les ARNm E6 et E6*I dans les fractions obtenues. Le vecteur exprimant E6 sert ici de référence de traduction car son expression protéique est facilement observable par western blotting. Si le transcrit E6*I se trouve dans les mêmes fractions que le transcrit E6, alors c'est une preuve indirecte de sa traduction en protéine. Avant de quantifier les transcrits HPV, la technique a été validée par mesure du profil de polysomes des cellules transfectées (figure 30).

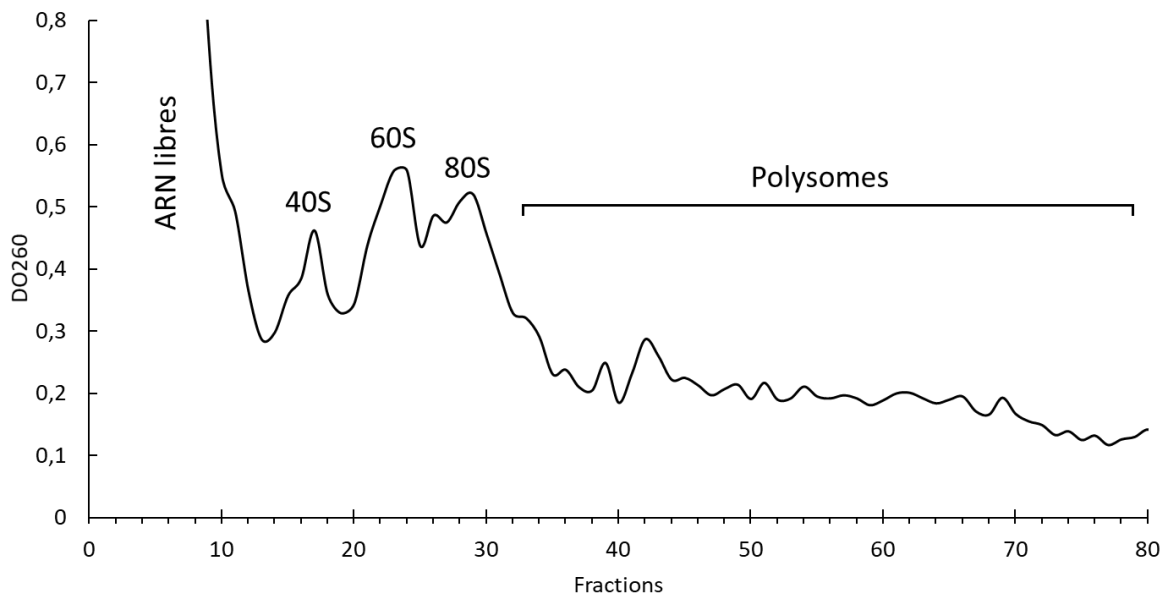


Figure 30 : Profil de polysomes des cellules U-2 OS exprimant E6 et E6*I.

Les cellules U-2 OS ont été transfectées avec les vecteurs pXJ40-E6*I et pXJ40-E6 pendant 48h. Après ultracentrifugation sur gradient de sucrose, 80 fractions ont été récupérées et dosées à 260nm.

Le profil obtenu est cohérent avec les résultats de la littérature et révèle une résolution correcte des pics d'intérêt avec d'abord une très forte présence d'ARNm libre dans les 10 premières fractions, un pic correspondant aux ARNm liés à des petites sous-unités ribosomiques (40S), puis à des grandes sous-unités (60S). Nous observons ensuite un pic correspondant aux ARNm liés à un seul ribosome (80S) appelé monosome et enfin, des pics moins marqués correspondant aux ARNm liés à 2, 3, 4, etc. ribosomes appelés les polysomes (Seimetz et al., 2019). Les fractions ont ensuite été regroupées par 10 et les ARN qu'elles

contiennent ont été extraits. Pour vérifier le fractionnement, les ARN extraits ont été déposés sur gel d'agarose afin d'observer la présence des ARNr 28S et 18S et 5,8S (figure 31).

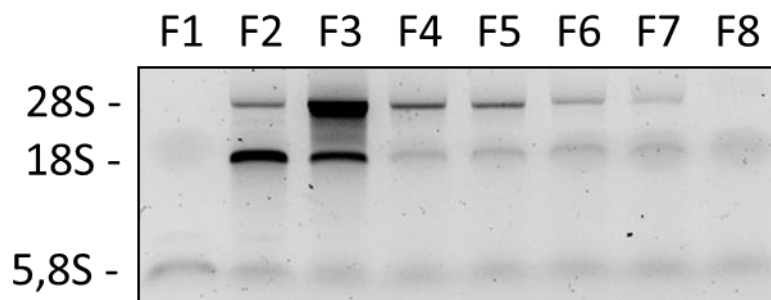


Figure 31 : Présence des ARNr dans les différentes fractions du gradient.
Après extraction, un volume fixe d'ARN a été déposé sur un gel d'agarose 1%.

La fraction 1 est enrichie en ARNr 5,8S libre et la fraction 2 en ARNr 18S contenus dans les petites sous-unités libres. La fraction 3 est fortement enrichie en ARNr 28S et correspond aux grandes sous-unités libres et aux monosomes. Les fractions 4 à 8 présentent un profil avec approximativement deux fois plus de 28S que de 18S et contiennent les polysomes. Ici encore, la présence des ARNr dans les différentes fractions est cohérente avec la littérature (Karamysheva et al., 2018). La traduction des ARNm suit une loi normale qui implique que les ARNm fortement traduits sont peu représentés dans le transcriptome, d'où la diminution progressive des ARNr 28S et 18S dans les fractions 4 à 7. Les ARNr sont non détectables dans la fraction 8 du fait de la faible quantité d'ARN total récoltée dans cette fraction. Enfin, la présence des transcrits E6 et E6*I a été quantifiée par RT-qPCR dans les 8 fractions selon la méthode décrite par (Pringle et al., 2019) (figure 32).

Nous observons (figure 32) que 60% des transcrits E6 et E6*I sont présents dans les fractions F3 à F8 contenant les monosomes et les polysomes, ce qui semble indiquer une bonne prise en charge de ces transcrits par les ribosomes de prime abord. En revanche, nous observons un enrichissement similaire de 60% pour les transcrits E6 et E6*I muté sur les codons d'initiation, ainsi que pour l'ARN non-codant PICSAR. Le transcrit de la GAPDH (Glycéraldéhyde-3-Phosphate Déshydrogénase) quant à lui est présent à 95% dans les monosomes et polysomes témoignant de sa forte traduction. Nous observons une différence d'enrichissement entre monosomes et polysomes pour les transcrits E6 et E6*I. En effet, le transcrit E6 est présent à hauteur de 45% dans les polysomes contre 30% pour le transcrit E6*I. Cette différence démarque le transcrit E6 des transcrits non-codant, là où le transcrit E6*I leur est semblable. Bien que plus présent dans les polysomes que E6*I, le transcrit E6 y est largement sous-représenté par rapport au transcrit de la GAPDH, reflétant une traduction

moins efficace. L'absence d'un contrôle négatif franc ne nous permet pas de statuer sur la traduction de E6*I. Si le transcrit E6*I est effectivement traduit il semble l'être moins efficacement que E6.

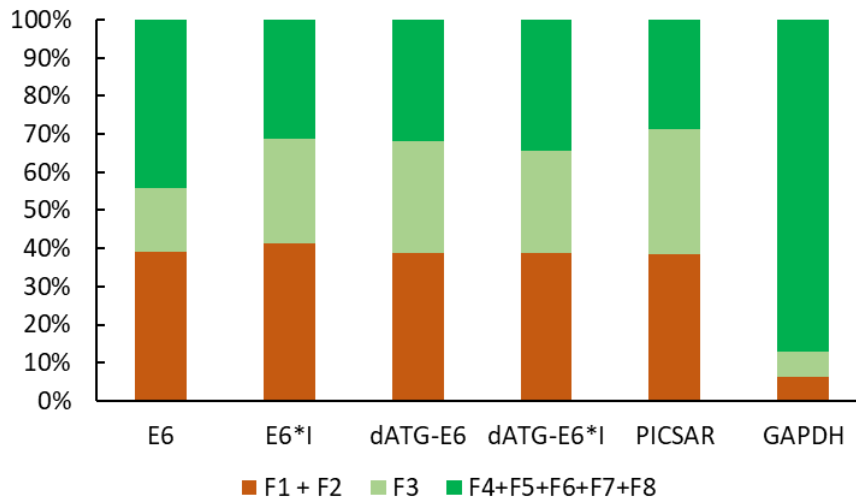


Figure 32 : Quantification des transcrits E6 et E6*I exprimés transitoirement après fractionnement.

Analyse par RT-qPCR des 8 fractions obtenues après fractionnement et après 24h de transfection. Le % de transcrit est obtenu par la méthode 2^{-ddCt} en divisant le RQ d'une fraction par la somme des RQ des 8 fractions pour un transcrit donné. L'enrichissement des transcrits est représenté sous forme d'histogrammes. Les transcrits deltaATG ont été quantifiés à partir de transfactions transitoires indépendantes des plasmides pXJ40- Δ ATG E6 et pXJ40- Δ ATG E6*I dans des cellules U-2 OS, soumises au même protocole. PICSAR est un ARN non-codant cellulaire et sert de contrôle négatif tandis que le transcrit de la GAPDH sert de contrôle positif. Est représentée la moyenne de trois expériences indépendantes (+ ou - S.D.).

Nous avons enfin réalisé ces expériences de profilage polysomique sur les clones stables, U-CTL, U-E6All et U-E6*I (figure 33). Les profils de polysomes (courbe DO260) n'ont pas révélé de différence significative entre les trois conditions (données non-montrées).

La quantification des transcrits E6 et E6*I dans les différentes fractions des clones stables U-E6All et U-E6*I révèlent un enrichissement des transcrits E6 et E6*I différent comparé aux transfactions transitoires (figure 33). Dans les clones stables, les transcrits E6 et E6*I sont moins présents dans les fractions libres (20%) comparé aux transitoires (40%). Nous observons aussi globalement un enrichissement de la fraction monosome dans les clones U-E6All et U-E6*I. Il se peut que la transfection transitoire surcharge les cellules en transcrit E6 et E6*I par rapport à l'expression des clones stables, entraînant au final leur surreprésentation dans les fractions libres. Pour les clones U-E6All, nous observons comme dans les transfactions transitoires que le transcrit E6 est plus présents dans les polysomes que le transcrit E6*I. De manière très surprenante en revanche, nous observons dans les clones U-E6*I que la proportion de transcrit E6*I double dans les polysomes comparés aux clones U-

E6All. Le profil des transcrits E6*I dans les clones U-E6*I semble indiqué une meilleure prise en charge par la machinerie traductionnelle que dans les clones U-E6All. De plus, les transcrits E6*I semble même mieux pris en charge dans les clones U-E6*I, que les transcrits E6 dans les clones U-E6All. Il se pourrait que la protéine E6 ou ses effets cellulaires entraînent une inhibition de la traduction de E6*I. Le transcrypt non-codant PICSAR présente quant à lui le même profil d'enrichissement entre transitoires et stables, que ce soit pour les fractions libres, les monosomes ou les polysomes. Le transcrypt GAPDH est pour finir moins enrichi dans les polysomes issus des clones stables, bien que cette proportion reste au moins deux fois supérieure aux autres transcrits quantifiés. Sa traduction pourrait être partiellement empêchée par les effets cellulaires de E6 et/ou E6*I dans les clones U-E6All et U-E6*I.

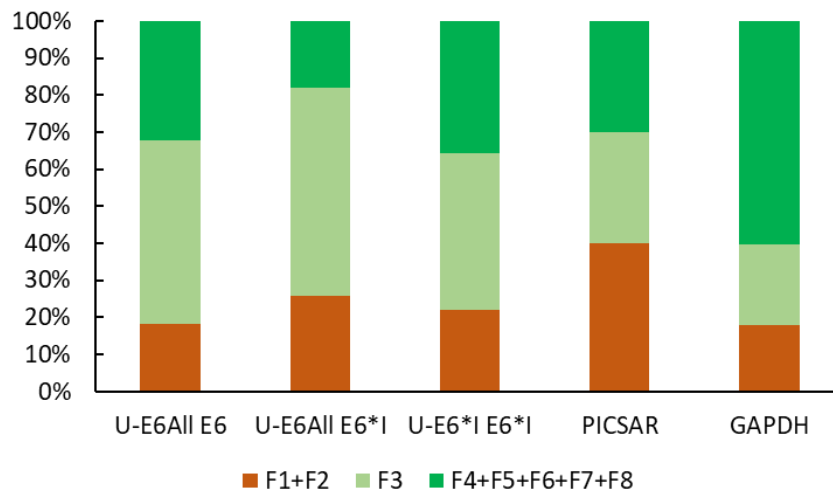


Figure 33 : Quantification des transcrits E6 et E6*I après fractionnement par polysome profiling dans les clones stables U-E6All et U-E6*I.

Analyse par RT-qPCR des 8 fractions obtenues après fractionnement. Le % de transcrypt est obtenu après méthode des 2^{-ddCt} en divisant le RQ d'une fraction par la somme des RQ des 8 fractions pour un transcrypt donné. L'enrichissement des transcrits E6 et E6*I est représenté sous forme d'histogrammes tandis que celui des transcrits PICSAR et GAPDH est représenté sous forme de courbes. Est représenté la moyenne du taux de transcrits de trois clones (+ ou - S.D.). Mann-Whitney, n.s. non-significatif.

Les données sur la prise en charge des transcrits E6*I par les ribosomes dans les transfections transitoires et les clones stables ne nous permettent pas de démontrer si le transcrypt est effectivement traduit ou non. Nous observons une différence marquée entre les profils de la GAPDH et ceux de E6 et E6*I qui témoigne de la meilleure prise en charge / traduction du transcrypt cellulaire par rapport aux transcrits viraux. En revanche, les profils des ARNm E6 et E6*I sont très proches des profils de transcrits non-traduits, alors même que l'expression protéique de E6 est avérée à la fois dans les transfections transitoires et les clones stables U-E6All.

3 Etude des effets des isoformes E6 dans une cohorte de cancers de l'oropharynx

Nous disposons au laboratoire d'une cohorte de 45 échantillons de tumeurs de l'oropharynx qui nous a été fournie par le Dr. Alain C. Jung dans le cadre d'une collaboration sur la thématique de l'épissage de l'ORF E6. Ces 45 tumeurs sont positives à l'ADN d'HPV16. En effet, 30% dans cancers de l'oropharynx sont causés par HPV16. Ces 45 échantillons font partis d'une plus large cohorte qui a été testée qualitativement pour la présence du transcrit E6E7 (Jung et al., 2010). Cette étude montre que la stratification des patients en fonction du statut transcriptionnellement actif d'HPV16 dans les tumeurs révèle des différences phénotypiques par rapport aux tumeurs n'exprimant pas les transcrits E6 et E7. Les patients avec des tumeurs exprimant l'ARNm E6E7 présentent un meilleur taux de survie à 5 ans et une signature hypoxique moindre comparé aux tumeurs positives pour HPV16 mais n'exprimant pas le transcrit E6E7 ou alors négatives pour HPV16. Cette cohorte réduite est constituée de 22 échantillons négatifs pour l'ARNm E6E7 (ARN E6 -), et de 23 échantillons exprimant cet ARNm (ARN E6 +). Les données cliniques de ces deux sous-groupes sont présentées tableau I.

Tableau 1 : Description des caractéristiques de patients et de leur tumeur HPV16+ en fonction de leur statut transcriptionnellement actif.

	ADN HPV16 +	
	ARN E6 - N=22	ARN E6 + N=23
Genre		
Homme	17	17
Femme	5	6
Age		
≥60 ans	12	12
<60 ans	10	11
Survie (mois)		
≥36	11	14
<36	11	9
Tabac		
non	3	5
oui	16	17
Alcool		
non	5	6
oui	14	14
Taille tumeur		
T1	2	0

T2	1	1
T3	4	9
T4	15	13
Statut lymphatique		
N0	3	4
>N1	19	19
Métastases		
non	16	16
oui	6	7

3.1 Expression des transcrits E6 dans les tumeurs de l'oropharynx

Nous avons utilisé les 23 échantillons ARN E6 + pour déterminer si les niveaux des transcrits E6 et E6*I, et le ratio E6*I/E6 pouvaient nous permettre de stratifier d'avantage cette cohorte afin d'en tirer une valeur pronostique. Pour se faire, nous avons quantifié le nombre de transcrits E6 et E6*I par RT-qPCR (figure 34).

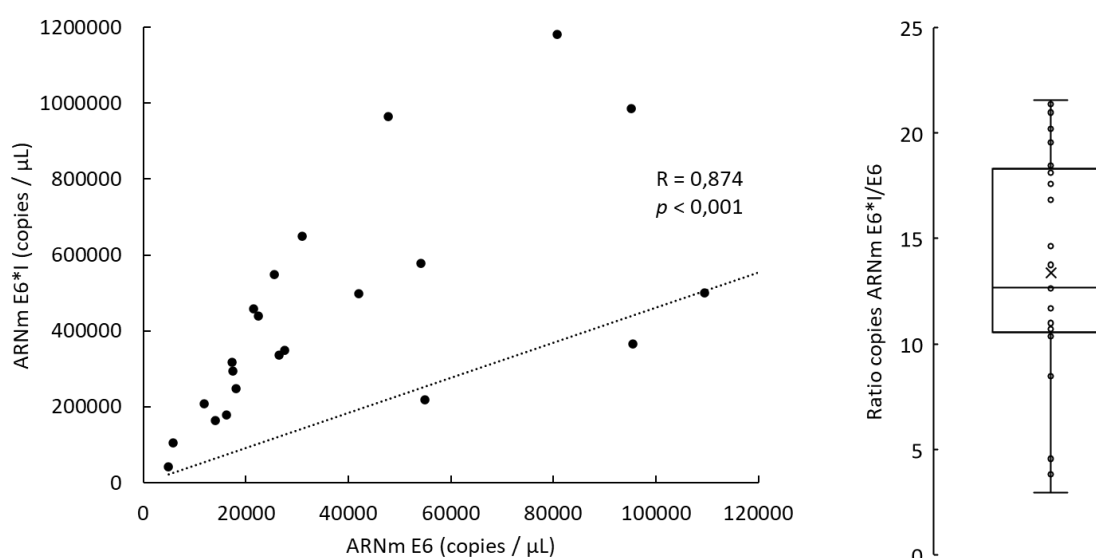


Figure 34 : Expression des ARNm E6 et E6*I dans une cohorte de cancers de l'oropharynx HPV16+.

(Gauche) Quantification absolue par RT-qPCR des ARNm E6 et E6*I dans les 23 échantillons transcriptionnellement actifs. Le graphique représente les niveaux absolus des ARNm E6 et E6*I de 22 échantillons. Le 23^{ème} échantillon n'est pas représenté à cause de ses valeurs extrêmes d'expression des ARNm E6 et E6*I. En revanche, l'équation de la droite de régression linéaire en pointillée est calculée sur les 23 échantillons exprimant les transcrits E6 (R = coefficient de corrélation de Spearman). (Droite) Boîte à moustache représentant le ratio d'expression des transcrits E6*I par rapport aux transcrits E6 dans les 23 échantillons.

Nous observons une forte variabilité de l'expression des transcrits E6 et E6*I entre les échantillons. En revanche, nous n'observons pas de modifications drastiques du profil d'épissage de l'ORF E6 dans ces tumeurs. En effet, les échantillons contiennent en moyenne

14 fois plus de transcrits E6*I que de transcrits E6 (figure 34). Ces résultats indiquent un épissage très efficace de l'intron 1 de E6 entre le SD226 et le SA409 dans cette cohorte de cancers de l'oropharynx. Il existe une forte corrélation entre les niveaux d'ARNm E6 et E6*I ($R = 0,874$; $p < 0,001$). Cette corrélation rend les échantillons très homogènes au regard du ratio E6*I/E6. Au final, cette homogénéité ne nous a pas permis de stratifier les 23 échantillons entre les patients dont les tumeurs présentent un ratio E6*I/E6 élevé ou au contraire faible.

3.2 Dérégulation des ARNm cellulaires en fonction de l'expression des transcrits E6

Nous avons ensuite mesuré par RT-qPCR l'expression relative de certains transcrits d'intérêt issus du RNAseq (Paget-Bailly et al., 2019) dans ces 23 échantillons ARN E6 + afin de valider leur dérégulation par l'expression des isoformes E6. Les échantillons ARN E6 + ont été séparés en deux groupes (E6 faible $n = 11$; E6 fort $n = 12$) en fonction de la médiane du niveau d'expression de l'ARNm E6 et afin de déterminer si il y a corrélation entre l'augmentation de l'expression des transcrits E6 et l'expression des transcrits cellulaires. Le niveau des transcrits a également été quantifié dans les 22 échantillons ARN E6 - pour servir de point de référence (figure 35).

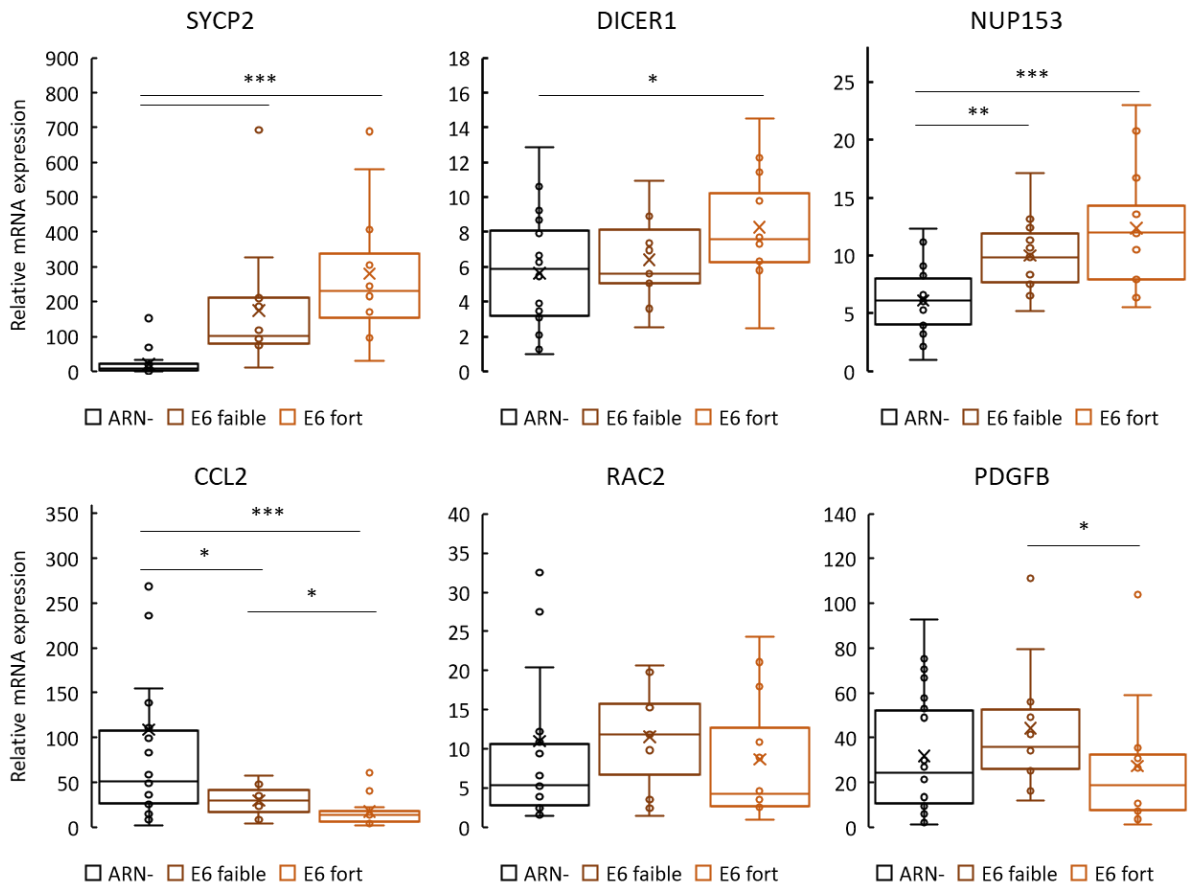


Figure 35 : Expression de transcrits cellulaires en fonction de l'expression des transcrits E6. Analyse par RT-qPCR du taux d'ARNm de SYCP2, DICER1, NUP153, CCL2, RAC2 et PDGFB en fonction de l'expression de E6. Les échantillons (n=45) sont stratifiés en fonction de l'absence des transcrits E6 (n=22) ou de leur présence (n=23). Ces derniers sont sous-stratifiés en fonction de la médiane d'expression des transcrits E6 (marron = E6 faible ; orange = E6 fort). RPLP0 et UBB sont utilisés comme contrôles. Test statistique Mann-Whitney unilatéral *p<0,05, **p<0,01, ***p<0,001.

Nous confirmons l'augmentation de l'expression des transcrits cellulaires SYCP2, DICER1 et NUP153 entre les échantillons ARN- et ARN+. Le transcrit codant SYCP2 a été choisi comme référence car il est fortement surexprimé dans la condition U-E6All et parce que sa surexpression a déjà été observée dans des échantillons de cancers cervicaux et de l'oropharynx HPV16+ (Pyeon et al., 2007) (den Boon et al., 2015) (Masterson et al., 2015). Le transcrit codant DICER1, également surexprimé dans la condition U-E6All, a été montré comme surexprimé dans des lignées de cancers cervicaux HPV16+ (Harden and Munger, 2017). NUP153 est aussi augmenté dans les clones U-E6All. L'augmentation du transcrit codant NUP153 n'a pas été précédemment montré dans la littérature à notre connaissance. Nous observons ici que son expression et celle du transcrit codant SYCP2 corrèle avec l'expression des transcrits E6 dans les échantillons de cancers de l'oropharynx (figure 35). Le fait que ces transcrits soient dérégulés dans les clones stables exprimant uniquement les ARNm E6 nous permet de conclure que E6 seul suffit à déréguler ces transcrits dans ces

tumeurs. E7 pourrait également impacter l'expression de ces trois transcrits dans ces tumeurs, comme cela a été montré pour SYCP2 (Pyeon et al., 2007) et DICER1 (Harden and Munger, 2017).

Nous avons proposé dans notre article que la surexpression des transcrits CCL2, RAC2 et PDGFB dans les clones U-E6*I puisse contribuer à l'augmentation des ROS. Dans la condition U-E6All en revanche, CCL2 est réprimé et, RAC2 et PDGFB ne varient pas par rapport aux U-CTL. Nous retrouvons des résultats similaires à la condition U-E6All avec une diminution d'expression de CCL2 et pas de différence significative pour RAC2 et PDGFB entre les tumeurs ARN E6 - et ARN E6 +. De manière intéressante, le transcrit codant CCL2 est d'autant plus réprimé que le transcrits E6 est exprimé. Nous émettons encore une fois l'hypothèse que les effets observés sont provoqués par E6, malgré l'expression majoritaire du transcrit épissé E6*I dans ces tumeurs comme dans les clones stables.

4 Rôles de l'épissage dans l'expression de l'oncoprotéine E7

L'épissage entre le SD226 et le SA409 du transcrit E6E7 d'HPV16 sert à réguler négativement l'expression de l'oncoprotéine E6, à produire l'isoforme E6*I et à réguler la traduction de l'ORF E7 en aval. Les études détaillées en introduction portant sur l'effet de l'épissage E6*I sur la traduction de E7 ont rapporté des données contradictoires. Il a été montré que l'épissage du transcrit prématuré E6E7 entre le SD226 et le SA409 n'a pas d'effet sur la traduction de E7, c'est-à-dire que E7 est aussi efficacement traduite à partir des ARNm E6E7 que E6*IE7. Mais il a également été montré que E7 est préférentiellement traduite à partir de l'ARN non-épissé ou à l'inverse à partir de l'ARN épissé. Nous avons étudié la régulation de l'expression de la protéine E7 dans le modèle U-2 OS.

4.1 Le transcrit E6*IE7 favorise la traduction de E7

Afin d'étudier l'impact de l'épissage de l'ORF E6 sur la traduction de l'ORF E7, nous avons généré un vecteur pXJ40-E6AllE7 qui contient la séquence E6E7 sauvage d'HPV16, obtenu par clonage de l'ADNc retrouvé dans le modèle cellulaire de cancer cervical HPV16+ SiHa. Nous avons aussi généré les plasmides pXJ40-E6E7 et pXJ40-E6*IE7 en mutant le SD226 et en supprimant l'intron 1 de E6, respectivement. Le transcrit produit à partir du vecteur pXJ40-E6E7 ne peut pas être épissé et correspond au seul transcrit bicistronique capable de produire l'isoforme entière E6. Le transcrit produit par le vecteur pXJ40-E6*IE7 ne possède plus l'intron 1 de E6 et correspond donc au transcrit bicistronique épissé entre le SD226 et le SA409 que l'on retrouve majoritaire dans les cancers.

Nous avons transfecté les vecteurs pXJ40-E6AllE7, -E6E7 et -E6*IE7 dans les U-2 OS afin d'observer dans quelle proportion chaque transcrit, épissé ou non, est capable de produire l'oncoprotéine E7 (figure 36). La condition transfectée avec le vecteur pXJ40-E6AllE7 sert ici de référence.

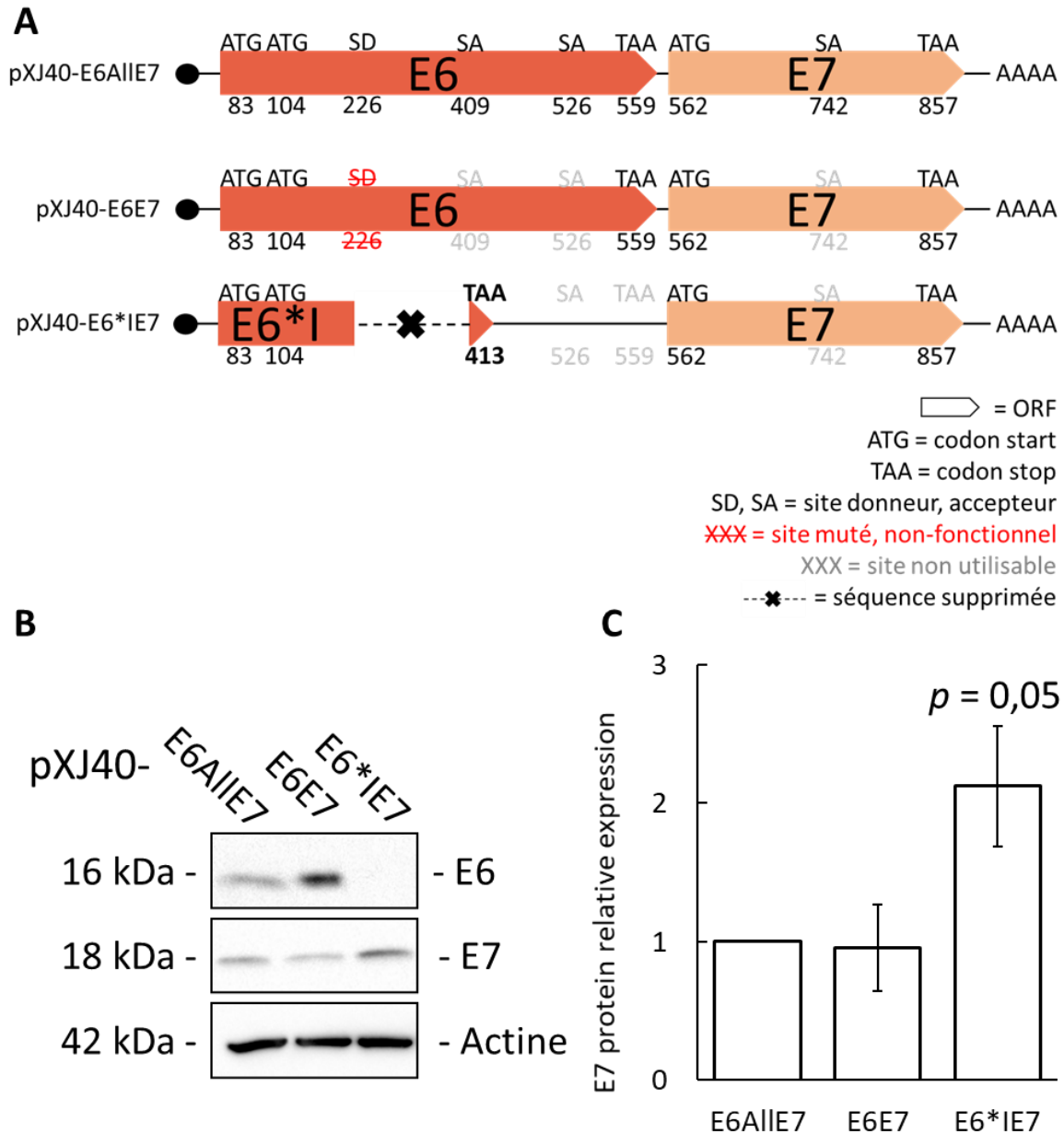


Figure 36 : Expression de la protéine E7 d'HPV16 à partir des différents transcrits E6E7 dans la lignée U-2 OS.

(A) Représentation des ARNm produits à partir des vecteurs pXJ40 utilisés. (B) Les cellules U-2 OS ont été transfectées transitoirement avec les vecteurs pXJ40-E6A1IE7, -E6E7 et -E6*IE7 pendant 24h. Les protéines E6 et E7 ont été détectées par western blotting, l'actine est utilisée comme contrôle de charge. (C) L'histogramme représente la moyenne de la quantification de l'expression de la protéine E7 normalisée par rapport à l'actine sur trois expériences indépendantes (+ ou - S.D.). Test de Mann-Whitney unilatéral.

La quantification de la protéine E7 montre une expression plus forte dans les U-2 OS transfectées avec le pXJ40-E6*IE7, qu'avec le pXJ40-E6E7. Le transcrit épissé E6*IE7 permettrait donc une meilleure traduction de E7. De manière intéressante, les transcrits primaires de la condition sauvage sont épissés à 90% et l'expression de E7 devrait donc être similaire à la condition n'exprimant que le transcrit épissé. L'absence d'expression de la

protéine E6 dans cette dernière condition, ou l'expression de l'isoforme tronquée E6*I, pourrait favoriser la traduction de E7. A l'inverse, la forte expression de la protéine E6 dans les cellules transfectées par le pXJ40-E6E7 pourrait perturber l'expression de la protéine E7 par la machinerie traductionnelle cellulaire.

4.2 Les ATG de E6 et E6*I sont préférentiellement reconnus par le ribosome

Nous avons ensuite produit des versions des vecteurs pXJ40-E6E7 et pXJ40-E6*IE7 mutés au niveau des codons d'initiation de l'ORF E6 ou E6*I afin d'étudier l'impact de la présence et de la traduction d'une première ORF avant l'ORF E7. Deux codons d'initiation ont été mutés, l'ATG83 et 104. Les deux ATG ont été rendus non-fonctionnels par mutation ponctuelle AAG. Nous avons transfecté dans les U-2 OS les vecteurs pXJ40- Δ ATG E6E7 et pXJ40- Δ ATG E6*IE7 afin d'étudier l'impact de la présence d'une ORF en amont de E7, qu'elle soit épissée ou non (figure 37). Les vecteurs pXJ40-E6E7, -E6*IE7 et -E7 servent de référence.

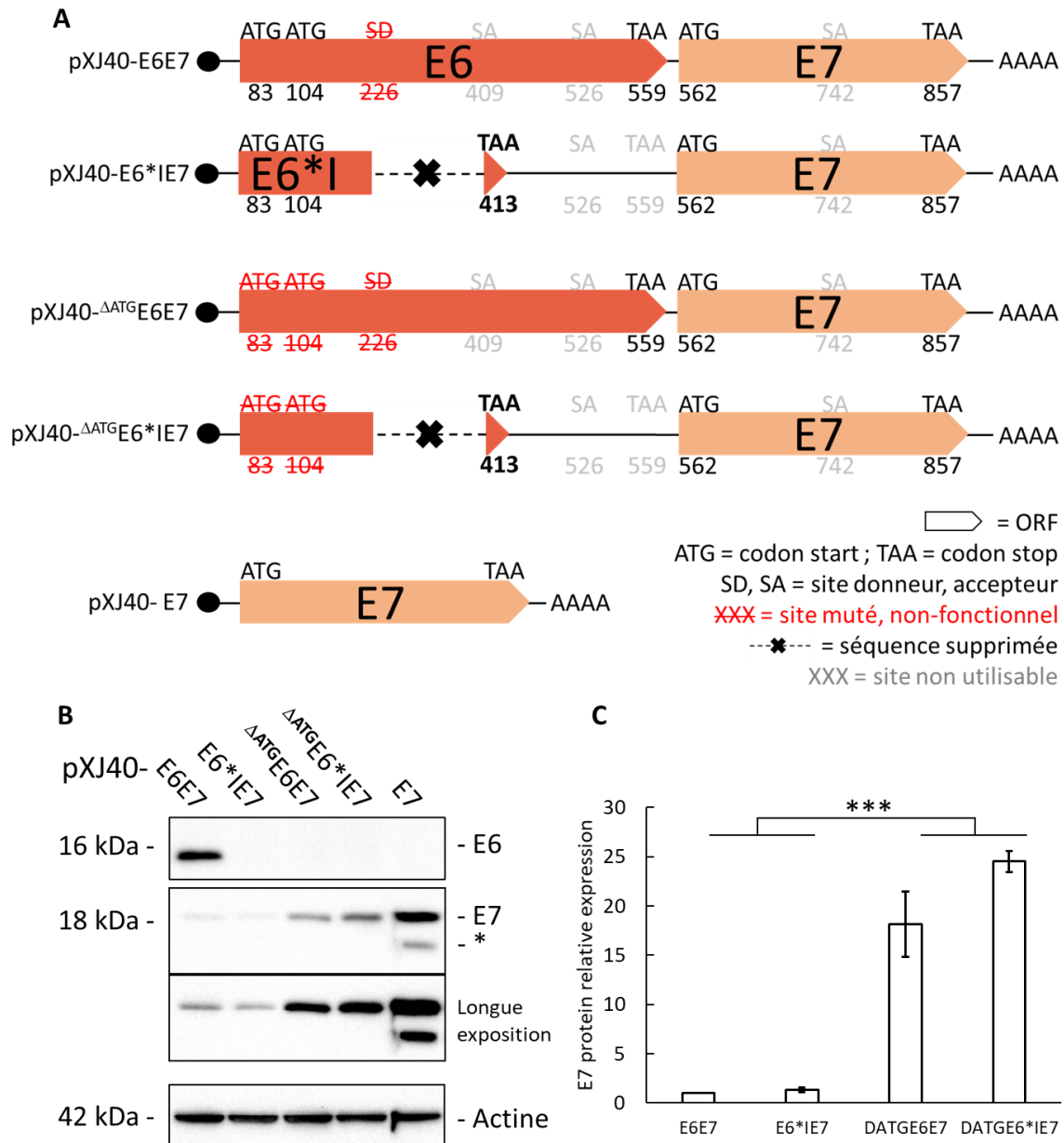


Figure 37 : Expression de la protéine E7 d'HPV16 à partir des transcrits E6E7 sans codon d'initiation pour E6 ou E6*I.

(A) Représentation des ARNm produits à partir des vecteurs pXJ40 utilisés. (B) Les cellules U-2 OS ont été transfectées transitoirement avec les vecteurs pXJ40-E6E7 et -E6*IE7, ou avec leur version mutée sur les ATG de E6, pXJ40- Δ ATGE6E7 et - Δ ATGE6*IE7 pendant 24h. La transfection du pXJ40-E7 sert de contrôle positif. La protéine E7 a été détectée par western blotting, La bande en dessous de E7 désignée par un astérisque correspond à une isoforme E7 traduite à partir d'un AUG en aval de l'AUG consensus de E7 (Stacey et al., 2000). L'actine est utilisée comme contrôle de charge. (C) L'histogramme représente la moyenne de la quantification de l'expression de la protéine E7 normalisée par rapport à l'actine sur trois expériences indépendantes (+ ou - S.D.). *** : $p < 0,001$ après test de Kruskal-Wallis bilatéral.

La mutation des codons d'initiation de l'ORF E6 ou E6*I en amont de l'ORF E7 aboutit à une expression de E7 20 fois supérieure aux versions avec ATG (Figure 37), différence d'expression de E7 entre les deux vecteurs Δ ATG, $p = 0,05$). Le taux de protéine

E7 obtenu par la transfection du vecteur pXJ40-E7 est encore plus important que ceux obtenus avec les vecteurs mutés sur les ATG de E6 mais reste dans le même ordre de grandeur. Ces résultats confirment que la machinerie traductionnelle cellulaire privilégie fortement le premier AUG rencontré sur un transcrit. L'AUG de l'ORF E6 en amont est préférentiellement choisi par le complexe d'initiation de la traduction. De plus, les autres AUG présents dans l'ORF E6 ou E6*I sur les trois cadres de lecture ne semblent pas reconnus par le ribosome qui initie la traduction à l'AUG de E7.

L'augmentation de l'expression de E7 entre transcrit $\Delta\text{AUG}^{\text{E6E7}}$ et $\Delta\text{AUG}^{\text{E6*IE7}}$ est à la limite de la significativité. Cette augmentation pourrait indiquer que le rapprochement de l'ORF E7 par rapport à la coiffe, suite à l'épissage E6*I, favorise sa traduction. De manière intéressante, la similarité des résultats entre transcrits $\Delta\text{AUG}^{\text{E6E7}}$ et $\Delta\text{AUG}^{\text{E6*IE7}}$ indique que l'AUG de l'ORF E6*I est aussi efficacement reconnu par le ribosome.

4.3 L'espace intergénique entre E6*I et E7 est indispensable à la traduction de E7

Nous avons enfin étudié le rôle de l'espace intergénique présent entre les ORF E6*I et E7. Tang et collaborateurs ont proposé que l'augmentation de la traduction de E7 à partir du transcrit épissé puisse s'expliquer par la facilitation de la réinitiation du ribosome s'étant détaché après traduction de l'ORF E6*I (Tang et al., 2006). En effet, l'espace intergénique entre le codon STOP de E6 et l'AUG de E7 n'est que de 2 nt et ne permet théoriquement pas la réinitiation du ribosome (Tang et al., 2006). L'épissage entre le SD226 et le SA409 avec le décalage du cadre ouvert de lecture augmente grandement cet espace intergénique par l'apparition d'un nouveau codon STOP 144 nt en amont de l'AUG E7. Cet espace de 144 nt est présent sur le transcrit produit à partir du vecteur pXJ40-E6*IE7. Afin d'étudier l'impact de cet espace intergénique sur la traduction de E7, nous avons produit le vecteur pXJ40-E6*IshortE7 qui comprend l'ORF E6*I suivi de l'ORF E7 avec une séparation de seulement 12 nucléotides entre les deux. Cet espace intergénique est composé d'un site de clonage de 6 nucléotides suivi des 6 nucléotides formant la séquence kozak nécessaire à la reconnaissance du codon d'initiation de E7. Nous avons transfecté le vecteur pXJ40-E6*IE7 et le vecteur pXJ40-E6*IshortE7 dans les U-2 OS (figure 38).

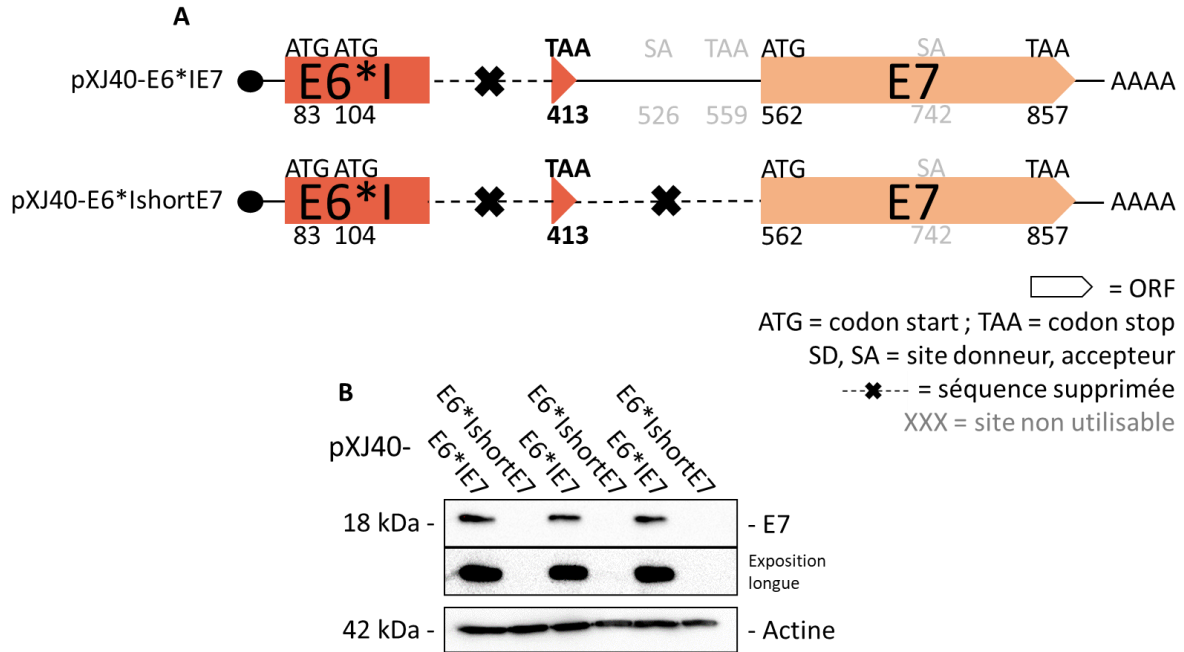


Figure 38 : Expression de la protéine E7 d'HPV16 à partir d'un transcrit E6*IE7 sans espace intergénique entre les deux ORF.

(A) Représentation des ARNm produits à partir des vecteurs pXJ40 utilisés. (B) Les cellules U-2 OS ont été transfectées transitoirement avec les vecteurs pXJ40-E6*IE7 et -E6*IshortE7 pendant 24h. La protéine E7 a été détectée par western blotting dans un triplicata biologique. L'actine est utilisée comme contrôle de charge.

Nous pouvons observer figure 37 que la réduction de l'espace intergénique entre les ORF E6*I et E7 abolit complètement la traduction de E7, indiquant l'importance de cet espace pour la traduction de E7. En revanche, cette expérience ne permet pas de statuer si l'abolition d'expression de E7 observée est due au rapprochement entre codon STOP de E6*I et AUG de E7, ou à la suppression d'éléments *cis* régulateurs présents dans l'espace intergénique et nécessaire à sa traduction. En effet, il a également été proposé que la reconnaissance de l'AUG de E7 se fasse par *leaky scanning* (Stacey et al., 2000) et l'ablation de l'espace intergénique et d'éventuelles éléments *cis* régulateurs en son sein pourrait aussi expliquer la perte d'expression de E7. En revanche, la réduction de la distance entre l'AUG de E7 et la coiffe penche théoriquement en faveur d'une traduction accrue (Zheng et al., 2004) (Davuluri et al., 2000) ce qu'on ne retrouve pas ici.

Discussion

L'étude de la biologie des HPV au cours des 30 dernières années a permis de comprendre les mécanismes du cycle viral et de la carcinogenèse associée aux HPVhr et notamment HPV16. Les rôles des oncoprotéines E6 et E7 sont bien définis dans ces deux processus. En revanche, les fonctions des protéines tronquées E6* sont difficiles à identifier et à replacer dans la réalisation du cycle viral et de la carcinogenèse. Par extension, certaines choses restent incertaines concernant le rôle et la régulation de l'épissage des transcrits E6E7. Les sites d'épissage présents dans les ORF E6 et E7 sont une particularité évolutive des HPVhr qui ne possèdent qu'un seul promoteur précoce (p97), comparés aux HPVbr possédant un promoteur pour l'ORF E6 et un promoteur pour l'ORF E7. Cet épissage permet donc de réguler l'expression de plusieurs protéines à partir d'un seul transcrit primaire. La présence de ces sites d'épissage exclusivement chez les HPVhr à l'origine des cancers induits par HPV a naturellement posé la question du lien entre cet épissage alternatif, le processus de carcinogenèse et la biologie des HPVhr en général.

L'objectif de ce travail de doctorat a été d'étudier les différents aspects de l'épissage du transcrit E6E7 d'HPV16 entre le SD226 et le SA409. Le premier constat établi dans la littérature propose que l'épissage très efficace de l'intron 1 de E6 permet de réprimer l'expression de l'oncoprotéine E6 tout en permettant l'expression des autres ORF présentes sur le transcrit mature. Les questions plus controversées de la littérature concernent le devenir de l'ORF épissée E6*I et l'impact de l'épissage sur la traduction de l'ORF E7 en aval. L'épissage entraîne un décalage du cadre ouvert de lecture faisant apparaître un nouveau codon stop, créant ainsi l'ORF E6*I. Plusieurs études ont démontré des fonctions biologiques à la protéine tronquée E6*I qui oscillent entre effets anti- ou pro-tumorales. Cependant, l'expression ou du moins la détection de la protéine tronquée est aussi très controversée. L'effet de l'épissage sur la traduction de E7 est aussi discuté dans la littérature tantôt favorisant, inhibant ou n'ayant pas d'effet sur la traduction de E7. Les travaux discutés ci-après apportent des éléments de réponse sur ces trois aspects de l'épissage.

1 Etude de l'impact de E6*I sur le transcriptome cellulaire

Notre premier objectif a été d'étudier l'impact de E6*I d'HPV16 sur le transcriptome cellulaire. Pour se faire, nous avons établi des clones stables à partir de la lignée cellulaire U-2 OS HPV négative. Les clones U-E6All expriment E6 et E6*I, et les clones U-E6*I uniquement E6*I. Le transcriptome de ces clones a été quantifié par RNAseq et comparé au transcriptome des clones U-CTL afin d'identifier des transcrits cellulaires dérégulés par l'expression ectopique des différentes isoformes de E6, et ainsi en apprendre plus sur leur rôle dans la carcinogenèse. C'est à notre connaissance, la première tentative d'étude de l'impact de E6*I sur le transcriptome cellulaire par RNAseq.

1.1 Les clones stables exprimant les isoformes E6

Cette lignée cellulaire dérivée d'un ostéosarcome semble de prime abord assez éloignée du tropisme muqueux des HPV. Plusieurs études ont cependant montré cette lignée comme convenant à l'étude des HPV et de leur métabolisme ARN (Geimanen et al., 2011) (Toots et al., 2014). Elle est aussi largement utilisée dans les études sur 16E6*I et son impact sur le métabolisme des ROS (Chen Wongworawat et al., 2016). Pour finir, l'expression dans cette lignée de p53 et pRb sauvage la rend propice à l'étude des fonctions de E6 et E7. Nous avons donc transfecté de manière stable nos différentes constructions pXJ40-vide, pXJ40-E6All et pXJ40-E6*I dans cette lignée. Les clones obtenus ont été sélectionnés dans un premier temps par l'expression au niveau ARN des différents transcrits E6. Les clones U-E6All expriment bien les transcrits E6, E6*I et E6*II et la proportion de chaque transcrit est similaire aux lignées HPV positives SiHa et Ca Ski, avec une expression très forte de l'ARNm E6*I témoignant d'un épissage entre le SD226 et le SA409 très efficace, et une expression faible des ARNm E6 et E6*II. De plus, les clones U-E6All et U-E6*I expriment des niveaux similaires d'ARNm E6*I. Ensuite, les clones U-E6All ont été sélectionnés en fonction de la dégradation de p53 par rapport aux clones contrôle, témoignant de la fonctionnalité de la protéine E6. En revanche, la plus forte expression de E6 dans les clones U-E6All comparée aux lignées HPV16+ entraîne pourtant une plus faible dégradation de p53. En effet, p53 est indétectable dans les SiHa et Ca Ski alors que la protéine E6 est elle-même quasiment indétectable. La forte expression de E6 ne suffit pas à complètement dégrader p53 dans les clones U-E6All, ce qui semble indiquer une limite du modèle U-2 OS. Pour finir, la protéine E6*I est indétectable que ce soit dans les clones U-E6All ou U-E6*I. La détection de l'isoforme tronquée n'est possible qu'en transfection transitoire après traitement au MG132.

Les causes probables de cette absence de détection ont été étudiées séparément et sont traitées plus loin dans la discussion.

Nous n'avons pas pu établir de clones stables exprimant uniquement E6 en transfectant les U-2 OS avec le vecteur pXJ40-E6, muté au niveau du SD226 et ne pouvant pas être épissé. Une condition exprimant uniquement E6 aurait permis d'attribuer directement les effets transcriptomiques à chacune des isoformes de E6, avec la condition U-E6All « physiologique » et les conditions U-E6 et U-E6*I montrant les effets propres à chaque isoforme. La mort des cellules transfectées stablement par le pXJ40-E6 est le reflet de la première fonction de l'épissage de l'intron 1 de E6. Il semble en effet que l'épissage efficace des transcrits primaires réduise drastiquement le taux d'ARNm codant E6. L'épissage assure donc une expression minimale de E6, suffisante pour réaliser ses fonctions pléiotropes sans être cytotoxique. L'utilisation d'un système inductible, dont l'expression de E6 serait contrôlable comme utilisé dans (Filippova et al., 2004), pourrait permettre de surpasser ce problème de létalité à l'établissement des clones stables. L'absence de cette condition nous a quand même permis d'étudier les effets propres à E6*I et les interactions entre E6 et de E6*I.

1.2 Modifications phénotypiques induites par les isoformes E6

Nous avons ensuite caractérisé les clones stables par divers tests fonctionnels afin de relier le phénotype des différentes conditions aux transcrits dérégulés identifiés par RNAseq. En accord avec la littérature, l'expression de E6*I ou des deux isoformes E6 et E6*I n'entraîne pas de modification de la prolifération cellulaire (Pim et al., 1997b) (Hawley-Nelson et al., 1989). En revanche, l'expression ectopique de E6*I diminue la capacité des U-2 OS à former des colonies. La clonogénicité des U-E6All est aussi réduite dans une moindre mesure. L'expression de l'ORF E6 sauvage a déjà été montrée comme diminuant la clonogénicité (Ristriani et al., 2009) mais pas l'expression de E6*I seul. A l'inverse, l'expression de E6 (Marullo et al., 2015) comme de E6*I (Williams et al., 2014) ont été montrées comme induisant une surproduction cellulaire de ROS. Mais, la quantification des niveaux d'H₂O₂ produits par nos différents clones montre que seule l'expression de E6*I entraîne une surproduction de ROS. Les effets des protéines virales sur le métabolisme des ROS sont fortement modèle dépendant (Cruz-Gregorio et al., 2018a). Nous reviendrons ci-après plus en détail sur le métabolisme des ROS. Il est communément admis dans la littérature que la dérégulation d'une voie de réponse au stress ou génératrice de stress cellulaire va impacter la capacité des autres voies à retourner à l'homéostasie cellulaire (Marullo et al.,

2015) (Filippova et al., 2014b). Pourtant, la surproduction de ROS des clones U-E6*I ne les sensibilise pas plus que les clones U-CTL au rayonnement ionisant. A noter que l'on observe tout de même une augmentation de la radiosensibilité des clones U-E6*I par rapport aux U-CTL pour la dose de 8 Gy, un fait déjà observé par (Pang et al., 2011). Un stress osmotique au sorbitol n'a pas non plus montré d'effet significatif sur la viabilité des différents clones. En revanche, la coexpression de E6 et E6*I sensibilise les cellules à la radiothérapie. E6 inhibe plusieurs effecteurs de la réparation de l'ADN et la charge mutationnelle qui en résulte, additionnée de celle causée par les rayons, diminue la viabilité cellulaire (Wallace et al., 2017). Pour compléter la caractérisation des clones, il serait intéressant de mesurer la viabilité des clones U-E6*I exposés à divers chimiothérapies car la surproduction de ROS devrait théoriquement s'accompagner d'une sensibilisation aux drogues (Filippova et al., 2014b). Nous observons donc que chaque isoforme a des capacités propres à déréguler le phénotype cellulaire et que l'expression simultanée de E6 avec E6*I entraîne la perte de certains effets propre à E6*I.

1.3 Dérégulations du transcriptome cellulaire induites par les isoformes E6

Sur les 53 gènes dérégulés par l'expression ectopique de E6*I, seul 11 gènes sont communs avec les quelques 419 gènes dérégulés dans la condition U-E6All. Ce qui indique premièrement que les effets de E6*I sont minimes sur le transcriptome, et deuxièmement que ces effets disparaissent en partie quand l'isoforme tronquée est coexprimée avec E6. Le clustering hiérarchique des neuf clones indique également que les effets de E6*I sur le transcriptome sont minimes du fait que les profils de dérégulation des clones U-E6*I se classent avec ceux des clones U-CTL. Il semble donc que la majorité des 419 transcrits dérégulés dans la condition U-E6All le soit à cause de E6 et pas de E6*I. Les larges études de transcriptomique réalisées sur des cohortes de cancers cervicaux ou de l'oropharynx HPV16+ révèlent beaucoup de cibles communes avec notre RNAseq (Pyeon et al., 2007) (Masterson et al., 2015) (den Boon et al., 2015). SYCP2 est fortement surexprimé dans U-E6All et a été identifié dans la littérature comme biomarqueur potentiel de l'évolution des lésions précancéreuses du col. L'avantage de notre modèle est que nous n'exprimons que le gène E6 comparé à ces larges études réalisées sur des cohortes d'échantillons cliniques exprimant E6 et E7. Nous pouvons donc directement lier la dérégulation de nos cibles à E6, et apporter ainsi de nouvelles informations par rapport à la littérature (Pyeon et al., 2007). Ainsi comme (den Boon et al., 2015), nous observons dans les U-E6All la répression de l'expression de la

majorité des transcrits codant des sous-unités ribosomales, attribuable ici à la seule expression de E6. De manière plus générale, les autres cibles dérégulées par E6 sont impliquées dans la réparation de l'ADN, le cycle cellulaire, la transcription et la réponse immunitaire.

Les études précédentes portant sur E6*I ont démontré l'impact de son expression sur le métabolisme des ROS. Notre étude de transcriptomique a identifié un ensemble de transcrits dérégulés par l'expression de E6*I et codant des protéines impliquées dans le métabolisme des ROS. Notre RNAseq a aussi révélé de nouvelles voies cellulaires potentiellement dérégulées par E6*I comme l'organisation de la MEC (Matrice Extracellulaire). Des protéines à domaines PDZ impliquées dans l'organisation de la MEC ont déjà été identifiées comme dérégulées par 18E6*I. Il a été montré que son expression entraîne la dégradation de PATJ, une protéine à domaine PDZ. Ni 18E6*I, ni 16E6*I ne possèdent de PBM et il est donc proposé que 18E6*I altère la stabilité de la protéine PATJ de manière indirecte *via* un autre intermédiaire protéique (Storrs and Silverstein, 2007). Toujours en absence de E6, il a ensuite été montré que 18E6*I entraîne une diminution des taux protéiques de hDlg et MAGI-1 et, de hScrib et MAGI-2 dans une moindre mesure *via* le protéasome mais sans pour autant, ici non plus, que 18E6*I interagisse directement avec ces protéines cellulaires (Pim et al., 2009). Il serait intéressant d'étudier les mécanismes à l'origine de ces dérégulations et leur impact sur le phénotype de cellules exprimant E6*I d'HPV16 et d'HPV18.

Bien que le métabolisme des ROS ait déjà été identifié comme dérégulé par E6*I, nous nous sommes intéressés à celui-ci car certaines cibles du RNAseq n'ont jamais été décrites dans la littérature et pourraient faire partie d'un nouveau mécanisme à l'origine de cette surproduction de ROS.

2 L'impact de E6*I d'HPV16 sur le métabolisme des ROS

2.1 ROS et homéostasie cellulaire

Les ROS sont produits par divers systèmes cellulaires. Certains produisent des ROS en tant que déchet inévitable de leur fonctionnement alors que d'autres servent exclusivement à la production de ces ROS. La respiration mitochondriale est la première source de ROS en tant que produit du couplage entre oxydations phosphorylantes et ATP synthase. La stœchiométrie parfois imparfaite des réactions entraîne la formation de radicaux hydroxyles OH⁻ et de superoxyde O₂⁻ à partir de l'oxygène moléculaire servant d'« accepteur final

d'électron » et normalement converti en H₂O. La production de ROS est étroitement liée au métabolisme énergétique. Mais la cellule produit également des complexes protéiques dont l'unique rôle est de produire des ROS. Il s'agit des complexes NAD(P)H oxydases qui sont enchâssés dans les diverses membranes lipidiques de la cellule et produisent du H₂O₂ en oxydant du NADPH. La cellule produit aussi des protéines régulant les niveaux de ROS. Ces protéines sont oxydées par les ROS et permettent leur neutralisation. La cellule produit enfin des enzymes (SOD, Gpx, catalase) chargés de convertir ces espèces réactives en H₂O₂ puis en H₂O inerte. Ces différents effecteurs maintiennent une homéostasie entre production et élimination des ROS.

En condition pathologique, le taux de ROS cellulaire peut être dérégulé. Cette dérégulation peut provenir d'une surproduction de ROS et/ou d'une diminution des défenses cellulaires antioxydantes, on parle alors de stress oxydatif ou de perte d'homéostasie redox. En effet, les ROS produits en trop grande quantité ont des effets délétères sur la cellule en attaquant chimiquement protéines, lipides et acides nucléiques causant respectivement des modifications post-traductionnelles non-désirées, des dommages aux membranes et des dommages à l'ADN. Les ROS sont aussi utilisés comme second messenger impliqués dans le retour à l'homéostasie redox, mais également dans de nombreux processus physiopathologiques. Les ROS sont ainsi impliqués dans la prolifération et la survie cellulaire, l'hypoxie, les maladies neurodégénératives, le diabète, les accidents vasculaires, le vieillissement et les cancers (Davies et al., 2017). Comme la plupart des voies de réponse au stress, l'homéostasie redox est à double tranchant car elle protège l'intégrité et la viabilité des tissus sains comme des tumeurs (Valko et al., 2007).

2.2 Les oncoprotéines E6 et E7 dérégulent le métabolisme des ROS

La question de la place du stress oxydant dans l'infection et la carcinogenèse a été étudiée précocement. Il a été montré que l'induction exogène d'un stress oxydant entraîne une augmentation de la réplication virale (Schmitt et al., 1989), et que l'infection productive dépend du gradient redox tissulaire, indispensable au bon assemblage des virions (Conway et al., 2009). Il a été ensuite démontré que la réplication et la transcription virale sont favorisées par le stress oxydatif. En effet, les dommages à l'ADN provoqués par le stress oxydant mènent à l'activation des voies de réparation de l'ADN (Moody and Laimins, 2009). E6 et E7 détournent ces voies pour favoriser la transcription et la réplication virale, tout en inhibant la réparation de l'ADN, *via* la dégradation de p53 induite par E6 et l'interaction de E7 avec les

protéines des voies de réparations ATR/ATM (Gillespie et al., 2012). Le stress oxydant est donc indispensable à la réalisation du cycle viral en favorisant la réplication et l'expression des gènes viraux. Les dommages à l'ADN vont également favoriser l'intégration des épisomes viraux au génome cellulaire et favoriser l'accumulation de mutations génétiques (De Marco, 2013). De plus, favoriser la réplication et augmenter le nombre d'épisomes augmentent de fait la probabilité de leur intégration. Ce mécanisme est partagé avec l'HBV (Hépatitis B Virus) qui induit un stress oxydant pour favoriser l'intégration de son matériel génétique au génome hôte (Kgatle et al., 2017).

Les oncoprotéines virales sont aussi capables de moduler le métabolisme des ROS afin de faire face aux stress exogènes induits par le système immunitaire et endogènes induits par la carcinogénèse. Ainsi, la protéine E7 va avoir des effets antioxydants en augmentant les taux de GSH réductrice et l'activité de la Catalase (Mileo et al., 2009). De plus, le maintien d'un taux minimal de protéine p53, malgré la dégradation induite par E6, permet de conserver certaines de ces fonctions de retour à l'homéostasie (Butz et al., 1999). Les cellules SiHa HPV16+ par exemple, dépendent de p53 pour transactiver des gènes du système antioxydant en cas d'exposition exogène aux ROS (Ding et al., 2007). Qu'elles soient pro- ou antioxydantes, les propriétés de E6 et E7 participent d'une part à la survie accrue des cellules infectées et d'autre part, au processus de carcinogénèse.

2.3 E6*I augmente la production de ROS

La protéine E6*I module également le métabolisme des ROS. Nos données et celles de la littérature montrent que l'expression ectopique de E6*I d'HPV16 induit une surproduction de ROS et ce dans divers modèles cellulaires. Cette surproduction semble médiée par une diminution des protéines antioxydantes GSH, GPx et SOD2 (Williams et al., 2014) et la dérégulation de protéines de la respiration mitochondriale (Evans et al., 2016). A noter que la première étude à s'intéresser au lien entre E6*I et le stress oxydant a cependant montré une corrélation entre son expression et une résistance à un stress oxydant exogène en augmentant GSH, GPx et Catalase (Mouret et al., 2005). L'étude ne prend pas en compte les effets de E7 également exprimé dans leurs modèles (Mileo et al., 2009). Notre RNAseq n'a pas révélé de transcrits dérégulés par E6*I codant ces protéines ce qui semble indiquer une régulation au niveau post-traductionnelle de la GSH, GPx et SOD2 par E6*I.

En revanche, le RNAseq révèle que l'expression de E6*I augmente l'expression de PDGFB, RAC2 et CCL2, trois cibles impliquées dans l'activation du complexe Nox2 (Chen et al., 2007) (Kietzmann and Görlach, 2005) (Cruz-Gregorio et al., 2018b) (Chun et al., 2015) (Chen et al., 2004). Nous avons donc étudié la possible implication des complexes Nox et de Nox2 en particulier dans l'augmentation des ROS provoquée par E6*I. Pour se faire, nous avons utilisé un inhibiteur chimique spécifique des complexes Nox aux doses utilisées, le DPI (Reis et al., 2020) (Augsburger et al., 2019). Nous avons dans un premier temps traité les clones stables avec 10 μ M de DPI afin d'inhiber l'activation de tous les complexes Nox. Nous avons observé une réduction du taux d' H_2O_2 produit de l'ordre de 20% dans les clones U-E6*I indiquant l'implication de ces complexes dans l'induction du stress oxydant. Le traitement au DPI 10 μ M ne permet cependant pas un retour du taux d' H_2O_2 au niveau basal des clones U-CTL ou U-E6All. Les complexes Nox sont donc impliqués dans l'augmentation des ROS induite par l'expression de E6*I, mais d'autres effecteurs semblent également participer à l'augmentation du taux de ROS dans notre modèle. Il se peut que le taux de ROS soit aussi augmenté *via* l'inhibition de protéines antioxydantes par E6*I, comme cela a déjà été décrit dans la littérature (Williams et al., 2014). Nous avons ensuite traité les clones U-E6*I avec 1 μ M de DPI afin d'inhiber spécifiquement le complexe Nox2. Nous observons une baisse non-significative du taux d' H_2O_2 de l'ordre de 10% comparée aux clones U-E6*I non-traités. Cela semble indiquer que l'activation du complexe Nox2 n'est pas l'acteur majeur à l'origine du stress oxydant induit par E6*I. En revanche, cette expérience seule ne nous permet pas de conclure sur l'implication du complexe Nox2. Des expériences complémentaires sont en cours au laboratoire pour déterminer si la surexpression de CCL2, RAC2 et PDGFB active Nox2 et si ce complexe participe effectivement au stress oxydant dans notre modèle.

Notre analyse RNAseq ne fait pas ressortir le gène CYBB codant Nox2, ni les gènes codant les protéines des autres membres de la famille Nox, comme dérégulé par l'expression de E6*I. Aucun transcrite dérégulé par E6*I ne code d'ailleurs de protéines impliquées directement dans la génération ou la catalyse des ROS. Cela indique que l'expression de E6*I augmente la production des ROS en activant ces complexes mais sans induire leur surexpression. Le complexe Nox2 a déjà été identifié comme à l'origine d'une part importante du stress oxydant retrouvé dans les principaux modèles cellulaires de cancers de l'oropharynx HPV16+ (Marullo et al., 2015). Dans ces lignées, le traitement par 1 μ M de DPI ou par un shRNA anti-NOX2 entraîne une diminution du taux de ROS intracellulaire de l'ordre de 50%.

Les auteurs de cette étude proposent d'ailleurs que E6*I puisse jouer un rôle dans l'activation de Nox2 du fait que les lignées étudiées expriment toutes très fortement E6*I par rapport à E6. En revanche, cette étude montre également que les transcrits codant les autres complexes Nox (Nox 1, 3, 4 et 5) ne sont pas exprimés dans les lignées de cancer de l'oropharynx HPV16+. Il se peut que dans d'autres modèles, du col de l'utérus ou notre modèle U-2 OS par exemple, les autres complexes Nox puissent être exprimés et activés par E6 ou E6*I pour participer au stress oxydant. L'expression des deux derniers membres de la famille Nox n'a pas été mesurée dans ces lignées de l'oropharynx. Or, les transcrits codant Duox1 et Duox2 (ainsi que leur protéine de maturation DuoxA1 et DuoxA2) ont été montrés comme surexprimés dans des tumeurs de l'oropharynx HPV16 + exprimant fortement l'ARNm E6*I par rapport à des tumeurs exprimant plus faiblement l'ARNm E6*I (Zhang et al., 2016). Ces deux protéines pourraient participer, avec l'activation de Nox2, à la signature de stress oxydant retrouvée dans ces tumeurs exprimant fortement E6*I et dans les lignées cellulaires de cancers de l'oropharynx induites par HPV16. Il serait également intéressant d'étudier leur implication dans notre modèle U-2 OS. En effet, l'activité de ces deux protéines est inhibée par le DPI à une concentration supérieure à celle de Nox2 (IC₅₀ à 2 µM contre 0,1 µM pour Nox2). Cela peut potentiellement expliquer pourquoi nous observons une diminution significative du taux d'H₂O₂ dans les U-E6*I traités avec 10 µM de DPI, comparés aux U-E6*I traités avec 1 µM.

2.4 Le rôle de E6*I dans la carcinogenèse

Mais alors quel est l'impact de E6*I et du stress oxydant dans l'infection et la carcinogenèse ? Il a d'abord été montré que l'expression de E6 et E7 induit un stress oxydant menant à l'accumulation de dommages à l'ADN et instaurant un milieu pro-inflammatoire qui corrèle avec l'évolution des lésions précancéreuses (Williams et al., 2010) (Visalli et al., 2016). Cet état pro-inflammatoire a d'ailleurs été spécifiquement lié aux complexes Nox (Martinon, 2010). Il a ensuite été montré que E6*I seul suffit à induire les ROS et à provoquer des dommages à l'ADN (Williams et al., 2014) menant à une facilitation de l'intégration (Chen Wongworawat et al., 2016). Ces données fondamentales ont été validées dans des échantillons de cancers de l'oropharynx HPV16+ stratifiés en fonction du ratio de l'expression E6/E6*I. Il a été montré par RNAseq l'implication d'E6*I dans l'induction du stress oxydant, de l'inflammation, de l'intégration et de la charge mutationnelle dans ces tumeurs. A l'inverse les tumeurs exprimant plus fortement E6 sont préférentiellement

impliquées dans la TEM et l'échappement immunitaire (Zhang et al., 2016) (Qin et al., 2020). Nos données (Paget-Bailly et al., 2019) et celles de (Filippova et al., 2007) indiquent aussi que l'expression ectopique de 16E6*I induit une augmentation du taux de protéine p53. Il est plausible que E6*I servent à garder un taux minimal de p53 au cours du cycle cellulaire pour favoriser l'amplification des épisomes. Il a d'ailleurs été montré que l'expression d'une protéine 18E6*I taguée est dépendante du cycle cellulaire et est multipliée par trois en phase G2/M (Guccione et al., 2004). Il est également possible que l'induction du stress oxydant par E6*I entraîne l'expression de p53 soit pour tenter de retourner à l'homéostasie cellulaire, soit pour provoquer la mort des cellules en condition de stress oxydant. En effet, p53 transactive de nombreux gènes codant des protéines antioxydantes tel GPX1 et SOD et des protéines impliquées dans la réparation des dommages à l'ADN, mais p53 transactive également des gènes codant des protéines régulant négativement le cycle cellulaire (Desaint et al., 2004), des protéines pro-oxydantes et des protéines pro-apoptotiques (Liu et al., 2008) (Shi and Dansen, 2020). Nous n'avons cependant pas observé de différence de viabilité ou de prolifération entre les conditions U-CTL, U-E6All et U-E6*I. Il serait intéressant d'étudier plus en détail cette induction de p53 afin de savoir s'il s'agit d'une réponse cellulaire au stress oxydant ou alors d'un mécanisme induit par E6*I pour favoriser le cycle viral.

E6*I favoriserait donc les cassures doubles brins des génomes viraux et cellulaires grâce aux stress oxydant et induirait potentiellement la restauration partielle et ponctuelle de p53 pour favoriser la recombinaison, menant à l'intégration des épisomes viraux au génome cellulaire. Les mécanismes cellulaires détournés par E6*I d'HPV16, dans un premier temps pour favoriser la maintenance et l'amplification des épisomes viraux, serviraient au final à favoriser la carcinogénèse.

3 Expression de la protéine E6*I

3.1 Détection de la protéine E6*I

La protéine E6*I est très difficile à détecter. Sa détection n'est possible qu'en transfection transitoire et après inhibition du protéasome par le MG132. E6*I n'est pas détectable dans les clones stables U-E6All et U-E6*I malgré ce traitement au MG132. De même, E6*I est indétectable dans les lignées cellulaires de cancers cervicaux HPV16+ SiHa et Ca Ski (données non montrées). E6 est elle-même difficilement détectable en expression endogène alors même que ces effets cellulaires sont plus qu'établis. E6*I d'HPV16 a

précédemment été détectée avec l'anticorps 6F4 anti-16E6N-ter (Giovane et al., 1999) dans des clones stables U-2 OS exprimant E6*I et après traitement au MG132 (Filippova et al., 2009). E6*I d'HPV18 a également été détectée en transfection transitoire et dans des clones stables à l'aide d'un anticorps spécifique de E6 d'HPV18 (Arborvita), sans traitement au MG132 cette fois ci (Pim et al., 2009). Cette détection sans nécessité d'inhiber le protéasome peut s'expliquer par les modèles cellulaires utilisés, par un meilleur taux d'expression de la protéine E6*I d'HPV18 ou par l'anticorps utilisé plus sensible. Il se peut que notre difficulté à détecter E6*I d'HPV16 dans divers modèles cellulaires s'explique simplement par la faible sensibilité des anticorps utilisés. Un autre anticorps (clone 4C6) reconnaissant aussi la partie N-terminale de E6 pourrait permettre la détection de E6*I mais n'a pas été testé car il présente une moins bonne affinité pour E6 que l'anticorps 6F4 (Masson et al., 2003).

L'ARNm E6*I, du fait de sa formation par épissage alternatif, est très proche de l'ARNm non-épisé E6, ajouté à l'efficacité de l'épissage qui aboutit à son expression 10 fois supérieure à l'ARNm E6, l'on comprend encore moins les difficultés rencontrées pour détecter la protéine E6*I. Des facteurs cellulaires reconnaissant E6*I pourraient assurer sa dégradation rapide, potentiellement plus rapide que pour E6 du fait de sa nature polypeptidique et peu structurée. Cela expliquerait aussi pourquoi E6*I n'est détectable qu'après inhibition du protéasome. L'intron retenu sur le transcrit E6 peut aussi potentiellement contenir des éléments *cis* régulateurs favorables à la traduction de l'isoforme entière. En revanche, la transfection d'une construction E6*I-tag ou E6-tag révèle une expression non seulement détectable de E6*I et E6 mais aussi dans un même ordre de grandeur *via* un anticorps anti-flag pour E6*I d'HPV16 (Filippova et al., 2014a) ou anti-HA pour E6*I d'HPV18 (Guccione et al., 2004). Aucune donnée de la littérature scientifique ne laisse penser que l'ajout d'un tag puisse drastiquement modifier une expression protéique. Le développement de nouveaux anticorps monoclonaux pourrait permettre d'étudier la protéine E6*I endogène (Togtema et al., 2019).

3.2 La protéine E6*I induit la production de ROS

Néanmoins, la difficulté à détecter E6*I nous a poussé à nous demander si l'augmentation du taux d'H₂O₂ mesurée dans les U-E6*I, et par extension la dérégulation des cibles du RNAseq, puisse provenir directement d'un effet de l'ARNm E6*I. Nous avons donc muté le vecteur pXJ40-E6*I au niveau des codons d'initiation ATG83 et ATG104 afin qu'ils ne puissent exprimer que l'ARNm E6*I. La transfection transitoire révèle que seul

l'expression du vecteur pXJ40-E6*I induit une augmentation du taux d'H₂O₂ produit par rapport au contrôle, là où la transfection du vecteur pXJ40- Δ ATG E6*I n'engendre aucune modification. Ce résultat indique que la protéine E6*I d'HPV16 est bien à l'origine de l'augmentation du taux d'H₂O₂, et du stress oxydant décrit dans la littérature (Chen Wongworawat et al., 2016).

3.3 Etude de la traduction de l'ARNm E6*I par profilage polysomique

Nous avons enfin étudié l'expression de E6*I par la méthode du profilage polysomique. Cette technique a été premièrement utilisée pour étudier des variations de la traduction dans son ensemble en comparant les profils de polysomes de plusieurs conditions. Les profils polysomiques (courbe DO260) réalisés sur les clones stables ne semblent pas indiquer de modification significative de la traduction induite par l'expression des deux isoformes E6 ou de seulement E6*I (données non montrées).

Cette technique est également utilisée pour étudier la traduction d'un ARNm spécifique. Sa variante plus moderne en association avec du RNAseq permet d'étudier le translatome, qui représente l'ensemble des ARNm engagés dans la traduction et avec quelle efficacité. Le profilage polysomique suivi d'un RNAseq sur les fractions obtenues permet de savoir quel ARNm est spécifiquement traduit et apporte ainsi une meilleure corrélation avec le phénotype que le RNAseq seul (Chassé et al., 2017) (Schwanhäusser et al., 2011). Cette approche holistique a été utilisée pour étudier l'effet de E6 d'HPV18 notamment, et montre par exemple que si l'expression de E6 inhibe bien la transcription de p21, elle entraîne de surcroît une baisse de la traduction du pool d'ARNm p21 restants (Zhao et al., 2019).

Nous avons ici utilisé cette technique pour étudier spécifiquement la traduction du transcrit E6*I en nous servant du transcrit E6 comme référence, car son expression protéique est facilement détectable dans les transfections transitoires et les clones stables. Cette approche spécifique a déjà été utilisée pour étudier des variations de la traduction de E7 d'HPV16 (Oh et al., 2006). Nous observons dans les transfections transitoires comme dans les clones stables que les transcrits E6 et E6*I présentent un profil d'enrichissement plus proche des transcrits non-codants que des transcrits cellulaires efficacement traduits. Il est difficile de conclure définitivement sur ces résultats étant donné l'absence d'un véritable contrôle négatif. En effet, 50% du taux de transcrits non-codants quantifiés est retrouvé associé à des ribosomes. Beaucoup de transcrits non-codants s'associent aux ribosomes du fait qu'ils

présentent de nombreuses caractéristiques communes avec les ARNm. De plus, certains transcrits prétendument non-codants présentent des ORF (Carlevaro-Fita et al., 2016).

De manière intéressante enfin, les transcrits E6*I des clones U-E6All semblent moins pris en charge par la machinerie traductionnelle que les transcrits E6*I exprimés dans les clones U-E6*I. En accord avec ce résultat, la protéine E6*I est détectable après inhibition du protéasome dans les cellules transfectées avec le pXJ40-E6*I, mais elle ne l'est pas dans les cellules transfectées avec le pXJ40-E6All. La traduction facilitée de E6*I dans les clones U-E6*I comparés aux clones U-E6All peut potentiellement expliquer pourquoi les effets de E6*I sur le transcriptome et le phénotype des U-E6*I ne sont pas retrouvés dans les clones U-E6All exprimant E6.

Le fait que le transcrits E6 présente un enrichissement réduit dans les fractions traduites proches des ARN non-codant alors même que son expression protéique est démontrée peut s'expliquer par l'utilisation de codons rares dans le génome des HPV (Zhao and Chen, 2011). Cette inadéquation entre les codons des ORF E6 et E6*I et le pool cellulaire d'ARNt limite leur traduction et peut expliquer les profils d'enrichissement obtenus après profil polysomique (Hanson and Coller, 2018). Cela pourrait également expliquer pourquoi (Shally et al., 1996) ont pu exprimer la protéine E6*I dans un système végétal mais pas animal, du fait du contexte différent en ARNt. L'étude des biais de codons est en train de révolutionner la conception de « code génétique dégénéré » selon laquelle plusieurs codons servent seulement à coder un acide aminé. Pour un même acide aminé, ces différents codons servent en fait à réguler la stabilité des ARNm, la traduction et le repliement des protéines (Hia et al., 2019) (Yu et al., 2015).

L'utilisation de codons rares comme facteur limitant de la traduction de E6 (et E6*I) ne devrait intuitivement pas pouvoir être contrebalancée par une présence massive de transcrits E6, comme c'est le cas dans nos transfections stables ou transitoires. Si le contexte cellulaire en ARNt n'est pas activement modifié par E6 (ou E7), on peut supposer que l'expression massive des oncoprotéines vu en blot doit automatiquement titrer le pool d'ARNt correspondant aux codons rares. Les ARNt accaparés pour produire E6 et E7 doivent l'être au détriment de la production de protéines cellulaires nécessitant ces même ARNt. Il est même étonnant que nos systèmes puissent finalement traduire autant d'ARNm suboptimaux si les ARNt nécessaires sont quantitativement peu présent dans la cellule. Ce constat implique

qu'une partie des effets observés dans notre RNA-seq pourrait résulter non pas d'un rôle direct de E6 mais indirect, modifiant le protéome cellulaire par sa forte traduction.

L'« humanisation » des séquences E6 et E6*I en remplaçant leurs codons rares par les codons les plus utilisés par les gènes cellulaires pourrait représenter une approche innovante pour trancher la question concernant l'expression et la détection de la protéine E6*I.

4 Relation dynamique entre E6 et E6*I

Nos données et celles de la littérature indiquent que E6 et E6*I ont chacune des propriétés propres. L'expression concomitante des deux isoformes aboutit en revanche à la perte des effets de l'une ou l'autre.

D'une part, les effets propres à E6*I ne sont pas retrouvés dans des cellules exprimant E6*I et E6. Nous observons en effet, l'absence d'induction de stress oxydant et des gènes cibles associés dans les U-E6All par rapport aux U-E6*I. L'expression transitoire de E6 seul n'a pas non plus d'effet significatif sur les ROS dans notre modèle U-2 OS, comme précédemment montré (Williams et al., 2014). De plus, cette étude montre dans les SiHa que l'inhibition de E6 comme la surexpression de E6*I entraîne un stress oxydant. Cela implique que l'absence de surproduction de ROS dans les U-E6All vient d'une inhibition directe de E6*I par E6 et non pas d'effets opposés propre à chaque isoforme sur des effecteurs du métabolisme des ROS. La perte des effets de E6*I pourrait provenir d'une inhibition de la traduction de E6*I provoquée par E6, comme le laisse penser nos résultats de profilage polysomique dans les clones stables et de détection de E6*I par western blotting en transfection transitoire.

D'autre part, des effets propres à E6 sont également inhibés par l'expression ectopique de E6*I. Il a en effet été montré que 18E6*I (Pim and Banks, 1999) et 16E6*I (Filippova et al., 2007) inhibe la dégradation de p53 induite par E6, des données confirmées par (Mesplède et al., 2012) pour les protéines E6*I de ces deux génotypes. E6*I stabilise aussi la procaspase 8, à l'inverse de E6 (Filippova et al., 2007). Le cas des effets sur l'apoptose démontre à quel point cette relation entre les deux isoformes est complexe et dépend du niveau d'expression de chaque isoforme. Quand elle est faiblement exprimée, E6 inhibe l'apoptose induite par le TNF en s'associant avec le récepteur de mort TNFR1 et en induisant la dégradation de la procaspase 8. E6*I se lie aussi à la procaspase 8 mais protège également les cellules de

l'apoptose induite par le TNF. En revanche l'expression des deux isoformes et le ratio d'expression entre E6 et E6*I sensibilise à la mort par le TNF via la formation d'une plateforme protéique activant l'apoptose, plus E6 est faible par rapport à E6*I, plus les cellules sont sensibles à la mort (Filippova et al., 2009).

Ces études montrent la nécessité de prendre en compte le ratio d'expression des deux isoformes quand on étudie les effets de leur expression. La littérature justifie souvent les effets contradictoires de E6 et E6*I en invoquant des effets modèles dépendants, ou l'absence d'effet de E6*I par rapport à E6 (Cruz-Gregorio et al., 2018b). La majorité des études moléculaires s'intéressant aux effets de l'oncoprotéine E6, c'est-à-dire plusieurs centaines, prennent d'ailleurs rarement en compte les produits d'épissage et les protéines tronquées pourtant exprimées dans les modèles utilisés. Les études moléculaires et les données cliniques de (Zhang et al., 2016) et (Qin et al., 2020) montrent pourtant clairement que E6*I d'HPV16 joue également un rôle d'oncoprotéine et que son expression et ses effets corrélient avec un pronostic défavorable dans les cancers de l'oropharynx induits par HPV16. Il est maintenant nécessaire de valider ces observations récentes dans de plus grandes cohortes, incluant notamment des échantillons de cancers cervicaux.

Nous avons tenté d'étudier les effets de E6 et E6*I dans une cohorte d'échantillons cliniques de cancers de l'oropharynx induits par HPV16. Notre objectif était de stratifier les tumeurs en fonction du ratio d'expression de E6*I/E6 afin de voir son impact sur les données cliniques ou sur la dérégulation de transcrits cellulaires. Malheureusement, les échantillons à notre disposition sont très homogènes au regard de ce ratio et nous n'avons pas observé de dérégulation de gènes cellulaires attribuables aux effets de l'expression de E6*I dans cette cohorte réduite de 23 échantillons, à la différence de (Zhang et al., 2016) et (Qin et al., 2020). Nous avons cependant retrouvé une dérégulation similaire entre gènes cellulaires identifiés en RNAseq de la condition U-E6All et tumeurs exprimant E6 et E6*I.

5 L'épissage dans la régulation de l'expression de E7

L'impact de l'épissage entre le SD226 et le SA409 sur la traduction de l'ORF E7 est toujours débattu dans la littérature. Il a été proposé que la reconnaissance de l'AUG de l'ORF E7 par leaky scanning permette la production de l'oncoprotéine de manière égale à partir de tous les transcrits alternatifs E6E7 (Stacey et al., 2000). Dans ce cas, la traduction dépendante de la coiffe des transcrits E6E7 ne reconnaîtrait pas l'AUG de E6 et scannerait l'ARNm

jusqu'à l'AUG de E7. D'autres études proposent au contraire que la traduction de E7 se fasse préférentiellement à partir des transcrits épissés E6*IE7 via le mécanisme de terminaison-réinitiation. Il est proposé que l'augmentation de l'espace intercistronique favoriserait la réinitiation de la traduction au niveau de l'AUG de E7 après avoir fini la traduction de E6*I en amont (Tang et al., 2006). L'étude de l'impact de l'épissage de l'intron 1 sur la traduction de E7 d'HPV18 a d'ailleurs montré que l'épissage E6*I corrèle avec la traduction de E7 (Brant et al., 2019)

Afin de terminer notre étude sur les rôles de l'épissage de l'ORF E6, nous avons étudié son impact sur la traduction de E7 dans le modèle U-2 OS. Nous observons dans les U-2 OS que la protéine E7 est traduite à partir des deux principaux transcrits E6E7 et E6*IE7. Le taux de protéine E7 est deux fois supérieur lorsque l'on exprime le transcrit E6*IE7 par rapport au transcrit E6E7. Ce résultat nuance certains résultats de la littérature montrant une expression quasi exclusive de E7 à partir, soit du transcrit E6E7, soit du transcrit E6*IE7 (del Moral-Hernández et al., 2010) (Sedman et al., 1991). Cependant, les modèles d'études varient également dans la littérature. En revanche, l'expression des deux transcrits aboutit à une expression de E7 égale à l'expression du seul transcrit E6E7. Il se peut que la protéine E6 régule négativement la traduction de l'ORF E7. En effet, E6 a été montré comme interagissant avec les ARN E6E7 (Bodaghi et al., 2009).

L'expression de transcrits mutés $\Delta^{ATG}E6E7$ et $\Delta^{ATG}E6*IE7$ sur les codons d'initiation de la traduction de E6 provoque une augmentation de la traduction de la protéine E7. Cette expérience montre que les transcrits atypiques que sont les transcrits E6E7 sont pris en charge par la cellule comme la majorité des transcrits cellulaires. Le premier codon d'initiation rencontré est privilégié par rapport aux autres. E6 et E6*I sont préférentiellement traduites par rapport à E7. Cette expérience semble aussi indiquer qu'aucun autres AUG contenu dans l'ORF E6 n'a de contexte favorable pour être scanné par le ribosome.

Enfin, la traduction de E7 est impossible à partir d'un transcrit E6*IshortE7 ne possédant plus les 140 nucléotides d'espace intergénique normalement présent entre les ORF E6*I et E7. Le fait que la protéine E7 puisse être produite à partir du transcrit E6E7 possédant un très faible espace intergénique semble indiquer que la distance entre les deux ORF ne joue pas de rôle dans la traduction de E7. En revanche, la présence d'éléments *cis* régulateurs dans les 140 derniers nucléotides de l'ORF E6 semble indispensable à la traduction de l'ORF E7. Une étude récente a identifié deux séquences dans E6 en amont de l'AUG de E7 présentant

une complémentarité avec l'ARNr 18S (Auslander et al., 2019). Ces séquences s'apparentant à des IRES (Internal Ribosome Entry Sites) et favorisant le recrutement du ribosome sont importantes pour la traduction indépendante de la coiffe, mais aussi pour le leaky scanning et le mécanisme de terminaison-réinitiation, en particulier lorsque le codon stop est proche du codon d'initiation (Gritsenko et al., 2017). La délétion de ces séquences pourrait expliquer pourquoi la traduction de E7 est abolie lorsque nous transfectons le vecteur E6*IshortE7. De plus, la présence de ces séquences sur l'ARN E6E7 présentant un très faible espace intergénique n'exclut pas que la protéine E7 soit quand même traduite par le mécanisme de terminaison-réinitiation.

Ces séquences pourraient aussi expliquer comment la protéine E7 peut être produite à partir d'un circRNA (ARN circulaire). Il a en effet été montré récemment qu'HPV16 est capable de produire un circRNA contenant toute la séquence codante de E7, mais pouvant en plus traduire l'oncoprotéine *via* un mécanisme indépendant de la coiffe. Là où l'épissage alternatif fait intervenir un SD puis un SA, la génération de circRNA se fait par « back-splicing » c.a.d. l'emploi des mêmes sites d'épissage mais dans un ordre inverse (Starke et al. 2015) (X.-O. Zhang et al. 2016). Le SD880 et le SA409 sont utilisés pour générer ce circRNA E7. Ce circRNA est retrouvé dans la lignée cellulaire Ca Ski et dans les échantillons de cancers cervicaux du TCGA (J. Zhao et al. 2019). Une étude sur une cohorte réduite de cancer du canal anal (HPV16+) a également détecté ce circRNA E7 et a montré de manière surprenante que son expression corrèle avec un meilleur pronostic de survie, là où l'ARNm E6*I n'a pas de valeur prédictive (Chamseddin et al. 2019).

L'expression des oncoprotéines est difficilement quantifiable dans des échantillons cliniques et nous nous reposons donc sur l'expression ARN de E6, E6*I et E7 pour y corrélérer des effets cellulaires. Un moyen permettrait peut-être de corrélérer l'expression des différents transcrits E6E7 à l'expression protéique de E7 afin de statuer sur le rôle de l'épissage dans l'expression de l'oncoprotéine. Ne pouvant pas mesurer l'expression de l'oncoprotéine E6, les études cliniques portant sur les effets de E6*I sur le transcriptome de tumeurs de l'oropharynx HPV16+ ont utilisé une méthode permettant de définir l'expression de E6 dans les échantillons de patients (Zhang et al., 2016). L'expression de E6 a été corrélée à la dérégulation d'un groupe de 30 gènes cellulaires, identifiés comme cibles de E6. Une forte dérégulation de ces gènes dans une tumeur forme un marqueur de la forte expression protéique de E6. Nous pouvons imaginer étudier un témoin similaire de l'expression de E7 en mesurant la dérégulation de ces cibles cellulaires. Ainsi dans une tumeur donnée, nous

pourrions corrélérer le ratio ARNm E6*I/E6 à l'expression de l'oncoprotéine E7 et définir à partir de quel ARN, E6E7 ou E6*I/E7, la protéine E7 est préférentiellement traduite.

Conclusion générale

Après plus de 40 ans d'études sur le sujet, de nombreux éléments de réponse restent encore à découvrir pour comprendre les mécanismes d'expressions des oncoprotéines virales et les rôles qu'elles jouent dans la dérégulation de l'homéostasie cellulaire. Ce travail de thèse a permis de renforcer ces connaissances sur la régulation de l'expression des transcrits et protéines virales impliqués dans la carcinogenèse associée aux HPV.

Nous avons montré dans ce travail que l'expression ectopique de E6*I d'HPV16 suffit à impacter significativement le phénotype et le transcriptome cellulaire, en dérégulant le métabolisme des ROS notamment. Nous avons également montré que cette dérégulation nécessite l'expression protéique de E6*I. En revanche, la coexpression de E6 avec E6*I ne permet pas de retrouver ces effets propres à E6*I. La littérature existante tend à montrer que E6*I est capable d'inhiber les effets de E6. Notre étude à l'inverse fait ressortir la disparition des effets propres à E6*I quand E6 est exprimée avec elle. De plus, nos données de profilage polysomique semblent indiquer que E6 pourrait inhiber la traduction de E6*I. Toutes ces données, si elles nous en apprennent plus sur les fonctions et la régulation de l'expression de E6*I, ne nous permettent pas de définir un rôle clair dans la carcinogenèse. Se faisant, il est difficile d'imaginer, dans le contexte d'une infection où la protéine E6 est produite, comment E6*I pourrait avoir un rôle à jouer dans le processus de carcinogenèse. Il se peut cependant que notre incapacité à appréhender la place de E6*I dans la biologie des HPV provienne d'une incompatibilité entre la réalité biologique et nos modèles d'études.

Une approche diamétralement opposée à celle proposée dans ce travail permettrait peut-être plus directement de statuer sur l'importance de E6*I dans le cycle viral et le processus de carcinogenèse. En effet, notre approche a consisté à introduire la variable E6*I dans un système cellulaire HPV négatif et d'en mesurer les conséquences. A l'inverse, nous pourrions créer un modèle d'étude dans lequel nous introduirions toutes les variables HPV sauf E6*I, en exprimant par exemple une version du génome complet d'HPV16 incapable de produire E6*I. La comparaison des effets cellulaires de l'expression de ce génome modifié par rapport à l'expression du génome sauvage permettrait peut-être de conclure sur la place de E6*I. L'élaboration d'un tel modèle pourrait cependant ne pas être des plus simples. Car non seulement nous étudions un problème aux variables multiples, mais ces variables sont de surcroît intriquées ensembles. Il serait difficile par exemple, d'inhiber l'épissage pour empêcher l'expression de E6*I sans déréguler l'expression de E6 et possiblement de E7.

Les difficultés à détecter la protéine E6*I, ainsi que E6 et E7 dans une moindre mesure, rendent également l'étude de leurs fonctions biologiques compliquée. A ce titre, une des limites dont il faudra s'affranchir pour comprendre la place de E6*I dans le processus de carcinogenèse concerne le développement de nouveaux anticorps efficace pour reconnaître la protéine E6*I. Ce problème d'ordre technique nous empêche d'avancer sur cette problématique et a surement participé au désintérêt d'une partie de la communauté scientifique HPV concernant les isoformes tronquées de E6.

Matériel et méthodes

Culture Cellulaire. Les lignées cellulaires U-2 OS (ATCC, HTB-96), SiHa, HaCaT ont été cultivées en milieu DMEM (Lonza) supplémenté de 10% SVF (Eurobio) à 37°C, sous atmosphère humidifiée avec 5% de CO₂. Toutes les lignées ont été cultivées sans antibiotique. Pour le maintien des lignées cellulaires stables U-2 OS, le milieu est additionné de G418 à 625 µg/mL (ThermoFisher). La présence de mycoplasmes dans les cultures de cellules a été inspectée de manière routinière.

Transfection transitoire. Les cellules ont étéensemencées dans des boîtes 6 puits à raison de 200 000 cellules par puits. Après 24 heures d'incubation, les cellules ont été transfectées par 1 µg d'ADN plasmidique ou 20 pmol de siRNA en utilisant l'agent de transfection Lipofectamine 2000 (ThermoFisher Scientific), respectant un ratio ADN/agent de transfection de 2/1 et 3/1, respectivement, dans du milieu Opti-MEM (ThermoFisher Scientific) selon les recommandations du fabricant.

Prolifération cellulaire. Cinq mille cellules de chaque clone stable ont étéensemencées en plaque 24 puits. Les cellules ont ensuite été comptées à l'aide du compteur numérique EveTM automated cell counter tous les deux jours jusqu'à 14 jours. Les cellules ont été comptées une dernière fois à 17 jours.

Réponse à un stress osmotique. Deux mille cinq cent ou 5 000 cellules de chaque clone stable ont étéensemencées en plaque 96 puits opaques. Après 24h, le milieu a été supplémenté de 200mM de D-sorbitol. L'effet du stress osmotique engendré par le sorbitol sur la viabilité cellulaire a ensuite été mesuré à l'aide du kit Celltiter Glo® (Promega) selon les recommandations du fabricant de 0 à 12h. Brièvement, un volume de réactif Celltiter Glo a été ajouté au volume de milieu dans chaque puits. Les plaques ont été préalablement ramenées à température ambiante. Ensuite, les plaques ont été agitées sur un shaker orbital pendant 2 min pour lyser les cellules et mélanger le lysat au réactif. Après 10 min d'incubation à l'obscurité, la luminescence a enfin été mesurée à l'aide d'un lecteur de plaque Synergy H1 (Biotek). La viabilité cellulaire des clones stables à t₀ a été normalisée par rapport à la viabilité du clone U-CTL#par.

Test de blessure cicatrisation. Pour chaque clone stable, 300 000 cellules ont étéensemencées en plaque 12 puits. Après 24h, des blessures verticale et horizontale ont été réalisées dans chaque puits à l'aide d'une pointe de p1000. Le recouvrement de ces blessures a ensuite été observé au microscope à contraste de phase sur 24h. Les données représentent le pourcentage de recouvrement par rapport à la blessure initiale pour chaque condition.

Mesure du taux d'H₂O₂. Les taux d'H₂O₂ produit par les clones stables ou par les U-2 OS transfectées transitoirement ont été mesurés 24h après ensemencement ou 24h après transfection, respectivement, à l'aide de la trousse Amplex® Red selon les recommandations du fabricant. Concernant les expériences incluant des clones stables, chaque répliquats biologiques a consisté à déposer en triplicat le surnageant provenant de chaque clone stable. Au total, neuf mesures par condition (U-CTL, U-E6All et U-E6*I) ont donc été obtenues pour chaque expérience indépendante. Pour un clone stable donné, les mesures de production d'H₂O₂ témoin ou traité au DPI ont été réalisés sur deux ensemencements différents et les données ont été traitées comme non-appariées.

Extraction d'ARN totaux. Après lavages au PBS 1X, les cellules ont été récoltées avec 500 µL de RiboZol (AMRESCO) et les ARN totaux ont été extraits selon les recommandations du fabricant. Brièvement, 0,2 volume de chloroforme ont été ajoutés au lysat cellulaire. Après 10 min d'incubation, les lysats ont été centrifugés à 12 000 g pendant 15 min à 4°C. Les ARN totaux ont été prélevés dans la phase aqueuse et précipités dans 3 volumes d'isopropanol pendant 10 min. Les tubes ont ensuite été centrifugés à 12 000 g pendant 10 min à 4°C, le surnageant a été éliminé et les culots d'ARN ont été lavés à l'éthanol 75% à -20°C. Après centrifugation à 7 600 g pendant 5 min, les culots ont été repris dans de l'eau qualité biologie moléculaire. Pour finir, la quantité d'ARN a été dosée à l'aide du NanoDropOne (Thermo Scientific).

Séquençage ARN. Les 9 clones stables U-2 OS ont été ensemencés en flasque T75 et cultivés 72h en absence de G418 avant extraction des ARN totaux. Les ARN totaux des lignées stables ont été extraits et confiés à la société HELIXIO (Biopôle Clermont-Limagne 63360 Saint-Beauzire, France) pour séquençage et traitement des données. Les bibliothèques d'ADN complémentaire (ADNc) ont été générées à partir de 1 µg d'ARN totaux en utilisant la trousse TruSeq Stranded mRNA Sample Preparation (Illumina). Enfin, le séquençage a été réalisé à l'aide du NextSeq 500 (Illumina) par la technologie SBS deux canaux en utilisant le kit NextSeq500 High Output v2 75 cycles (Illumina). L'analyse des données de séquençage a été réalisée à l'aide des logiciels BowTie (Johns Hopkins University, v2.3.2), TopHat (Johns Hopkins University, v2.1.1), CuffLinks (Trapnell Lab, v2.2.1), SeqMonk (Babraham Bioinformatics, logiciel v1.37.1) et R (Lucent technologies, logiciel v3.3.3). Le clustering hiérarchique a été obtenu en comparant le transcriptome des 9 clones stables sans classement préalable. Les gènes différentiellement exprimés ont été identifiés en comparant la moyenne d'expression (FPKM) d'un gène donné dans la condition U-CTL par rapport à la moyenne

d'expression dans la condition U-E6All, ou dans la condition U-E6*I. Un gène est considéré comme différentiellement exprimé si la variabilité entre deux conditions est supérieure à la variabilité interne des trois clones d'une même condition.

Echantillons cliniques. Les ARN totaux des 45 échantillons cliniques HPV16+ (23 ARN E6E7+ et 22 ARN-) utilisés ont gracieusement été donnés par le Dr. A. JUNG du CRLCC Paul Strauss. Ces échantillons proviennent de carcinomes de l'oropharynx. Les conditions d'obtention et de génotypage (HPV+/HPV-) des échantillons sont détaillées dans (Jung et al., 2010).

Transcription inverse. Les ADNc ont été obtenus par transcription inverse à partir de 500 ng d'ARN totaux *via* la trousse Maxima first strand cDNA Synthesis for RT-qPCR with dsDNase (Thermo Scientific) selon les recommandations du fabricant. Brièvement, après le traitement des échantillons à la DNase, la transcriptase inverse a été ajoutée avec son tampon dans un volume final de 20 µL. Le milieu de réaction a été incubé pendant 10 min à 25°C, 15 min à 50°C et 5 min à 85°C.

PCR en point final. La présence ou l'absence des transcrits E6 dans les lignées stables a été évaluée en utilisant la trousse DreamTaq PCR Master Mix 2X (ThermoFisher) selon les recommandations du fabricant. La réaction de PCR a été réalisée dans un volume final de 25µL contenant 2,5 µL d'ADNc dilués au 1:5, le master mix 1X et les amorces E6orf_F et E6orf_R à 1 µM. L'amplification a consisté en une étape de dénaturation initiale de 2 min à 95°C, puis 35 cycles de 30 sec de dénaturation à 95°C, 20 sec d'hybridation à 60°C et 30 sec d'élongation à 72°C et enfin une étape d'élongation finale de 10 min à 72°C.

Les produits de PCR ont été repris dans du tampon de charge Gel Loading Dye, Purple 1X (NEB inc.), puis analysés par électrophorèse en gel d'agarose 2% additionné de SYBR Safe DNA Gel Stain 1X (ThermoScientific). Un marqueur de taille (ThermoScientific, GeneRuler 100 bp Plus DNA Ladder) a été déposé en parallèle des échantillons. La migration a été effectuée dans du tampon de migration TBE 1X (Tris 89 mM, Borate 89 mM, EDTA 2 mM) à 100 V pendant 30 min. Le gel a été révélé à l'aide du ChemiDoc MP System (Bio-Rad).

Test mycoplasme. La présence de mycoplasmes dans les cultures de cellules a été inspectée de manière routinière à l'aide d'une détection par PCR en utilisant les amorces universelles Forward : 5'-GTGGGGAGCAAAYAGGATTAGA-3' et Reverse : 5'-GGCATGATGATTTGACGTCRT-3'. L'amplification consistait en une dénaturation initiale

de 3 min à 94°C puis, 32 cycles de de 60 sec de dénaturation à 94°C, 30 sec d'hybridation à 55°C et de 60 sec d'élongation à 72°C, suivi d'une élongation finale de 3 min à 72°C. Les produits de PCR ont ensuite été analysés sur un gel d'agarose 1% (m/v).

PCR quantitative SYBR Green. Les taux des transcrits E6 dans les lignées stables et des cibles du séquençage ARN ont été quantifiés en utilisant la trousse Power SYBR Green PCR master mix (Invitrogen) selon les recommandations du fabricant. Le mélange réactionnel de 15 µL contenait 3 µL d'ADNc dilué au 1:40, le master mix 1X et les amorces à 200 nM (annexes, tableau V). L'amplification, réalisée à l'aide du thermocycleur StepOne Plus Real-Time PCR system (Applied Biosystems), a consisté en une étape de dénaturation initiale de 5 min à 95°C, puis 40 cycles de 15 sec de dénaturation à 95°C et 60 sec d'hybridation/élongation à 60°C. Les résultats ont été analysés selon la méthode 2^{-ddCt} à l'aide du logiciel StepOne (ThermoFisher Scientific, logiciel v2.3) et qBase+ (Biogazelle, logiciel v3.1) dans le cas des échantillons cliniques.

Extraction des protéines. Après plusieurs lavages dans du PBS 1X, les cellules ont été récoltées par décollement mécanique, dans 1 mL de PBS 1X puis centrifugées 5 min à 1500 g à 4°C. Les culots cellulaires ont été lysés en présence de tampon RIPA (Tris-HCl 50 mM pH7.4, NaCl 150 mM, NP40 1%, Déoxycholate de sodium 0,5%, EDTA 1 mM) additionné d'inhibiteurs de protéases cOmplete 1X (Roche) et de PMSF 0.5 mM (ThermoFisher) pendant 20 min à 4°C. Après une sonication de 5 sec, le lysat cellulaire a été centrifugé pendant 10 min à 10 000 g à 4°C. Le surnageant contenant les protéines a été récupéré et dosé selon la méthode de Bradford (Bio-rad, Bradford Protein Assay). La lecture de l'absorbance à 595 nm a été réalisée à l'aide du lecteur de plaque Synergy H1 (BioTek) en plaque de 96 puits.

Western Blotting. Les expériences de western blotting ont été réalisées à partir de 20 µg d'extraits protéiques totaux additionnés de tampon Laemmli (Tris-HCl 125 mM pH 6.8, SDS 2%, β-mercaptoéthanol 5%, Glycérol 10% et bleu de bromophénol). Après dénaturation pendant 5 min à 95°C, les échantillons ont été déposés sur un gel de polyacrylamide en conditions dénaturantes. Un marqueur de poids moléculaire (Bio-Rad, Precision Plus Protein Dual Color Standards) a été déposé en parallèle des échantillons. La migration a été effectuée à 100 V pendant 90 min dans du tampon de migration 1X (Tris 25 mM pH 8.5, Glycine 40 mM, SDS 0.5%). Les protéines ont ensuite été transférées sur membrane de PVDF 0.2 µm (GE Life Sciences, Amersham Hybond P 0.2 PVDF) à 100 V pendant 120 min dans du

tampon de transfert 1X (Tris 2.5 mM, Glycine 19.2 mM, méthanol 10 ou 20%). Les membranes ont ensuite été saturées dans du TBS-T-lait 5% (Tris-HCl 20 mM pH 7.4, 137 mM, Tween 20 0.1%, Régilait 5%) pendant 1h. Après lavage au TBS-T, les membranes ont été hybridées avec l'anticorps primaire (tableau II). Après 3 lavages au TBS-T, les membranes ont été hybridées pendant 1 h avec l'anticorps secondaire anti-souris (Dako, P0447) couplé à la HRP dans du TBS-T-lait 0.5%. Enfin, après 4 lavages au TBS-T et un lavage au TBS, les membranes ont été révélées à l'aide du Chemidoc en déposant préalablement 1 mL d'ECL sur la membrane (ThermoFisher Scientific, Pierce ECL Plus Western Blotting Substrate) pendant 5 min à l'obscurité.

Tableau 2 : Anticorps primaires utilisés en Western blotting.

Cible	Clone	Dilution	Incubation
p53	DO-1	1/1000	1h
E6 C-ter	E2-3F8	1/1000	2h
E6 N-ter	E2-6F4	1/1000	2h
E7	NM-2	1/1000	1h
Actine	AC-15	1/10000	1h

Analyse statistique. A moins qu'il soit précisé autrement dans les légendes des figures, les données ont été analysées à l'aide du logiciel XLSTAT (Addinsoft, logiciel v2017.2) par test non-paramétrique U de Mann-Whitney. Les valeurs $p < 0,05$ ont été retenues comme significatives.

Profilage polysomique. Un million de cellules de chaque clone stable a été ensemencé en boîte 10 cm. Ou, 500 000 cellules U-2 OS ont été ensemencées en plaque 6 puits, puis transfectées par les plasmides pXJ40-CTL, pXJ40-E6, pXJ40-E6*I et pXJ40-E7. Après 24h de transfection, les cellules ont été passées en boîte 10 cm. Après 48h d'incubation, les ribosomes ont été fixés sur les ARNm en cours de traduction en supplémentant le milieu avec 100 µg/mL de *Cycloheximide* (CHX). Après 15 min d'incubation, les cellules ont été lavées au PBS-CHX. Puis, les cellules ont été scrapées dans 2 mL de PBS-CHX. Après centrifugation à 1500 g, le culot cellulaire a été lysé dans 200 µL de tampon de lyse (NaCl 300 mM, Tris-HCl pH7,4 20 mM, MgCl₂ 10 mM, Trition X-100 1%, Sodium Déoxycholate 0,5%, DTT 1mM, AP cOmplete 1X, RNaseOUT 100 U/mL et CHX 100 µg/mL) pendant 10 min à 4°C. Le lysat a ensuite été centrifugé 2 minutes à 2000 g pour culotter les noyaux. Puis à 12 000 g pour culotter les divers débris cellulaires. Le surnageant obtenu a été délicatement déposé sur un gradient de sucrose 50-10% de 4mL (NaCl 300 mM, Tris-HCl pH7,4 20 mM, MgCl₂ 10 mM, Sucrose 50-10%, AP complete 1X, RNaseOUT 100 U/mL et CHX 100

µg/mL) après dosage au Nanodrop. La migration des polysomes en gradient de sucrose a ensuite été réalisé à 260 000 g (44 000 rpm) pendant 1h dans une ultracentrifugeuse Sorvall WX80 (ThermoFisher Scientific) et un rotor TH-660 swing (ThermoFisher Scientific) avec une phase d'accélération moyenne (réglage 6) et une phase de décélération lente (réglage 1) afin de ne pas perturber le gradient.

Ne disposant pas de fractionneur UV permettant d'analyser les gradients de sucrose de manière continue (ARNm libres, monosomes et les différents polysomes), les gradients ont été fractionnés « manuellement » à la micropipette. Au final, 80 fractions par gradient ont été délicatement prélevées et dosées au nanodrop afin d'obtenir les profils de polysomes. Après identification du pic de monosome, les 80 fractions ont été regroupées par 10 et les ARN ont été extraits des 8 fractions résultantes. Chaque fraction a été supplémentée de 5 µg d'ARN bactérien (*P. Aeruginosa*) et de 10 µg de glycogène (Invitrogen). La transcription inverse a été réalisé sur 8 µL d'ARN pour chaque fraction et les gènes d'intérêt ont enfin été analysés par qPCR. Les données présentées représentent le pourcentage d'enrichissement de chaque fraction par rapport à la quantité totale de transcrits. Les données de RT-qPCR analysées par la méthode des 2^{-ddCt} ne sont pas normalisées par rapport à un gène de référence cellulaire mais à un gène de référence bactérien, témoin de l'efficacité d'extraction et de RT (ARN ajouté en début d'extraction).

Bibliographie

- Ajiro, M., Jia, R., Zhang, L., Liu, X., and Zheng, Z.-M. (2012). Intron Definition and a Branch Site Adenosine at nt 385 Control RNA Splicing of HPV16 E6*1 and E7 Expression. *PLOS ONE* 7, e46412.
- Ajiro, M., Tang, S., Doorbar, J., and Zheng, Z.-M. (2016). Serine/Arginine-Rich Splicing Factor 3 and Heterogeneous Nuclear Ribonucleoprotein A1 Regulate Alternative RNA Splicing and Gene Expression of Human Papillomavirus 18 through Two Functionally Distinguishable cis Elements. *J. Virol.* 90, 9138–9152.
- Alonso, L.G., Smal, C., Garcia-Alai, M.M., Chemes, L., Salame, M., and de Prat-Gay, G. (2006). Chaperone Holdase Activity of Human Papillomavirus E7 Oncoprotein. *Biochemistry* 45, 657–667.
- Androphy, E. j., Hubbert, N. l., Schiller, J. t., and Lowy, D. r. (1987). Identification of the HPV-16 E6 protein from transformed mouse cells and human cervical carcinoma cell lines. *EMBO J.* 6, 989–992.
- Arbyn, M., Weiderpass, E., Bruni, L., de Sanjosé, S., Saraiya, M., Ferlay, J., and Bray, F. (2020). Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis. *Lancet Glob. Health* 8, e191–e203.
- Artaza-Irigaray, C., Molina-Pineda, A., Aguilar-Lemarroy, A., Ortiz-Lazareno, P., Limón-Toledo, L.P., Pereira-Suárez, A.L., Rojo-Contreras, W., and Jave-Suárez, L.F. (2019). E6/E7 and E6* From HPV16 and HPV18 Upregulate IL-6 Expression Independently of p53 in Keratinocytes. *Front. Immunol.* 10.
- Ashrafi, G.H., Haghshenas, M., Marchetti, B., and Campo, M.S. (2006). E5 protein of human papillomavirus 16 downregulates HLA class I and interacts with the heavy chain via its first hydrophobic domain. *Int. J. Cancer* 119, 2105–2112.
- Atamna, H., Cheung, I., and Ames, B.N. (2000). A method for detecting abasic sites in living cells: Age-dependent changes in base excision repair. *Proc. Natl. Acad. Sci.* 97, 686–691.
- Augsburger, F., Filippova, A., Rasti, D., Seredenina, T., Lam, M., Maghzal, G., Mahiout, Z., Jansen-Dürr, P., Knaus, U.G., Doroshov, J., et al. (2019). Pharmacological characterization of the seven human NOX isoforms and their inhibitors. *Redox Biol.* 26, 101272.
- Auslander, N., Wolf, Y.I., Shabalina, S.A., and Koonin, E.V. (2019). A unique insert in the genomes of high-risk human papillomaviruses with a predicted dual role in conferring oncogenic risk. *F1000Research* 8.
- Aydin, I., Weber, S., Snijder, B., Samperio Ventayol, P., Kühbacher, A., Becker, M., Day, P.M., Schiller, J.T., Kann, M., Pelkmans, L., et al. (2014). Large scale RNAi reveals the requirement of nuclear envelope breakdown for nuclear import of human papillomaviruses. *PLoS Pathog.* 10, e1004162.
- Baker, C.C., and Calef, C. (1996). Maps of Papillomavirus mRNA Transcripts. *N Hum. Papillomaviruses Compend.* 1–30.

- Baker, C.C., Phelps, W.C., Lindgren, V., Braun, M.J., Gonda, M.A., and Howley, P.M. (1987). Structural and transcriptional analysis of human papillomavirus type 16 sequences in cervical carcinoma cell lines. *J. Virol.* *61*, 962–971.
- Barbosa, M.S., Edmonds, C., Fisher, C., Schiller, J.T., Lowy, D.R., and Vousden, K.H. (1990). The region of the HPV E7 oncoprotein homologous to adenovirus E1a and Sv40 large T antigen contains separate domains for Rb binding and casein kinase II phosphorylation. *EMBO J.* *9*, 153–160.
- Barbosa, M.S., Vass, W.C., Lowy, D.R., and Schiller, J.T. (1991). In vitro biological activities of the E6 and E7 genes vary among human papillomaviruses of different oncogenic potential. *J. Virol.* *65*, 292–298.
- Barlev, N.A., Liu, L., Chehab, N.H., Mansfield, K., Harris, K.G., Halazonetis, T.D., and Berger, S.L. (2001). Acetylation of p53 Activates Transcription through Recruitment of Coactivators/Histone Acetyltransferases. *Mol. Cell* *8*, 1243–1254.
- Barreau, C., Paillard, L., and Osborne, H.B. (2005). AU-rich elements and associated factors: are there unifying principles? *Nucleic Acids Res.* *33*, 7138–7150.
- Bedell, M.A., Jones, K.H., Grossman, S.R., and Laimins, L.A. (1989). Identification of human papillomavirus type 18 transforming genes in immortalized and primary cells. *J. Virol.* *63*, 1247–1255.
- Berezutskaya, E., and Bagchi, S. (1997). The human papillomavirus E7 oncoprotein functionally interacts with the S4 subunit of the 26 S proteasome. *J. Biol. Chem.* *272*, 30135–30140.
- Berget, S.M., Moore, C., and Sharp, P.A. (1977). Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc. Natl. Acad. Sci. U. S. A.* *74*, 3171–3175.
- Berman, T.A., and Schiller, J.T. (2017). Human papillomavirus in cervical cancer and oropharyngeal cancer: One cause, two diseases. *Cancer* *123*, 2219–2229.
- Bernard, H.-U. (2002). Gene expression of genital human papillomaviruses and considerations on potential antiviral approaches. *Antivir. Ther.* *7*, 219–237.
- Bernard, H.-U., Burk, R.D., Chen, Z., van Doorslaer, K., Hausen, H. zur, and de Villiers, E.-M. (2010). Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology* *401*, 70–79.
- Blin, A. (2020). Prévenir les infections à papillomavirus humains. *Actual. Pharm.* *59*, 53–56.
- Bodaghi, S., Jia, R., and Zheng, Z.-M. (2009). Human papillomavirus type 16 E2 and E6 are RNA-binding proteins and inhibit in vitro splicing of pre-mRNAs with suboptimal splice sites. *Virology* *386*, 32–43.
- Bodelon, C., Untereiner, M.E., Machiela, M.J., Vinokurova, S., and Wentzensen, N. (2016). Genomic characterization of viral integration sites in HPV-related cancers. *Int. J. Cancer* *139*, 2001–2011.

- Bodily, J.M., Mehta, K.P.M., and Laimins, L.A. (2011). Human papillomavirus E7 enhances hypoxia-inducible factor 1-mediated transcription by inhibiting binding of histone deacetylases. *Cancer Res.* *71*, 1187–1195.
- Bonetta, A.C., Maily, L., Robinet, E., Travé, G., Masson, M., and Deryckere, F. (2015). Artificial microRNAs against the viral E6 protein provoke apoptosis in HPV positive cancer cells. *Biochem. Biophys. Res. Commun.* *465*, 658–664.
- Boon, S.S., Tomaić, V., Thomas, M., Roberts, S., and Banks, L. (2015). Cancer-Causing Human Papillomavirus E6 Proteins Display Major Differences in the Phospho-Regulation of Their PDZ Interactions. *J. Virol.* *89*, 1579–1586.
- den Boon, J.A., Pyeon, D., Wang, S.S., Horswill, M., Schiffman, M., Sherman, M., Zuna, R.E., Wang, Z., Hewitt, S.M., Pearson, R., et al. (2015). Molecular transitions from papillomavirus infection to cervical precancer and cancer: Role of stromal estrogen receptor signaling. *Proc. Natl. Acad. Sci. U. S. A.* *112*, E3255–E3264.
- Bouvard, V., Baan, R., Straif, K., Grosse, Y., Secretan, B., El Ghissassi, F., Benbrahim-Tallaa, L., Guha, N., Freeman, C., Galichet, L., et al. (2009). A review of human carcinogens-Part B: biological agents. *Lancet Oncol.* *10*, 321–322.
- Boyer, S.N., Wazer, D.E., and Band, V. (1996). E7 Protein of Human Papilloma Virus-16 Induces Degradation of Retinoblastoma Protein through the Ubiquitin-Proteasome Pathway. *Cancer Res.* *56*, 4620–4624.
- Brant, A.C., Majerciak, V., Moreira, M.A.M., and Zheng, Z.-M. (2019). HPV18 Utilizes Two Alternative Branch Sites for E6*I Splicing to Produce E7 Protein. *Virol. Sin.* *34*, 211–221.
- Bravo, I.G., and Alonso, Á. (2007). Phylogeny and evolution of papillomaviruses based on the E1 and E2 proteins. *Virus Genes* *34*, 249–262.
- Bravo, I.G., and Féllez-Sánchez, M. (2015). Papillomaviruses. *Evol. Med. Public Health* *2015*, 32–51.
- Bray, F., Loos, A.H., McCarron, P., Weiderpass, E., Arbyn, M., Møller, H., Hakama, M., and Parkin, D.M. (2005). Trends in Cervical Squamous Cell Carcinoma Incidence in 13 European Countries: Changing Risk and the Effects of Screening. *Cancer Epidemiol. Prev. Biomark.* *14*, 677–686.
- Bray, F., Lortet-Tieulent, J., Znaor, A., Brotons, M., Poljak, M., and Arbyn, M. (2013). Patterns and Trends in Human Papillomavirus-Related Diseases in Central and Eastern Europe and Central Asia. *Vaccine* *31*, H32–H45.
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer J. Clin.* *68*, 394–424.
- Burgers, W.A., Blanchon, L., Pradhan, S., de Launoit, Y., Kouzarides, T., and Fuks, F. (2007). Viral oncoproteins target the DNA methyltransferases. *Oncogene* *26*, 1650–1655.

- Burk, R.D., Chen, Z., Saller, C., Tarvin, K., Carvalho, A.L., Scapulatempo-Neto, C., Silveira, H.C., Fregnani, J.H., Creighton, C.J., Anderson, M.L., et al. (2017). Integrated genomic and molecular characterization of cervical cancer. *Nature* 543, 378–384.
- Butz, K., Whitaker, N., Denk, C., Ullmann, A., Geisen, C., and Hoppe-Seyler, F. (1999). Induction of the p53-target gene GADD45 in HPV-positive cancer cells. *Oncogene* 18, 2381–2386.
- Butz, K., Ristriani, T., Hengstermann, A., Denk, C., Scheffner, M., and Hoppe-Seyler, F. (2003). siRNA targeting of the viral E6 oncogene efficiently kills human papillomavirus-positive cancer cells. *Oncogene* 22, 5938–5945.
- Bzhalava, D., Eklund, C., and Dillner, J. (2015). International standardization and classification of human papillomavirus types. *Virology* 476, 341–344.
- Cabeça, T.K., De Mello Abreu, A., Andrette, R., De Souza Lino, V., Morale, M.G., Aguayo, F., Termini, L., Villa, L.L., Lepique, A.P., and Boccardo, E. (2019). HPV-Mediated Resistance to TNF and TRAIL Is Characterized by Global Alterations in Apoptosis Regulatory Factors, Dysregulation of Death Receptors, and Induction of ROS/RNS. *Int. J. Mol. Sci.* 20, 198.
- Carlevaro-Fita, J., Rahim, A., Guigó, R., Vardy, L.A., and Johnson, R. (2016). Cytoplasmic long noncoding RNAs are frequently bound to and degraded at ribosomes in human cells. *RNA* 22, 867–882.
- Celegato, M., Messa, L., Goracci, L., Mercorelli, B., Bertagnin, C., Spyraakis, F., Suarez, I., Cousido-Siah, A., Travé, G., Banks, L., et al. (2020). A novel small-molecule inhibitor of the human papillomavirus E6-p53 interaction that reactivates p53 function and blocks cancer cells growth. *Cancer Lett.* 470, 115–125.
- Cerasuolo, A., Buonaguro, L., Buonaguro, F.M., and Tornesello, M.L. (2020). The Role of RNA Splicing Factors in Cancer: Regulation of Viral and Human Gene Expression in Human Papillomavirus-Related Cervical Cancer. *Front. Cell Dev. Biol.* 8.
- Chassé, H., Boulben, S., Costache, V., Cormier, P., and Morales, J. (2017). Analysis of translation using polysome profiling. *Nucleic Acids Res.* 45, e15.
- Chen, J., Xue, Y., Poidinger, M., Lim, T., Chew, S.H., Pang, C.L., Abastado, J.-P., and Thierry, F. (2014a). Mapping of HPV transcripts in four human cervical lesions using RNAseq suggests quantitative rearrangements during carcinogenic progression. *Virology* 462–463, 14–24.
- Chen, J.J., Hong, Y., Rustamzadeh, E., Baleja, J.D., and Androphy, E.J. (1998). Identification of an alpha helical motif sufficient for association with papillomavirus E6. *J. Biol. Chem.* 273, 13537–13544.
- Chen, K.C.-W., Zhou, Y., Zhang, W., and Lou, M.F. (2007). Control of PDGF-induced reactive oxygen species (ROS) generation and signal transduction in human lens epithelial cells. *Mol. Vis.* 13, 374–387.

- Chen, X.-L., Zhang, Q., Zhao, R., and Medford, R.M. (2004). Superoxide, H₂O₂, and iron are required for TNF- α -induced MCP-1 gene expression in endothelial cells: role of Rac1 and NADPH oxidase. *Am. J. Physiol.-Heart Circ. Physiol.* 286, H1001–H1007.
- Chen, Y., Williams, V., Filippova, M., Filippov, V., and Duerksen-Hughes, P. (2014b). Viral Carcinogenesis: Factors Inducing DNA Damage and Virus Integration. *Cancers* 6, 2155–2186.
- Chen Wongworawat, Y., Filippova, M., Williams, V.M., Filippov, V., and Duerksen-Hughes, P.J. (2016). Chronic oxidative stress increases the integration frequency of foreign DNA and human papillomavirus 16 in human keratinocytes. *Am. J. Cancer Res.* 6, 764–780.
- Cheung, Y.-K., Cheng, S.C.-S., Sin, F.W.-Y., and Xie, Y. (2004). Plasmid encoding papillomavirus Type 16 (HPV16) DNA constructed with codon optimization improved the immunogenicity against HPV infection. *Vaccine* 23, 629–638.
- Chow, L.T., Roberts, J.M., Lewis, J.B., and Broker, T.R. (1977). A map of cytoplasmic RNA transcripts from lytic adenovirus type 2, determined by electron microscopy of RNA:DNA hybrids. *Cell* 11, 819–836.
- Chun, E., Lavoie, S., Michaud, M., Gallini, C.A., Kim, J., Soucy, G., Odze, R., Glickman, J.N., and Garrett, W.S. (2015). CCL2 Promotes Colorectal Carcinogenesis by Enhancing Polymorphonuclear Myeloid-Derived Suppressor Cell Population and Function. *Cell Rep.* 12, 244–257.
- Cid-Arregui, A., Juárez, V., and Hausen, H. zur (2003). A Synthetic E7 Gene of Human Papillomavirus Type 16 That Yields Enhanced Expression of the Protein in Mammalian Cells and Is Useful for DNA Immunization Studies. *J. Virol.* 77, 4928–4937.
- Clements, A., Johnston, K., Mazzei, J.M., Ricciardi, R.P., and Marmorstein, R. (2000). Oligomerization Properties of the Viral Oncoproteins Adenovirus E1A and Human Papillomavirus E7 and Their Complexes with the Retinoblastoma Protein. *Biochemistry* 39, 16033–16045.
- Conrady, M.C., Suarez, I., Gogl, G., Frecot, D.I., Bonhoure, A., Kostmann, C., Cousido-Siah, A., Mitschler, A., Lim, J., Masson, M., et al. (2020). Structure of High-Risk Papillomavirus 31 E6 Oncogenic Protein and Characterization of E6/E6AP/p53 Complex Formation. *J. Virol.* 95.
- Conway, M.J., Alam, S., Ryndock, E.J., Cruz, L., Christensen, N.D., Roden, R.B.S., and Meyers, C. (2009). Tissue-Spanning Redox Gradient-Dependent Assembly of Native Human Papillomavirus Type 16 Virions. *J. Virol.* 83, 10515–10526.
- Cornall, A.M., Roberts, J.M., Garland, S.M., Hillman, R.J., Grulich, A.E., and Tabrizi, S.N. (2013). Anal and perianal squamous carcinomas and high-grade intraepithelial lesions exclusively associated with “low-risk” HPV genotypes 6 and 11. *Int. J. Cancer* 133, 2253–2258.
- Cornelissen, M.T.E., Smits, H.L., Briët, M.A., van den Tweel, J.G., Struyk, A.P.H.B., van der Noordaa, J., and ter Schegget, J. (1990). Uniformity of the Splicing Pattern of the E6/E7

Transcripts in Human Papillomavirus Type 16-transformed Human Fibroblasts, Human Cervical Premalignant Lesions and Carcinomas. *J. Gen. Virol.* *71*, 1243–1246.

Cruz-Gregorio, A., Manzo-Merino, J., González-García, M.C., Pedraza-Chaverri, J., Medina-Campos, O.N., Valverde, M., Rojas, E., Rodríguez-Sastre, M.A., García-Cuellar, C.M., and Lizano, M. (2018a). Human Papillomavirus Types 16 and 18 Early-expressed Proteins Differentially Modulate the Cellular Redox State and DNA Damage. *Int. J. Biol. Sci.* *14*, 21–35.

Cruz-Gregorio, A., Manzo-Merino, J., and Lizano, M. (2018b). Cellular redox, cancer and human papillomavirus. *Virus Res.* *246*, 35–45.

Cruz-Gregorio, A., Aranda-Rivera, A.K., Aparicio-Trejo, O.E., Coronado-Martínez, I., Pedraza-Chaverri, J., and Lizano, M. (2019). E6 Oncoproteins from High-Risk Human Papillomavirus Induce Mitochondrial Metabolism in a Head and Neck Squamous Cell Carcinoma Model. *Biomolecules* *9*.

Davies, J.M.S., Cillard, J., Friguet, B., Cadenas, E., Cadet, J., Cayce, R., Fishmann, A., Liao, D., Bulteau, A.-L., Derbré, F., et al. (2017). The Oxygen Paradox, the French Paradox, and age-related diseases. *GeroScience* *39*, 499–550.

Davuluri, R.V., Suzuki, Y., Sugano, S., and Zhang, M.Q. (2000). CART Classification of Human 5' UTR Sequences. *Genome Res.* *10*, 1807–1816.

Day, P.M., Baker, C.C., Lowy, D.R., and Schiller, J.T. (2004). Establishment of papillomavirus infection is enhanced by promyelocytic leukemia protein (PML) expression. *Proc. Natl. Acad. Sci. U. S. A.* *101*, 14252–14257.

De la Cruz-Hernandez, E. (2005). Differential splicing of E6 within human papillomavirus type 18 variants and functional consequences. *J. Gen. Virol.* *86*, 2459–2468.

De la Rosa-Rios, M.A., Martínez-Salazar, M., Martínez-García, M., González-Bonilla, C., and Villegas-Sepúlveda, N. (2006). The intron 1 of HPV 16 has a suboptimal branch point at a guanosine. *Virus Res.* *118*, 46–54.

De Marco, F. (2013). Oxidative Stress and HPV Carcinogenesis. *Viruses* *5*, 708–731.

Desaint, S., Luriau, S., Aude, J.-C., Rousselet, G., and Toledano, M.B. (2004). Mammalian antioxidant defenses are not inducible by H₂O₂. *J. Biol. Chem.* *279*, 31157–31163.

DiGiuseppe, S., Bienkowska-Haba, M., Guion, L.G., and Sapp, M. (2017). Cruising the cellular highways: how human papillomavirus travels from the surface to the nucleus. *Virus Res.* *231*, 1–9.

Ding, B., Chi, S.G., Kim, S.H., Kang, S., Cho, J.H., Kim, D.S., and Cho, N.H. (2007). Role of p53 in antioxidant defense of HPV-positive cervical carcinoma cells following H₂O₂ exposure. *J. Cell Sci.* *120*, 2284–2294.

Doorbar, J. (2006). Molecular biology of human papillomavirus infection and cervical cancer. *Clin. Sci. Lond. Engl.* *199* *110*, 525–541.

- Doorbar, J., Quint, W., Banks, L., Bravo, I.G., Stoler, M., Broker, T.R., and Stanley, M.A. (2012). The biology and life-cycle of human papillomaviruses. *Vaccine 30 Suppl 5*, F55-70.
- Dosset, M., Castro, A., Carter, H., and Zanetti, M. (2020). Telomerase and CD4 T Cell Immunity in Cancer. *Cancers 12*, 1687.
- Duensing, S., and Münger, K. (2003). Human Papillomavirus Type 16 E7 Oncoprotein Can Induce Abnormal Centrosome Duplication through a Mechanism Independent of Inactivation of Retinoblastoma Protein Family Members. *J. Virol. 77*, 12331–12335.
- Duensing, A., Liu, Y., Tseng, M., Malumbres, M., Barbacid, M., and Duensing, S. (2006). Cyclin-dependent kinase 2 is dispensable for normal centrosome duplication but required for oncogene-induced centrosome overduplication. *Oncogene 25*, 2943–2949.
- Duensing, S., Duensing, A., Er, F., A, D., Pf, L., and K, M. (2001). Centrosome abnormalities and genomic instability by episomal expression of human papillomavirus type 16 in raft cultures of human keratinocytes. *J. Virol. 75*, 7712–7716.
- Duhoo, Y., Girault, V., Turchetto, J., Ramond, L., Durbesson, F., Fourquet, P., Nominé, Y., Cardoso, V., Sequeira, A.F., Brás, J.L.A., et al. (2019). High-Throughput Production of a New Library of Human Single and Tandem PDZ Domains Allows Quantitative PDZ-Peptide Interaction Screening Through High-Throughput Holdup Assay. *Methods Mol. Biol. Clifton NJ 2025*, 439–476.
- Dürst, M., Kleinheinz, A., Hotz, M., and Gissmann, L. (1985). The Physical State of Human Papillomavirus Type 16 DNA in Benign and Malignant Genital Tumours. *J. Gen. Virol. 66*, 1515–1522.
- Dürst, M., Croce, C.M., Gissmann, L., Schwarz, E., and Huebner, K. (1987). Papillomavirus sequences integrate near cellular oncogenes in some cervical carcinomas. *Proc. Natl. Acad. Sci. 84*, 1070–1074.
- Dyhr-Mikkelsen, H., and 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.
- Dyson, N. (1998). The regulation of E2F by pRB-family proteins. *Genes Dev. 12*, 2245–2262.
- Dyson, N., Howley, P.M., Munger, K., and Harlow, E. (1989). The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product. *Science 243*, 934–937.
- Egawa, N., and Doorbar, J. (2017). The low-risk papillomaviruses. *Virus Res. 231*, 119–127.
- Egawa, N., Egawa, K., Griffin, H., and Doorbar, J. (2015). Human Papillomaviruses; Epithelial Tropisms, and the Development of Neoplasia. *Viruses 7*, 3863–3890.
- Egawa, N., Wang, Q., Griffin, H.M., Murakami, I., Jackson, D., Mahmood, R., and Doorbar, J. (2017). HPV16 and 18 genome amplification show different E4-dependence, with 16E4 enhancing E1 nuclear accumulation and replicative efficiency via its cell cycle arrest and kinase activation functions. *PLoS Pathog. 13*, e1006282.

- Ehrig, F., Häfner, N., Driesch, C., Kraus Christiansen, I., Beer, K., Schmitz, M., Runnebaum, I.B., and Dürst, M. (2019). Differences in Stability of Viral and Viral-Cellular Fusion Transcripts in HPV-Induced Cervical Cancers. *Int. J. Mol. Sci.* 21.
- Eichten, A., Westfall, M., Pietenpol, J.A., and Münger, K. (2002). Stabilization and functional impairment of the tumor suppressor p53 by the human papillomavirus type 16 E7 oncoprotein. *Virology* 295, 74–85.
- Einstein, M.H., and Burk, R.D. (2001). Persistent human papillomavirus infection: Definitions and clinical implications. *Papillomavirus Rep.* 12, 119–123.
- Evans, W., Filippova, M., Filippov, V., Bashkirova, S., Zhang, G., Reeves, M.E., and Duerksen-Hughes, P. (2016). Overexpression of HPV16 E6* Alters β -Integrin and Mitochondrial Dysfunction Pathways in Cervical Cancer Cells. *Cancer Genomics - Proteomics* 13, 259–273.
- Fay, J., Kelehan, P., Lambkin, H., and Schwartz, S. (2009). Increased expression of cellular RNA-binding proteins in HPV-induced neoplasia and cervical cancer. *J. Med. Virol.* 81, 897–907.
- Filippova, M., Song, H., Connolly, J.L., Dermody, T.S., and Duerksen-Hughes, P.J. (2002). The Human Papillomavirus 16 E6 Protein Binds to Tumor Necrosis Factor (TNF) R1 and Protects Cells from TNF-induced Apoptosis. *J. Biol. Chem.* 277, 21730–21739.
- Filippova, M., Parkhurst, L., and Duerksen-Hughes, P.J. (2004). The Human Papillomavirus 16 E6 Protein Binds to Fas-associated Death Domain and Protects Cells from Fas-triggered Apoptosis. *J. Biol. Chem.* 279, 25729–25744.
- Filippova, M., Brown-Bryan, T.A., Casiano, C.A., and Duerksen-Hughes, P.J. (2005). The human papillomavirus 16 E6 protein can either protect or further sensitize cells to TNF: effect of dose. *Cell Death Differ.* 12, 1622–1635.
- Filippova, M., Johnson, M.M., Bautista, M., Filippov, V., Fodor, N., Tungteakkhun, S.S., Williams, K., and Duerksen-Hughes, P.J. (2007). The Large and Small Isoforms of Human Papillomavirus Type 16 E6 Bind to and Differentially Affect Procaspase 8 Stability and Activity. *J. Virol.* 81, 4116–4129.
- Filippova, M., Filippov, V.A., Kagoda, M., Garnett, T., Fodor, N., and Duerksen-Hughes, P.J. (2009). Complexes of Human Papillomavirus Type 16 E6 Proteins Form Pseudo-Death-Inducing Signaling Complex Structures during Tumor Necrosis Factor-Mediated Apoptosis. *J. Virol.* 83, 210–227.
- Filippova, M., Evans, W., Aragon, R., Filippov, V., Williams, V.M., Hong, L., Reeves, M.E., and Duerksen-Hughes, P. (2014a). The small splice variant of HPV16 E6, E6*, reduces tumor formation in cervical carcinoma xenografts. *Virology* 450–451, 153–164.
- Filippova, M., Filippov, V., Williams, V.M., Zhang, K., Kokoza, A., Bashkirova, S., and Duerksen-Hughes, P. (2014b). Cellular Levels of Oxidative Stress Affect the Response of Cervical Cancer Cells to Chemotherapeutic Agents (Hindawi).
- Fischer, M. (2017). Census and evaluation of p53 target genes. *Oncogene* 36, 3943–3956.

- Flores, E.R., and Lambert, P.F. (1997). Evidence for a switch in the mode of human papillomavirus type 16 DNA replication during the viral life cycle. *J. Virol.* *71*, 7167–7179.
- Funk, J.O., Waga, S., Harry, J.B., Espling, E., Stillman, B., and Galloway, D.A. (1997). Inhibition of CDK activity and PCNA-dependent DNA replication by p21 is blocked by interaction with the HPV-16 E7 oncoprotein. *Genes Dev.* *11*, 2090–2100.
- Gao, Y., Trivedi, S., Ferris, R.L., and Koide, K. (2014). Regulation of HPV16 E6 and MCL1 by SF3B1 inhibitor in head and neck cancer cells. *Sci. Rep.* *4*, 6098.
- García-Alai, M.M., Dantur, K.I., Smal, C., Pietrasanta, L., and de Prat-Gay, G. (2007). High-risk HPV E6 oncoproteins assemble into large oligomers that allow localization of endogenous species in prototypic HPV-transformed cell lines. *Biochemistry* *46*, 341–349.
- Geimanen, J., Isok-Paas, H., Pipitch, R., Salk, K., Laos, T., Orav, M., Reinson, T., Ustav, M., Ustav, M., and Ustav, E. (2011). Development of a Cellular Assay System To Study the Genome Replication of High- and Low-Risk Mucosal and Cutaneous Human Papillomaviruses. *J. Virol.* *85*, 3315–3329.
- Genther, S.M., Sterling, S., Duensing, S., Münger, K., Sattler, C., and Lambert, P.F. (2003). Quantitative Role of the Human Papillomavirus Type 16 E5 Gene during the Productive Stage of the Viral Life Cycle. *J. Virol.* *77*, 2832–2842.
- Ghittoni, R., Accardi, R., Hasan, U., Gheit, T., Sylla, B., and Tommasino, M. (2010). The biological properties of E6 and E7 oncoproteins from human papillomaviruses. *Virus Genes* *40*, 1–13.
- Gillespie, K.A., Mehta, K.P., Laimins, L.A., and Moody, C.A. (2012). Human Papillomaviruses Recruit Cellular DNA Repair and Homologous Recombination Factors to Viral Replication Centers. *J. Virol.* *86*, 9520–9526.
- Giovane, C., Travé, G., Briones, A., Lutz, Y., Wasylyk, B., and Weiss, E. (1999). Targetting of the N-terminal domain of the human papillomavirus type 16 E6 oncoprotein with monomeric scFvs blocks the E6-mediated degradation of cellular p53. *J. Mol. Recognit.* *12*, 141–152.
- Gissmann, L., Boshart, M., Dürst, M., Ikenberg, H., Wagner, D., and zur Hausen, H. (1984). Presence of human papillomavirus in genital tumors. *J. Invest. Dermatol.* *83*, 26s–28s.
- Glaunsinger, B.A., Lee, S.S., Thomas, M., Banks, L., and Javier, R. (2000). Interactions of the PDZ-protein MAGI-1 with adenovirus E4-ORF1 and high-risk papillomavirus E6 oncoproteins. *Oncogene* *19*, 5270–5280.
- Graham, S.V. (2017). The human papillomavirus replication cycle, and its links to cancer progression: a comprehensive review. *Clin. Sci. Lond. Engl.* *131*, 2201–2221.
- Graham, S.V., and Faizo, A.A.A. (2017). Control of human papillomavirus gene expression by alternative splicing. *Virus Res.* *231*, 83–95.
- Gritsenko, A.A., Weingarten-Gabbay, S., Elias-Kirma, S., Nir, R., de Ridder, D., and Segal, E. (2017). Sequence features of viral and human Internal Ribosome Entry Sites predictive of their activity. *PLoS Comput. Biol.* *13*.

- Grm, H.S., and Banks, L. (2004). Degradation of hDIg and MAGIs by human papillomavirus E6 is E6-AP-independent. *J. Gen. Virol.* 85, 2815–2819.
- Grm, H.S., Massimi, P., Gammoh, N., and Banks, L. (2005). Crosstalk between the human papillomavirus E2 transcriptional activator and the E6 oncoprotein. *Oncogene* 24, 5149–5164.
- Groves, I.J., and Coleman, N. (2015). Pathogenesis of human papillomavirus-associated mucosal disease: Mucosal HPV pathogenesis. *J. Pathol.* 235, 527–538.
- Guccione, E., Massimi, P., Bernat, A., and Banks, L. (2002). Comparative analysis of the intracellular location of the high- and low-risk human papillomavirus oncoproteins. *Virology* 293, 20–25.
- Guccione, E., Pim, D., and Banks, L. (2004). HPV-18 E6*I modulates HPV-18 full-length E6 functions in a cell cycle dependent manner. *Int. J. Cancer* 110, 928–933.
- Häfner, N., Driesch, C., Gajda, M., Jansen, L., Kirchmayr, R., Runnebaum, I.B., and Dürst, M. (2008). Integration of the HPV16 genome does not invariably result in high levels of viral oncogene transcripts. *Oncogene* 27, 1610–1617.
- Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of Cancer: The Next Generation. *Cell* 144, 646–674.
- Hanson, G., and Collier, J. (2018). Codon optimality, bias and usage in translation and mRNA decay. *Nat. Rev. Mol. Cell Biol.* 19, 20–30.
- Harden, M., and Munger, K. (2017). Perturbation of DROSHA and DICER expression by Human Papillomavirus 16 Oncoproteins. *Virology* 507, 192–198.
- Harden, M.E., Prasad, N., Griffiths, A., and Munger, K. (2017). Modulation of microRNA-mRNA Target Pairs by Human Papillomavirus 16 Oncoproteins. *MBio* 8.
- Hasan, U.A., Bates, E., Takeshita, F., Biliato, A., Accardi, R., Bouvard, V., Mansour, M., Vincent, I., Gissmann, L., Iftner, T., et al. (2007). TLR9 expression and function is abolished by the cervical cancer-associated human papillomavirus type 16. *J. Immunol. Baltim. Md* 1950 178, 3186–3197.
- Hasan, U.A., Zannetti, C., Parroche, P., Goutagny, N., Malfroy, M., Roblot, G., Carreira, C., Hussain, I., Müller, M., Taylor-Papadimitriou, J., et al. (2013). The Human papillomavirus type 16 E7 oncoprotein induces a transcriptional repressor complex on the Toll-like receptor 9 promoter. *J. Exp. Med.* 210, 1369–1387.
- zur Hausen, H. (2000). Papillomaviruses causing cancer: evasion from host-cell control in early events in carcinogenesis. *J. Natl. Cancer Inst.* 92, 690–698.
- zur Hausen, H. (2009). The search for infectious causes of human cancers: Where and why. *Virology* 392, 1–10.
- Hawley-Nelson, P., Vousden, K. h., Hubbert, N. l., Lowy, D. r., and Schiller, J. t. (1989). HPV16 E6 and E7 proteins cooperate to immortalize human foreskin keratinocytes. *EMBO J.* 8, 3905–3910.

- Heer, A., Alonso, L.G., and de Prat-Gay, G. (2011). E6*, the 50 Amino Acid Product of the Most Abundant Spliced Transcript of the E6 Oncoprotein in High-Risk Human Papillomavirus, Is a Promiscuous Folder and Binder. *Biochemistry* 50, 1376–1383.
- Helt, A.-M., and Galloway, D.A. (2001). Destabilization of the Retinoblastoma Tumor Suppressor by Human Papillomavirus Type 16 E7 Is Not Sufficient To Overcome Cell Cycle Arrest in Human Keratinocytes. *J. Virol.* 75, 6737–6747.
- Herfs, M., Yamamoto, Y., Laury, A., Wang, X., Nucci, M.R., McLaughlin-Drubin, M.E., Münger, K., Feldman, S., McKeon, F.D., Xian, W., et al. (2012). A discrete population of squamocolumnar junction cells implicated in the pathogenesis of cervical cancer. *Proc. Natl. Acad. Sci.* 109, 10516–10521.
- Herfs, M., Vargas, S.O., Yamamoto, Y., Howitt, B.E., Nucci, M.R., Hornick, J.L., Mckeon, F.D., Xian, W., and Crum, C.P. (2013a). A novel blueprint for ‘top down’ differentiation defines the cervical squamocolumnar junction during development, reproductive life, and neoplasia. *J. Pathol.* 229, 460–468.
- Herfs, M., Herran, C.P., Howitt, B., Laury, A., Nucci, M.R., Feldman, S., Jimenez, C.A., McKeon, F.D., Xian, W., and Crum, C.P. (2013b). Cervical squamocolumnar junction-specific markers define distinct, clinically relevant subsets of low-grade squamous intraepithelial lesions. *Am. J. Surg. Pathol.* 37, 1311–1318.
- Hia, F., Yang, S.F., Shichino, Y., Yoshinaga, M., Murakawa, Y., Vandenberg, A., Fukao, A., Fujiwara, T., Landthaler, M., Natsume, T., et al. (2019). Codon bias confers stability to human mRNAs. *EMBO Rep.* 20, e48220.
- Hong, D., Liu, J., Hu, Y., Lu, X., Li, B., Li, Y., Hu, D., Lu, W., Xie, X., and Cheng, X. (2017). Viral E6 is overexpressed via high viral load in invasive cervical cancer with episomal HPV16. *BMC Cancer* 17, 136.
- Howie, H.L., Katzenellenbogen, R.A., and Galloway, D.A. (2009). Papillomavirus E6 proteins. *Virology* 384, 324–334.
- Huh, K., Zhou, X., Hayakawa, H., Cho, J.-Y., Libermann, T.A., Jin, J., Harper, J.W., and Munger, K. (2007). Human Papillomavirus Type 16 E7 Oncoprotein Associates with the Cullin 2 Ubiquitin Ligase Complex, Which Contributes to Degradation of the Retinoblastoma Tumor Suppressor. *J. Virol.* 81, 9737–9747.
- Huh, K.-W., DeMasi, J., Ogawa, H., Nakatani, Y., Howley, P.M., and Münger, K. (2005). Association of the human papillomavirus type 16 E7 oncoprotein with the 600-kDa retinoblastoma protein-associated factor, p600. *Proc. Natl. Acad. Sci.* 102, 11492–11497.
- Huibregtse, J.M., Scheffner, M., and Howley, P.M. (1991). A cellular protein mediates association of p53 with the E6 oncoprotein of human papillomavirus types 16 or 18. *EMBO J.* 10, 4129–4135.
- Hung, A.Y., and Sheng, M. (2002). PDZ Domains: Structural Modules for Protein Complex Assembly. *J. Biol. Chem.* 277, 5699–5702.

Hyland, P.L., McDade, S.S., McCloskey, R., Dickson, G.J., Arthur, K., McCance, D.J., and Patel, D. (2011). Evidence for alteration of EZH2, BMI1, and KDM6A and epigenetic reprogramming in human papillomavirus type 16 E6/E7-expressing keratinocytes. *J. Virol.* *85*, 10999–11006.

IARC Cervix Cancer Screening.

Iftner, T., Elbel, M., Schopp, B., Hiller, T., Loizou, J.I., Caldecott, K.W., and Stubenrauch, F. (2002). Interference of papillomavirus E6 protein with single-strand break repair by interaction with XRCC1. *EMBO J.* *21*, 4741–4748.

Iftner, T., Haedicke-Jarboui, J., Wu, S.-Y., and Chiang, C.-M. (2017). Involvement of Brd4 in different steps of the papillomavirus life cycle. *Virus Res.* *231*, 76–82.

James, C.D., and Roberts, S. (2016). Viral Interactions with PDZ Domain-Containing Proteins-An Oncogenic Trait? *Pathog. Basel Switz.* *5*.

James, M.A., Lee, J.H., and Klingelutz, A.J. (2006a). Human Papillomavirus Type 16 E6 Activates NF- κ B, Induces cIAP-2 Expression, and Protects against Apoptosis in a PDZ Binding Motif-Dependent Manner. *J. Virol.* *80*, 5301–5307.

James, M.A., Lee, J.H., and Klingelutz, A.J. (2006b). HPV16-E6 associated hTERT promoter acetylation is E6AP dependent, increased in later passage cells and enhanced by loss of p300. *Int. J. Cancer* *119*, 1878–1885.

Jeon, S., and Lambert, P.F. (1995). Integration of human papillomavirus type 16 DNA into the human genome leads to increased stability of E6 and E7 mRNAs: implications for cervical carcinogenesis. *Proc. Natl. Acad. Sci. U. S. A.* *92*, 1654–1658.

Jing, M., Bohl, J., Brimer, N., Kinter, M., and Vande Pol, S.B. (2007). Degradation of tyrosine phosphatase PTPN3 (PTPH1) by association with oncogenic human papillomavirus E6 proteins. *J. Virol.* *81*, 2231–2239.

Johansson, C., and Schwartz, S. (2013a). Regulation of human papillomavirus gene expression by splicing and polyadenylation. *Nat. Rev. Microbiol.* *11*, 239–251.

Johansson, C., and Schwartz, S. (2013b). Regulation of human papillomavirus gene expression by splicing and polyadenylation. *Nat. Rev. Microbiol.* *11*, 239–251.

Johansson, C., Somberg, M., Li, X., Backström Winquist, E., Fay, J., Ryan, F., Pim, D., Banks, L., and Schwartz, S. (2012). HPV-16 E2 contributes to induction of HPV-16 late gene expression by inhibiting early polyadenylation. *EMBO J.* *31*, 3212–3227.

Johnson, K.M., Kines, R.C., Roberts, J.N., Lowy, D.R., Schiller, J.T., and Day, P.M. (2009). Role of heparan sulfate in attachment to and infection of the murine female genital tract by human papillomavirus. *J. Virol.* *83*, 2067–2074.

Jones, D.L., Alani, R.M., and Münger, K. (1997). The human papillomavirus E7 oncoprotein can uncouple cellular differentiation and proliferation in human keratinocytes by abrogating p21Cip1-mediated inhibition of cdk2. *Genes Dev.* *11*, 2101–2111.

- Jonkman, J.E.N., Cathcart, J.A., Xu, F., Bartolini, M.E., Amon, J.E., Stevens, K.M., and Colarusso, P. (2014). An introduction to the wound healing assay using live-cell microscopy. *Cell Adhes. Migr.* 8, 440–451.
- Jung, A.C., Briolat, J., Millon, R., de Reyniès, A., Rickman, D., Thomas, E., Abecassis, J., Clavel, C., and Wasylyk, B. (2010). Biological and clinical relevance of transcriptionally active human papillomavirus (HPV) infection in oropharynx squamous cell carcinoma. *Int. J. Cancer NA-NA*.
- Jung, H.M., Phillips, B.L., and Chan, E.K. (2014). miR-375 activates p21 and suppresses telomerase activity by coordinately regulating HPV E6/E7, E6AP, CIP2A, and 14-3-3 ζ . *Mol. Cancer* 13, 80.
- Jung, Y.-S., Kato, I., and Kim, H.-R.C. (2013). A novel function of HPV16-E6/E7 in epithelial–mesenchymal transition. *Biochem. Biophys. Res. Commun.* 435, 339–344.
- Kao, W.H., Beaudenon, S.L., Talis, A.L., Huibregtse, J.M., and Howley, P.M. (2000). Human Papillomavirus Type 16 E6 Induces Self-Ubiquitination of the E6AP Ubiquitin-Protein Ligase. *J. Virol.* 74, 6408–6417.
- Karamysheva, Z.N., Tikhonova, E.B., Grozdanov, P.N., Huffman, J.C., Baca, K.R., Karamyshev, A., Denison, R.B., MacDonald, C.C., Zhang, K., and Karamyshev, A.L. (2018). Polysome Profiling in Leishmania, Human Cells and Mouse Testis. *JoVE J. Vis. Exp.* e57600.
- Katzenellenbogen, R.A., Egelkrout, E.M., Vliet-Gregg, P., Gewin, L.C., Gafken, P.R., and Galloway, D.A. (2007). NFX1-123 and Poly(A) Binding Proteins Synergistically Augment Activation of Telomerase in Human Papillomavirus Type 16 E6-Expressing Cells. *J. Virol.* 81, 3786–3796.
- Kgatle, M.M., Spearman, C.W., Kalla, A.A., and Hairwadzi, H.N. (2017). DNA Oncogenic Virus-Induced Oxidative Stress, Genomic Damage, and Aberrant Epigenetic Alterations. *Oxid. Med. Cell. Longev.* 2017.
- Kietzmann, T., and Görlach, A. (2005). Reactive oxygen species in the control of hypoxia-inducible factor-mediated gene expression. *Semin. Cell Dev. Biol.* 16, 474–486.
- Kines, R.C., Thompson, C.D., Lowy, D.R., Schiller, J.T., and Day, P.M. (2009). The initial steps leading to papillomavirus infection occur on the basement membrane prior to cell surface binding. *Proc. Natl. Acad. Sci.* 106, 20458–20463.
- Kiyono, T., Hiraiwa, A., Fujita, M., Hayashi, Y., Akiyama, T., and Ishibashi, M. (1997). Binding of high-risk human papillomavirus E6 oncoproteins to the human homologue of the *Drosophila* discs large tumor suppressor protein. *Proc. Natl. Acad. Sci.* 94, 11612–11616.
- Knapp, A.A., McManus, P.M., Bockstall, K., and Moroianu, J. (2009). Identification of the nuclear localization and export signals of high risk HPV16 E7 oncoprotein. *Virology* 383, 60–68.

Koh, W.-J., Abu-Rustum, N.R., Bean, S., Bradley, K., Campos, S.M., Cho, K.R., Chon, H.S., Chu, C., Clark, R., Cohn, D., et al. (2019). Cervical Cancer, Version 3.2019, NCCN Clinical Practice Guidelines in Oncology. *J. Natl. Compr. Canc. Netw.* 17, 64–84.

Kösel, S., Burggraf, S., Engelhardt, W., and Olgemöller, B. (2007). Increased levels of HPV16 E6*I transcripts in high-grade cervical cytology and histology (CIN II+) detected by rapid real-time RT-PCR amplification. *Cytopathology* 18, 290–299.

Kozak, M. (2002). Pushing the limits of the scanning mechanism for initiation of translation. *Gene* 299, 1–34.

Kühne, C., and Banks, L. (1998). E3-Ubiquitin Ligase/E6-AP Links Multicopy Maintenance Protein 7 to the Ubiquitination Pathway by a Novel Motif, the L2G Box. *J. Biol. Chem.* 273, 34302–34309.

Kumar, A. (2007). The silent defense: Micro-RNA directed defense against HIV-1 replication. *Retrovirology* 4, 26.

Kumar, A., Zhao, Y., Meng, G., Zeng, M., Srinivasan, S., Delmolino, L.M., Gao, Q., Dimri, G., Weber, G.F., Wazer, D.E., et al. (2002). Human Papillomavirus Oncoprotein E6 Inactivates the Transcriptional Coactivator Human ADA3. *Mol. Cell. Biol.* 22, 5801–5812.

Lace, M.J., Anson, J.R., Thomas, G.S., Turek, L.P., and Haugen, T.H. (2008). The E8AE2 Gene Product of Human Papillomavirus Type 16 Represses Early Transcription and Replication but Is Dispensable for Viral Plasmid Persistence in Keratinocytes. *J. Virol.* 82, 10841–10853.

Lace, M.J., Isacson, C., Anson, J.R., Lörincz, A.T., Wilczynski, S.P., Haugen, T.H., and Turek, L.P. (2009). Upstream regulatory region alterations found in human papillomavirus type 16 (HPV-16) isolates from cervical carcinomas increase transcription, ori function, and HPV immortalization capacity in culture. *J. Virol.* 83, 7457–7466.

Lechner, M.S., and Laimins, L.A. (1994). Inhibition of p53 DNA binding by human papillomavirus E6 proteins. *J. Virol.* 68, 4262–4273.

Lee, C., and Laimins, L.A. (2004). Role of the PDZ Domain-Binding Motif of the Oncoprotein E6 in the Pathogenesis of Human Papillomavirus Type 31. *J. Virol.* 78, 12366–12377.

Lee, Y., and Rio, D.C. (2015). Mechanisms and Regulation of Alternative Pre-mRNA Splicing. *Annu. Rev. Biochem.* 84, 291–323.

Lee, S.S., Glaunsinger, B., Mantovani, F., Banks, L., and Javier, R.T. (2000). Multi-PDZ Domain Protein MUPP1 Is a Cellular Target for both Adenovirus E4-ORF1 and High-Risk Papillomavirus Type 18 E6 Oncoproteins. *J. Virol.* 74, 9680–9693.

Leon, K., and Ott, M. (2020). An ‘Arms Race’ between the Nonsense-mediated mRNA Decay Pathway and Viral Infections. *Semin. Cell Dev. Biol.*

Leverrier, S., Bergamaschi, D., Ghali, L., Ola, A., Warnes, G., Akgül, B., Blight, K., García-Escudero, R., Penna, A., Eddaoudi, A., et al. (2007). Role of HPV E6 proteins in preventing UVB-induced release of pro-apoptotic factors from the mitochondria. *Apoptosis* 12, 549–560.

- Li, Y., Cai, Q., Lin, L., and Xu, C. (2018). MiR-875 and miR-3144 switch the human papillomavirus 16 E6/E6* mRNA ratio through the EGFR pathway and a direct targeting effect. *Gene* 679, 389–397.
- Lin, S.-Y., and Elledge, S.J. (2003). Multiple Tumor Suppressor Pathways Negatively Regulate Telomerase. *Cell* 113, 881–889.
- Lin, L., Cai, Q., Zhang, X., Zhang, H., Zhong, Y., Xu, C., and Li, Y. (2015). Two less common human microRNAs miR-875 and miR-3144 target a conserved site of E6 oncogene in most high-risk human papillomavirus subtypes. *Protein Cell* 6, 575–588.
- Litwin, T.R., Clarke, M.A., Dean, M., and Wentzensen, N. (2017). Somatic Host Cell Alterations in HPV Carcinogenesis. *Viruses* 9, 206.
- Liu, B., Chen, Y., and St. Clair, D.K. (2008). ROS and p53: versatile partnership. *Free Radic. Biol. Med.* 44, 1529–1535.
- Liu, X., Clements, A., Zhao, K., and Marmorstein, R. (2006). Structure of the Human Papillomavirus E7 Oncoprotein and Its Mechanism for Inactivation of the Retinoblastoma Tumor Suppressor. *J. Biol. Chem.* 281, 578–586.
- Liu, X., Dakic, A., Zhang, Y., Dai, Y., Chen, R., and Schlegel, R. (2009). HPV E6 protein interacts physically and functionally with the cellular telomerase complex. *Proc. Natl. Acad. Sci.* 106, 18780–18785.
- López-Bueno, A., Mavian, C., Labella, A.M., Castro, D., Borrego, J.J., Alcamí, A., and Alejo, A. (2016). Concurrence of Iridovirus, Polyomavirus, and a Unique Member of a New Group of Fish Papillomaviruses in Lymphocystis Disease-Affected Gilthead Sea Bream. *J. Virol.* 90, 8768–8779.
- López-Ocejo, O., Vilorio-Petit, A., Bequet-Romero, M., Mukhopadhyay, D., Rak, J., and Kerbel, R.S. (2000). Oncogenes and tumor angiogenesis: the HPV-16 E6 oncoprotein activates the vascular endothelial growth factor (VEGF) gene promoter in a p53 independent manner. *Oncogene* 19, 4611–4620.
- López-Urrutia, E., Valdés, J., Bonilla-Moreno, R., Martínez-Salazar, M., Martínez-García, M., Berumen, J., and Villegas-Sepúlveda, N. (2012). A few nucleotide polymorphisms are sufficient to recruit nuclear factors differentially to the intron 1 of HPV-16 intratypic variants. *Virus Res.* 166, 43–53.
- Maglennon, G.A., McIntosh, P., and Doorbar, J. (2011). Persistence of viral DNA in the epithelial basal layer suggests a model for papillomavirus latency following immune regression. *Virology* 414, 153–163.
- Malanchi, I., Accardi, R., Diehl, F., Smet, A., Androphy, E., Hoheisel, J., and Tommasino, M. (2004). Human Papillomavirus Type 16 E6 Promotes Retinoblastoma Protein Phosphorylation and Cell Cycle Progression. *J. Virol.* 78, 13769–13778.
- Mantovani, F., and Banks, L. (1999). Inhibition of E6 induced degradation of p53 is not sufficient for stabilization of p53 protein in cervical tumour derived cell lines. *Oncogene* 18, 3309–3315.

- Manzo-Merino, J., Massimi, P., Lizano, M., and Banks, L. (2014). The human papillomavirus (HPV) E6 oncoproteins promotes nuclear localization of active caspase 8. *Virology* 450–451, 146–152.
- de Martel, C., Plummer, M., Vignat, J., and Franceschi, S. (2018). Worldwide burden of cancer attributable to HPV by site, country and HPV type. *Int. J. Cancer* 664–670.
- Martínez-Salazar, M., López-Urrutia, E., Arechaga-Ocampo, E., Bonilla-Moreno, R., Martínez-Castillo, M., Díaz-Hernández, J., Del Moral-Hernández, O., Cedillo-Barrón, L., Martínez-Juarez, V., De Nova-Ocampo, M., et al. (2014). Biochemical and proteomic analysis of spliceosome factors interacting with intron-1 of human papillomavirus type-16. *J. Proteomics* 111, 184–197.
- Martinon, F. (2010). Signaling by ROS drives inflammasome activation. *Eur. J. Immunol.* 40, 616–619.
- Marullo, R., Werner, E., Zhang, H., Chen, G.Z., Shin, D.M., and Doetsch, P.W. (2015). HPV16 E6 and E7 proteins induce a chronic oxidative stress response via NOX2 that causes genomic instability and increased susceptibility to DNA damage in head and neck cancer cells. *Carcinogenesis* 36, 1397–1406.
- Massimi, P., and Banks, L. (1997). Repression of p53 transcriptional activity by the HPV E7 proteins. *Virology* 227, 255–259.
- Massimi, P., and Banks, L. (2000). Differential phosphorylation of the HPV-16 E7 oncoprotein during the cell cycle. *Virology* 276, 388–394.
- Masson, M., Hindelang, C., Sibling, A.-P., Schwalbach, G., Travé, G., and Weiss, E. (2003). Preferential nuclear localization of the human papillomavirus type 16 E6 oncoprotein in cervical carcinoma cells. *J. Gen. Virol.* 84, 2099–2104.
- Masterson, L., Sorgeloos, F., Winder, D., Lechner, M., Marker, A., Malhotra, S., Sudhoff, H., Jani, P., Goon, P., and Sterling, J. (2015). Deregulation of SYCP2 predicts early stage human papillomavirus-positive oropharyngeal carcinoma: A prospective whole transcriptome analysis. *Cancer Sci.* 106, 1568–1575.
- Matlashewski, G., Schneider, J., Banks, L., Jones, N., Murray, A., and Crawford, L. (1987). Human papillomavirus type 16 DNA cooperates with activated ras in transforming primary cells. *EMBO J.* 6, 1741–1746.
- McBride, A.A. (2017). Oncogenic human papillomaviruses. *Philos. Trans. R. Soc. B Biol. Sci.* 372, 20160273.
- McBride, A.A., and Warburton, A. (2017). The role of integration in oncogenic progression of HPV-associated cancers. *PLoS Pathog.* 13.
- McFarlane, M., MacDonald, A.I., Stevenson, A., and Graham, S.V. (2015). Human Papillomavirus 16 Oncoprotein Expression Is Controlled by the Cellular Splicing Factor SRSF2 (SC35). *J. Virol.* 89, 5276–5287.

- McIntyre, M.C., Frattini, M.G., Grossman, S.R., and Laimins, L.A. (1993). Human papillomavirus type 18 E7 protein requires intact Cys-X-X-Cys motifs for zinc binding, dimerization, and transformation but not for Rb binding. *J. Virol.* *67*, 3142–3150.
- McLaughlin-Drubin, M.E., and Münger, K. (2009). The human papillomavirus E7 oncoprotein. *Virology* *384*, 335–344.
- McLaughlin-Drubin, M.E., Crum, C.P., and Münger, K. (2011). Human papillomavirus E7 oncoprotein induces KDM6A and KDM6B histone demethylase expression and causes epigenetic reprogramming. *Proc. Natl. Acad. Sci. U. S. A.* *108*, 2130–2135.
- McMurray, H.R., and McCance, D.J. (2003). Human Papillomavirus Type 16 E6 Activates TERT Gene Transcription through Induction of c-Myc and Release of USF-Mediated Repression. *J. Virol.* *77*, 9852–9861.
- McMurray, H.R., and McCance, D.J. (2004). Degradation of p53, Not Telomerase Activation, by E6 Is Required for Bypass of Crisis and Immortalization by Human Papillomavirus Type 16 E6/E7. *J. Virol.* *78*, 5698–5706.
- Mesplède, T., Gagnon, D., Bergeron-Labrecque, F., Azar, I., Sénéchal, H., Coutlée, F., and Archambault, J. (2012). p53 Degradation Activity, Expression, and Subcellular Localization of E6 Proteins from 29 Human Papillomavirus Genotypes. *J. Virol.* *86*, 94–107.
- Mileo, A.M., Abbruzzese, C., Mattarocci, S., Bellacchio, E., Pisano, P., Federico, A., Maresca, V., Picardo, M., Giorgi, A., Maras, B., et al. (2009). Human Papillomavirus-16 E7 Interacts with Glutathione S-Transferase P1 and Enhances Its Role in Cell Survival. *PLOS ONE* *4*, e7254.
- Milligan, S.G., Veerapraditsin, T., Ahamet, B., Mole, S., and Graham, S.V. (2007). Analysis of novel human papillomavirus type 16 late mRNAs in differentiated W12 cervical epithelial cells. *Virology* *360*, 172–181.
- Mirabello, L., Yeager, M., Yu, K., Clifford, G.M., Xiao, Y., Zhu, B., Cullen, M., Boland, J.F., Wentzensen, N., Nelson, C.W., et al. (2017). HPV16 E7 Genetic Conservation Is Critical to Carcinogenesis. *Cell* *170*, 1164-1174.e6.
- Miura, S., Kawana, K., Schust, D.J., Fujii, T., Yokoyama, T., Iwasawa, Y., Nagamatsu, T., Adachi, K., Tomio, A., Tomio, K., et al. (2010). CD1d, a sentinel molecule bridging innate and adaptive immunity, is downregulated by the human papillomavirus (HPV) E5 protein: a possible mechanism for immune evasion by HPV. *J. Virol.* *84*, 11614–11623.
- Moens, U., Calvignac-Spencer, S., Lauber, C., Ramqvist, T., Feltkamp, M.C.W., Daugherty, M.D., Verschoor, E.J., and Ehlers, B. (2017). ICTV Virus Taxonomy Profile: Polyomaviridae. *J. Gen. Virol.* *98*, 1159–1160.
- Mole, S., McFarlane, M., Chuen-Im, T., Milligan, S.G., Millan, D., and Graham, S.V. (2009). RNA splicing factors regulated by HPV16 during cervical tumour progression. *J. Pathol.* *219*, 383–391.

Moody, C.A., and Laimins, L.A. (2009). Human Papillomaviruses Activate the ATM DNA Damage Pathway for Viral Genome Amplification upon Differentiation. *PLOS Pathog.* 5, e1000605.

del Moral-Hernández, O., López-Urrutia, E., Bonilla-Moreno, R., Martínez-Salazar, M., Arechaga-Ocampo, E., Berumen, J., and 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.

Morel, A., Baguet, A., Perrard, J., Demeret, C., Jacquin, E., Guenat, D., Mougin, C., and Prétet, J.-L. (2017). 5azadC treatment upregulates miR-375 level and represses HPV16 E6 expression. *Oncotarget* 8, 46163–46176.

Mouret, S., Sauvaigo, S., Peinnequin, A., Favier, A., Beani, J.-C., and Leccia, M.-T. (2005). E6* oncoprotein expression of human papillomavirus type-16 determines different ultraviolet sensitivity related to glutathione and glutathione peroxidase antioxidant defence. *Exp. Dermatol.* 14, 401–410.

Münger, K., Werness, B.A., Dyson, N., Phelps, W.C., Harlow, E., and Howley, P.M. (1989). Complex formation of human papillomavirus E7 proteins with the retinoblastoma tumor suppressor gene product. *EMBO J.* 8, 4099–4105.

Münger, K., Basile, J.R., Duensing, S., Eichten, A., Gonzalez, S.L., Grace, M., and Zacny, V.L. (2001). Biological activities and molecular targets of the human papillomavirus E7 oncoprotein. *Oncogene* 20, 7888–7898.

Nahand, J.S., Moghoofei, M., Salmaninejad, A., Bahmanpour, Z., Karimzadeh, M., Nasiri, M., Mirzaei, H.R., Pourhanifeh, M.H., Bokharai-Salim, F., Mirzaei, H., et al. (2020). Pathogenic role of exosomes and microRNAs in HPV-mediated inflammation and cervical cancer: A review. *Int. J. Cancer* 146, 305–320.

Nakagawa, S., and Huibregtse, J.M. (2000). Human Scribble (Vartul) Is Targeted for Ubiquitin-Mediated Degradation by the High-Risk Papillomavirus E6 Proteins and the E6AP Ubiquitin-Protein Ligase. *Mol. Cell. Biol.* 20, 8244–8253.

Nicolaides, L., Davy, C., Raj, K., Kranjec, C., Banks, L., and Doorbar, J. (2011). Stabilization of HPV16 E6 protein by PDZ proteins, and potential implications for genome maintenance. *Virology* 414, 137–145.

Nishimura, A., Ono, T., Ishimoto, A., Dowhanick, J.J., Frizzell, M.A., Howley, P.M., and Sakai, H. (2000). Mechanisms of Human Papillomavirus E2-Mediated Repression of Viral Oncogene Expression and Cervical Cancer Cell Growth Inhibition. *J. Virol.* 74, 3752–3760.

Nominé, Y., Charbonnier, S., Ristriani, T., Stier, G., Masson, M., Cavusoglu, N., Van Dorsselaer, A., Weiss, É., Kieffer, B., and Travé, G. (2003). Domain Substructure of HPV E6 Oncoprotein: Biophysical Characterization of the E6 C-Terminal DNA-Binding Domain. *Biochemistry* 42, 4909–4917.

Nominé, Y., Masson, M., Charbonnier, S., Zanier, K., Ristriani, T., Deryckère, F., Sibler, A.-P., Desplancq, D., Atkinson, R.A., Weiss, E., et al. (2006). Structural and functional analysis

of E6 oncoprotein: insights in the molecular pathways of human papillomavirus-mediated pathogenesis. *Mol. Cell* 21, 665–678.

Oberg, D., Collier, B., Zhao, X., and Schwartz, S. (2003). Mutational inactivation of two distinct negative RNA elements in the human papillomavirus type 16 L2 coding region induces production of high levels of L2 in human cells. *J. Virol.* 77, 11674–11684.

Oberg, D., Fay, J., Lambkin, H., and Schwartz, S. (2005). A downstream polyadenylation element in human papillomavirus type 16 L2 encodes multiple GGG motifs and interacts with hnRNP H. *J. Virol.* 79, 9254–9269.

Oh, K.-J., Kalinina, A., Park, N.-H., and Bagchi, S. (2006). Deregulation of eIF4E: 4E-BP1 in Differentiated Human Papillomavirus-Containing Cells Leads to High Levels of Expression of the E7 Oncoprotein. *J. Virol.* 80, 7079–7088.

Oh, S.T., Kyo, S., and Laimins, L.A. (2001). 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.* 75, 5559–5566.

Olmedo-Nieva, L., Muñoz-Bello, J.O., Contreras-Paredes, A., and Lizano, M. (2018). The Role of E6 Spliced Isoforms (E6*) in Human Papillomavirus-Induced Carcinogenesis. *Viruses* 10.

Paaso, A., Jaakola, A., Syrjänen, S., and Louvanto, K. (2019). From HPV Infection to Lesion Progression: The Role of HLA Alleles and Host Immunity. *Acta Cytol.* 63, 148–158.

Paget-Bailly, P., Meznad, K., Bruyère, D., Perrard, J., Herfs, M., Jung, A.C., Mougin, C., Prétet, J.-L., and Baguet, A. (2019). Comparative RNA sequencing reveals that HPV16 E6 abrogates the effect of E6*I on ROS metabolism. *Sci. Rep.* 9, 5938.

Pang, E., Delic, N.C., Hong, A., Zhang, M., Rose, B.R., and Lyons, J.G. (2011). Radiosensitization of oropharyngeal squamous cell carcinoma cells by human papillomavirus 16 oncoprotein E6*I. *Int. J. Radiat. Oncol. Biol. Phys.* 79, 860–865.

Pappa, K.I., Kontostathi, G., Lygirou, V., Zoidakis, J., and Anagnostou, N.P. (2018). Novel structural approaches concerning HPV proteins: Insight into targeted therapies for cervical cancer (Review). *Oncol. Rep.* 39, 1547–1554.

Park, R.B., and Androphy, E.J. (2002). Genetic Analysis of High-Risk E6 in Episomal Maintenance of Human Papillomavirus Genomes in Primary Human Keratinocytes. *J. Virol.* 76, 11359–11364.

Parroche, P., Touka, M., Mansour, M., Bouvard, V., Thépot, A., Accardi, R., Carreira, C., Roblot, G.G., Sylla, B.S., Hasan, U., et al. (2011). Human papillomavirus type 16 E6 inhibits p21WAF1 transcription independently of p53 by inactivating p150Sal2. *Virology* 417, 443–448.

Peh, W.L., Middleton, K., Christensen, N., Nicholls, P., Egawa, K., Sotlar, K., Brandsma, J., Percival, A., Lewis, J., Liu, W.J., et al. (2002). Life cycle heterogeneity in animal models of human papillomavirus-associated disease. *J. Virol.* 76, 10401–10416.

- Peña, L. del M., and Laimins, L.A. (2001). Differentiation-Dependent Chromatin Rearrangement Coincides with Activation of Human Papillomavirus Type 31 Late Gene Expression. *J. Virol.* *75*, 10005–10013.
- Pentland, I., Campos-León, K., Cotic, M., Davies, K.-J., Wood, C.D., Groves, I.J., Burley, M., Coleman, N., Stockton, J.D., Noyvert, B., et al. (2018). Disruption of CTCF-YY1-dependent looping of the human papillomavirus genome activates differentiation-induced viral oncogene transcription. *PLoS Biol.* *16*.
- Pett, M.R., Alazawi, W.O.F., Roberts, I., Downen, S., Smith, D.I., Stanley, M.A., and Coleman, N. (2004). Acquisition of high-level chromosomal instability is associated with integration of human papillomavirus type 16 in cervical keratinocytes. *Cancer Res.* *64*, 1359–1368.
- Phelps, W.C., Yee, C.L., Münger, K., and Howley, P.M. (1988). The human papillomavirus type 16 E7 gene encodes transactivation and transformation functions similar to those of adenovirus E1A. *Cell* *53*, 539–547.
- Phillips, A.C., and Vousden, K.H. (1997). Analysis of the interaction between human papillomavirus type 16 E7 and the TATA-binding protein, TBP. *J. Gen. Virol.* *78* (Pt 4), 905–909.
- Pim, D., and Banks, L. (1999). HPV-18 E6*I protein modulates the E6-directed degradation of p53 by binding to full-length HPV-18 E6. *Oncogene* *18*, 7403–7408.
- Pim, D., Massimi, P., and Banks, L. (1997a). Alternatively spliced HPV-18 E6* protein inhibits E6 mediated degradation of p53 and suppresses transformed cell growth. *Oncogene* *15*, 257–264.
- Pim, D., Massimi, P., and Banks, L. (1997b). Alternatively spliced HPV-18 E6* protein inhibits E6 mediated degradation of p53 and suppresses transformed cell growth. *Oncogene* *15*, 257–264.
- Pim, D., Tomaic, V., and Banks, L. (2009). The Human Papillomavirus (HPV) E6* Proteins from High-Risk, Mucosal HPVs Can Direct Degradation of Cellular Proteins in the Absence of Full-Length E6 Protein. *J. Virol.* *83*, 9863–9874.
- Porta, C., Riboldi, E., and Sica, A. (2011). Mechanisms linking pathogens-associated inflammation and cancer. *Cancer Lett.* *305*, 250–262.
- Pringle, E.S., McCormick, C., and Cheng, Z. (2019). Polysome Profiling Analysis of mRNA and Associated Proteins Engaged in Translation. *Curr. Protoc. Mol. Biol.* *125*, e79.
- Pyeon, D., Newton, M.A., Lambert, P.F., den Boon, J.A., Sengupta, S., Marsit, C.J., Woodworth, C.D., Connor, J.P., Haugen, T.H., Smith, E.M., et al. (2007). Fundamental Differences in Cell Cycle Deregulation in Human Papillomavirus-Positive and Human Papillomavirus-Negative Head/Neck and Cervical Cancers. *Cancer Res.* *67*, 4605–4619.
- Pyeon, D., Pearce, S.M., Lank, S.M., Ahlquist, P., and Lambert, P.F. (2009). Establishment of human papillomavirus infection requires cell cycle progression. *PLoS Pathog.* *5*, e1000318.

- Qin, T., Koneva, L.A., Liu, Y., Zhang, Y., Arthur, A.E., Zarins, K.R., Carey, T.E., Chepeha, D., Wolf, G.T., Rozek, L.S., et al. (2020). Significant association between host transcriptome-derived HPV oncogene E6* influence score and carcinogenic pathways, tumor size, and survival in head and neck cancer. *Head Neck* *n/a*.
- Raad, H., Paciet, M.-H., Boussetta, T., Kroviarski, Y., Morel, F., Quinn, M.T., Gougerot-Pocidalò, M.-A., Dang, P.M.-C., and El-Benna, J. (2009). Regulation of the phagocyte NADPH oxidase activity: phosphorylation of gp91phox/NOX2 by protein kinase C enhances its diaphorase activity and binding to Rac2, p67phox, and p47phox. *FASEB J.* *23*, 1011–1022.
- Rak, R., Dahan, O., and Pilpel, Y. (2018). Repertoires of tRNAs: The Couplers of Genomics and Proteomics. *Annu. Rev. Cell Dev. Biol.* *34*, 239–264.
- Reinstein, E., Scheffner, M., Oren, M., Ciechanover, A., and Schwartz, A. (2000). Degradation of the E7 human papillomavirus oncoprotein by the ubiquitin-proteasome system: targeting via ubiquitination of the N-terminal residue. *Oncogene* *19*, 5944–5950.
- Reis, J., Massari, M., Marchese, S., Ceccon, M., Aalbers, F.S., Corana, F., Valente, S., Mai, A., Magnani, F., and Mattevi, A. (2020). A closer look into NADPH oxidase inhibitors: Validation and insight into their mechanism of action. *Redox Biol.* *32*, 101466.
- Reshkin, S.J., Bellizzi, A., Caldeira, S., Albarani, V., Malanchi, I., Poignee, M., Alunni-Fabbroni, M., Casavola, V., and Tommasino, M. (2000). Na⁺/H⁺ exchanger-dependent intracellular alkalization is an early event in malignant transformation and plays an essential role in the development of subsequent transformation-associated phenotypes. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* *14*, 2185–2197.
- Ribeiro, A.L., Caodaglio, A.S., and Sichero, L. (2018). Regulation of HPV transcription. *Clinics* *73*.
- Riethmuller, D., Jacquard, A.-C., Lacau St Guily, J., Aubin, F., Carcopino, X., Pradat, P., Dahlab, A., and Prétet, J.-L. (2015). Potential impact of a nonavalent HPV vaccine on the occurrence of HPV-related diseases in France. *BMC Public Health* *15*, 453.
- Rigalli, J.P., Reichel, M., Tocchetti, G.N., Reuter, T., Dyckhoff, G., Herold-Mende, C., and Weiss, J. (2018). Human papilloma virus (HPV) 18 proteins E6 and E7 up-regulate ABC transporters in oropharyngeal carcinoma. Involvement of the nonsense-mediated decay (NMD) pathway. *Cancer Lett.* *428*, 69–76.
- Ristriani, T., Fournane, S., Orfanoudakis, G., Travé, G., and Masson, M. (2009). A single-codon mutation converts HPV16 E6 oncoprotein into a potential tumor suppressor, which induces p53-dependent senescence of HPV-positive HeLa cervical cancer cells. *Oncogene* *28*, 762–772.
- Roberts, J.N., Buck, C.B., Thompson, C.D., Kines, R., Bernardo, M., Choyke, P.L., Lowy, D.R., and Schiller, J.T. (2007). Genital transmission of HPV in a mouse model is potentiated by nonoxynol-9 and inhibited by carrageenan. *Nat. Med.* *13*, 857–861.

- Rohlf, M., Winkenbach, S., Meyer, S., Rupp, T., and Dürst, M. (1991). Viral transcription in human keratinocyte cell lines immortalized by human papillomavirus type-16. *Virology* *183*, 331–342.
- Roman, A., and Munger, K. (2013). The papillomavirus E7 proteins. *Virology* *445*, 138–168.
- Rosenberger, S., Arce, J.D.-C., Langbein, L., Steenbergen, R.D.M., and Rösl, F. (2010). Alternative splicing of human papillomavirus type-16 E6/E6* early mRNA is coupled to EGF signaling via Erk1/2 activation. *Proc. Natl. Acad. Sci. U. S. A.* *107*, 7006–7011.
- Roy, R., Durie, D., Li, H., Liu, B.-Q., Skehel, J.M., Mauri, F., Cuorvo, L.V., Barbareschi, M., Guo, L., Holcik, M., et al. (2014). hnRNPA1 couples nuclear export and translation of specific mRNAs downstream of FGF-2/S6K2 signalling. *Nucleic Acids Res.* *42*, 12483–12497.
- Samorski, R., Gissmann, L., and Osen, W. (2006). Codon optimized expression of HPV 16 E6 renders target cells susceptible to E6-specific CTL recognition. *Immunol. Lett.* *107*, 41–49.
- Sanders, C.M., and Stenlund, A. (1998). Recruitment and loading of the E1 initiator protein: an ATP-dependent process catalysed by a transcription factor. *EMBO J.* *17*, 7044–7055.
- de Sanjosé, S., Diaz, M., Castellsagué, X., Clifford, G., Bruni, L., Muñoz, N., and Bosch, F.X. (2007). Worldwide prevalence and genotype distribution of cervical human papillomavirus DNA in women with normal cytology: a meta-analysis. *Lancet Infect. Dis.* *7*, 453–459.
- Sasagawa, T., Takagi, H., and Makinoda, S. (2012). Immune responses against human papillomavirus (HPV) infection and evasion of host defense in cervical cancer. *J. Infect. Chemother.* *18*, 807–815.
- Scheffner, M., Werness, B.A., Huibregtse, J.M., Levine, A.J., and Howley, P.M. (1990). The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell* *63*, 1129–1136.
- Scheffner, M., Huibregtse, J.M., Vierstra, R.D., and Howley, P.M. (1993). The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. *Cell* *75*, 495–505.
- Schiffman, M., Doorbar, J., Wentzensen, N., de Sanjosé, S., Fakhry, C., Monk, B.J., Stanley, M.A., and Franceschi, S. (2016). Carcinogenic human papillomavirus infection. *Nat. Rev. Dis. Primer* *2*, 1–20.
- Schmitt, J., Schlehofer, J.R., Mergener, K., Gissmann, L., and Hausen, H.Z. (1989). Amplification of bovine papillomavirus DNA by N-Methyl-N'-Nitro-N-nitrosoguanidine, ultraviolet irradiation, or infection with herpes simplex virus. *Virology* *172*, 73–81.
- Schmitz, M., Driesch, C., Jansen, L., Runnebaum, I.B., and Dürst, M. (2012). Non-random integration of the HPV genome in cervical cancer. *PLoS One* *7*, e39632.
- Schröder, K. (2020). NADPH oxidases: Current aspects and tools. *Redox Biol.* *34*.

- Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. *Nature* *473*, 337–342.
- Schwarz, E., Freese, U.K., Gissmann, L., Mayer, W., Roggenbuck, B., Stremlau, A., and zur Hausen, H. (1985). Structure and transcription of human papillomavirus sequences in cervical carcinoma cells. *Nature* *314*, 111–114.
- Seavey, S.E., Holubar, M., Saucedo, L.J., and Perry, M.E. (1999). The E7 oncoprotein of human papillomavirus type 16 stabilizes p53 through a mechanism independent of p19(ARF). *J. Virol.* *73*, 7590–7598.
- Sedman, S.A., Barbosa, M.S., Vass, W.C., Hubbert, N.L., Haas, J.A., Lowy, D.R., and Schiller, J.T. (1991). The full-length E6 protein of human papillomavirus type 16 has transforming and trans-activating activities and cooperates with E7 to immortalize keratinocytes in culture. *J. Virol.* *65*, 4860–4866.
- Seimetz, J., Arif, W., Bangru, S., Hernaez, M., and Kalsotra, A. (2019). Cell-type specific polysome profiling from mammalian tissues. *Methods* *155*, 131–139.
- Shafabakhsh, R., Reiter, R.J., Mirzaei, H., Teymoordash, S.N., and Asemi, Z. (2019). Melatonin: A new inhibitor agent for cervical cancer treatment. *J. Cell. Physiol.* *234*, 21670–21682.
- Shai, A., Brake, T., Somoza, C., and 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.
- Shally, M., Alloul, N., Jackman, A., Muller, M., Gissmann, L., and Sherman, L. (1996). The E6 variant proteins E6I–E6IV of human papillomavirus 16: expression in cell free systems and bacteria and study of their interaction with p53. *Virus Res.* *42*, 81–96.
- Shamanin, V.A., Sekaric, P., and Androphy, E.J. (2008). hAda3 Degradation by Papillomavirus Type 16 E6 Correlates with Abrogation of the p14ARF-p53 Pathway and Efficient immortalization of Human Mammary Epithelial Cells. *J. Virol.* *82*, 3912–3920.
- Sharma, S., and Munger, K. (2020). The Role of Long Noncoding RNAs in Human Papillomavirus-associated Pathogenesis. *Pathogens* *9*.
- Shay, J.W., and Wright, W.E. (2019). Telomeres and telomerase: three decades of progress. *Nat. Rev. Genet.* *20*, 299–309.
- Sherman, L., and Alloul, N. (1992). Human papillomavirus type 16 expresses a variety of alternatively spliced mRNAs putatively encoding the E2 protein. *Virology* *191*, 953–959.
- Shi, Y. (2017). Mechanistic insights into precursor messenger RNA splicing by the spliceosome. *Nat. Rev. Mol. Cell Biol.* *18*, 655–670.
- Shi, T., and Dansen, T.B. (2020). Reactive Oxygen Species Induced p53 Activation: DNA Damage, Redox Signaling, or Both? *Antioxid. Redox Signal.* *33*, 839–859.

- Shirasawa, H., Tomita, Y., Kubota, K., Kasai, T., Sekiya, S., Takamizawa, H., and Simizu, B. (1988). Transcriptional differences of the human papillomavirus type 16 genome between precancerous lesions and invasive carcinomas. *J. Virol.* *62*, 1022–1027.
- Shirasawa, H., Tanzawa, H., Matsunaga, T., and Simizu, B. (1991). Quantitative detection of spliced E6-E7 transcripts of human papillomavirus type 16 in cervical premalignant lesions. *Virology* *184*, 795–798.
- Shulzhenko, N., Lyng, H., Sanson, G.F., and Morgun, A. (2014). Ménage à trois: an evolutionary interplay between human papillomavirus, a tumor, and a woman. *Trends Microbiol.* *22*, 345–353.
- Singh, L., Gao, Q., Kumar, A., Gotoh, T., Wazer, D.E., Band, H., Feig, L.A., and Band, V. (2003). The High-Risk Human Papillomavirus Type 16 E6 Counters the GAP Function of E6TP1 toward Small Rap G Proteins. *J. Virol.* *77*, 1614–1620.
- Slack, F.J., and Chinnaiyan, A.M. (2019). The Role of Non-coding RNAs in Oncology. *Cell* *179*, 1033–1055.
- Smal, C., Wetzler, D.E., Dantur, K.I., Chemes, L.B., Garcia-Alai, M.M., Dellarole, M., Alonso, L.G., Gaston, K., and de Prat-Gay, G. (2009). The human papillomavirus E7-E2 interaction mechanism in vitro reveals a finely tuned system for modulating available E7 and E2 proteins. *Biochemistry* *48*, 11939–11949.
- Small, W., Bacon, M.A., Bajaj, A., Chuang, L.T., Fisher, B.J., Harkenrider, M.M., Jhingran, A., Kitchener, H.C., Mileskin, L.R., Viswanathan, A.N., et al. (2017). Cervical cancer: A global health crisis. *Cancer* *123*, 2404–2412.
- Smalley Rumfield, C., Roller, N., Pellom, S.T., Schlom, J., and Jochems, C. (2020). Therapeutic Vaccines for HPV-Associated Malignancies. *ImmunoTargets Ther.* *9*, 167–200.
- Smith, E.M., Rubenstein, L.M., Haugen, T.H., Pawlita, M., and Turek, L.P. (2012). Complex Etiology Underlies Risk and Survival in Head and Neck Cancer Human Papillomavirus, Tobacco, and Alcohol: A Case for Multifactor Disease (Hindawi).
- Smotkin, D., and Wettstein, F.O. (1986). Transcription of human papillomavirus type 16 early genes in a cervical cancer and a cancer-derived cell line and identification of the E7 protein. *Proc. Natl. Acad. Sci.* *83*, 4680–4684.
- Smotkin, D., and Wettstein, F.O. (1987). The major human papillomavirus protein in cervical cancers is a cytoplasmic phosphoprotein. *J. Virol.* *61*, 1686–1689.
- Smotkin, D., Prokoph, H., and Wettstein, F.O. (1989). Oncogenic and nononcogenic human genital papillomaviruses generate the E7 mRNA by different mechanisms. *J. Virol.* *63*, 1441–1447.
- Spanos, W.C., Hoover, A., Harris, G.F., Wu, S., Strand, G.L., Anderson, M.E., Klingelutz, A.J., Hendriks, W., Bossler, A.D., and Lee, J.H. (2008). The PDZ Binding Motif of Human Papillomavirus Type 16 E6 Induces PTPN13 Loss, Which Allows Anchorage-Independent Growth and Synergizes with Ras for Invasive Growth. *J. Virol.* *82*, 2493–2500.

- Srivenugopal, K.S., and Ali-Osman, F. (2002). The DNA repair protein, O⁶-Methylguanine-DNA methyltransferase is a proteolytic target for the E6 human papillomavirus oncoprotein. *Oncogene* 21, 5940–5945.
- Stacey, S.N., Jordan, D., Snijders, P.J., Mackett, M., Walboomers, J.M., and 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.
- Stacey, S.N., Jordan, D., Williamson, A.J., Brown, M., Coote, J.H., and Arrand, J.R. (2000). Leaky scanning is the predominant mechanism for translation of human papillomavirus type 16 E7 oncoprotein from E6/E7 bicistronic mRNA. *J. Virol.* 74, 7284–7297.
- Stanley, M.A. (2012). Epithelial cell responses to infection with human papillomavirus. *Clin. Microbiol. Rev.* 25, 215–222.
- Stoler, M.H., Rhodes, C.R., Whitbeck, A., Wolinsky, S.M., Chow, L.T., and Broker, T.R. (1992). Human papillomavirus type 16 and 18 gene expression in cervical neoplasias. *Hum. Pathol.* 23, 117–128.
- Stöppler, H., Hartmann, D.-P., Sherman, L., and Schlegel, R. (1997). The Human Papillomavirus Type 16 E6 and E7 Oncoproteins Dissociate Cellular Telomerase Activity from the Maintenance of Telomere Length. *J. Biol. Chem.* 272, 13332–13337.
- Storrs, C.H., and Silverstein, S.J. (2007). PATJ, a Tight Junction-Associated PDZ Protein, Is a Novel Degradation Target of High-Risk Human Papillomavirus E6 and the Alternatively Spliced Isoform 18 E6*. *J. Virol.* 81, 4080–4090.
- Straub, E., Fertey, J., Dreer, M., Iftner, T., and Stubenrauch, F. (2015). Characterization of the Human Papillomavirus 16 E8 Promoter. *J. Virol.* 89, 7304–7313.
- Stünkel, W., and Bernard, H.-U. (1999). The Chromatin Structure of the Long Control Region of Human Papillomavirus Type 16 Represses Viral Oncoprotein Expression. *J. Virol.* 73, 1918–1930.
- Tan, S.H., Leong, L.E., Walker, P.A., and Bernard, H.U. (1994a). The human papillomavirus type 16 E2 transcription factor binds with low cooperativity to two flanking sites and represses the E6 promoter through displacement of Sp1 and TFIID. *J. Virol.* 68, 6411–6420.
- Tan, T.M., Gloss, B., Bernard, H.U., and Ting, R.C. (1994b). Mechanism of translation of the bicistronic mRNA encoding human papillomavirus type 16 E6-E7 genes. *J. Gen. Virol.* 75 (Pt 10), 2663–2670.
- Tang, S., Tao, M., McCoy, J.P., and Zheng, Z.-M. (2006). The E7 Oncoprotein Is Translated from Spliced E6*I Transcripts in High-Risk Human Papillomavirus Type 16- or Type 18-Positive Cervical Cancer Cell Lines via Translation Reinitiation. *J. Virol.* 80, 4249–4263.
- Tao, M., Kruhlak, M., Xia, S., Androphy, E., and Zheng, Z.-M. (2003). Signals That Dictate Nuclear Localization of Human Papillomavirus Type 16 Oncoprotein E6 in Living Cells. *J. Virol.* 77, 13232–13247.

- Thain, A., Jenkins, O., Clarke, A.R., and Gaston, K. (1996). CpG methylation directly inhibits binding of the human papillomavirus type 16 E2 protein to specific DNA sequences. *J. Virol.* *70*, 7233–7235.
- Thomas, M., Glaunsinger, B., Pim, D., Javier, R., and Banks, L. (2001). HPV E6 and MAGUK protein interactions: determination of the molecular basis for specific protein recognition and degradation. *Oncogene* *20*, 5431–5439.
- Thomas, M., Laura, R., Hepner, K., Guccione, E., Sawyers, C., Lasky, L., and Banks, L. (2002). Oncogenic human papillomavirus E6 proteins target the MAGI-2 and MAGI-3 proteins for degradation. *Oncogene* *21*, 5088–5096.
- Thun, M., Linet, M.S., Cerhan, J.R., Haiman, C.A., and Schottenfeld, D. (2017). *Cancer Epidemiology and Prevention* (Oxford University Press).
- Togtema, M., Hussack, G., Dayer, G., Tegtmeyer, M.R., Raphael, S., Tanha, J., and Zehbe, I. (2019). Single-Domain Antibodies Represent Novel Alternatives to Monoclonal Antibodies as Targeting Agents against the Human Papillomavirus 16 E6 Protein. *Int. J. Mol. Sci.* *20*, 2088.
- Tomaić, V., Pim, D., and Banks, L. (2009). The stability of the human papillomavirus E6 oncoprotein is E6AP dependent. *Virology* *393*, 7–10.
- Toots, M., Männik, A., Kivi, G., Jr, M.U., Ustav, E., and Ustav, M. (2014). The Transcription Map of Human Papillomavirus Type 18 during Genome Replication in U2OS Cells. *PLOS ONE* *9*, e116151.
- Tungteakkhun, S.S., Filippova, M., Fodor, N., and Duerksen-Hughes, P.J. (2010). The Full-Length Isoform of Human Papillomavirus 16 E6 and Its Splice Variant E6* Bind to Different Sites on the Procaspase 8 Death Effector Domain. *J. Virol.* *84*, 1453–1463.
- Umbach, J.L., and Cullen, B.R. (2009). The role of RNAi and microRNAs in animal virus replication and antiviral immunity. *Genes Dev.* *23*, 1151–1164.
- Underbrink, M.P., Howie, H.L., Bedard, K.M., Koop, J.I., and Galloway, D.A. (2008). E6 Proteins from Multiple Human Betapapillomavirus Types Degrade Bak and Protect Keratinocytes from Apoptosis after UVB Irradiation. *J. Virol.* *82*, 10408–10417.
- Vaeteewoottacharn, K., Chamutpong, S., Ponglikitmongkol, M., and Angeletti, P.C. (2005). Differential localization of HPV16 E6 splice products with E6-associated protein. *Virol. J.* *2*, 50.
- Vaisman, C.E., Del Moral-Hernandez, O., Moreno-Campuzano, S., Aréchaga-Ocampo, E., Bonilla-Moreno, R., Garcia-Aguiar, I., Cedillo-Barron, L., Berumen, J., Nava, P., and Villegas-Sepúlveda, N. (2018). C33-A cells transfected with E6*I or E6*II the short forms of HPV-16 E6, displayed opposite effects on cisplatin-induced apoptosis. *Virus Res.* *247*, 94–101.
- Valko, M., Leibfritz, D., Moncol, J., Cronin, M.T.D., Mazur, M., and Telser, J. (2007). Free radicals and antioxidants in normal physiological functions and human disease. *Int. J. Biochem. Cell Biol.* *39*, 44–84.

- Vambutas, A., Bonagura, V.R., and Steinberg, B.M. (2000). Altered expression of TAP-1 and major histocompatibility complex class I in laryngeal papillomatosis: correlation of TAP-1 with disease. *Clin. Diagn. Lab. Immunol.* 7, 79–85.
- Van Doorslaer, K. (2013). Evolution of the Papillomaviridae. *Virology* 445, 11–20.
- Vazquez-Vega, S., Sanchez-Suarez, L.P., Andrade-Cruz, R., Castellanos-Juarez, E., Contreras-Paredes, A., Lizano-Soberon, M., Garcia-Carranca, A., and Bribiesca, L.B. (2013). Regulation of p14ARF expression by HPV-18 E6 variants. *J. Med. Virol.* 85, 1215–1221.
- Veress, G., Szarka, K., Dong, X.P., Gergely, L., and Pfister, H. (1999). Functional significance of sequence variation in the E2 gene and the long control region of human papillomavirus type 16. *J. Gen. Virol.* 80, 1035–1043.
- de Villiers, E.-M. (2013). Cross-roads in the classification of papillomaviruses. *Virology* 445, 2–10.
- de Villiers, E.-M., Fauquet, C., Broker, T.R., Bernard, H.-U., and zur Hausen, H. (2004). Classification of papillomaviruses. *Virology* 324, 17–27.
- Vincentelli, R., Luck, K., Poirson, J., Polanowska, J., Abdat, J., Blémont, M., Turchetto, J., Iv, F., Ricquier, K., Straub, M.-L., et al. (2015). Quantifying domain-ligand affinities and specificities by high-throughput holdup assay. *Nat. Methods* 12, 787–793.
- Visalli, G., Riso, R., Facciola, A., Mondello, P., Caruso, C., Picerno, I., Di Pietro, A., Spataro, P., and Bertuccio, M.P. (2016). Higher levels of oxidative DNA damage in cervical cells are correlated with the grade of dysplasia and HPV infection. *J. Med. Virol.* 88, 336–344.
- Vousden, K.H., and Jat, P. (1989). Functional similarity between HPV16E7, SV40 large T and adenovirus E1a proteins. *Oncogene* 4, 153–158.
- Vousden, K.H., Doniger, J., DiPaolo, J.A., and Lowy, D.R. (1988). The E7 open reading frame of human papillomavirus type 16 encodes a transforming gene. *Oncogene Res.* 3, 167–175.
- Wallace, N.A., and Galloway, D.A. (2015). Novel Functions of the Human Papillomavirus E6 Oncoproteins. *Annu. Rev. Virol.* 2, 403–423.
- Wallace, N.A., Khanal, S., Robinson, K.L., Wendel, S.O., Messer, J.J., and Galloway, D.A. (2017). High-Risk Alphapapillomavirus Oncogenes Impair the Homologous Recombination Pathway. *J. Virol.* 91.
- Wang, Q., Griffin, H., Southern, S., Jackson, D., Martin, A., McIntosh, P., Davy, C., Masterson, P.J., Walker, P.A., Laskey, P., et al. (2004). Functional analysis of the human papillomavirus type 16 E1=E4 protein provides a mechanism for in vivo and in vitro keratin filament reorganization. *J. Virol.* 78, 821–833.
- Wanichwatanadecha, P., Sirisrimangkorn, S., Kaewprag, J., and Ponglikitmongkol, M. (2012). Transactivation activity of human papillomavirus type 16 E6*I on aldo-keto reductase genes enhances chemoresistance in cervical cancer cells. *J. Gen. Virol.* 93, 1081–1092.

- Weiderpass, E., Ye, W., Tamimi, R., Trichopolous, D., Nyren, O., Vainio, H., and Adami, H.-O. (2001). Alcoholism and Risk for Cancer of the Cervix Uteri, Vagina, and Vulva. *Cancer Epidemiol. Prev. Biomark.* *10*, 899–901.
- Westrich, J.A., Warren, C.J., and Pyeon, D. (2017). Evasion of host immune defenses by human papillomavirus. *Virus Res.* *231*, 21–33.
- Will, C.L., and Lührmann, R. (2011). Spliceosome Structure and Function. Cold Spring Harb. *Perspect. Biol.* *3*.
- Williams, V.M., Filippova, M., Soto, U., and Duerksen-Hughes, P.J. (2010). HPV-DNA integration and carcinogenesis: putative roles for inflammation and oxidative stress. *Future Virol.* *6*, 45–57.
- Williams, V.M., Filippova, M., Filippov, V., Payne, K.J., and Duerksen-Hughes, P. (2014). Human Papillomavirus Type 16 E6* Induces Oxidative Stress and DNA Damage. *J. Virol.* *88*, 6751–6761.
- Wooldridge, T.R., and Laimins, L.A. (2008). Regulation of human papillomavirus type 31 gene expression during the differentiation-dependent life cycle through histone modifications and transcription factor binding. *Virology* *374*, 371–380.
- Wu, P., Mo, Y., Peng, M., Tang, T., Zhong, Y., Deng, X., Xiong, F., Guo, C., Wu, X., Li, Y., et al. (2020). Emerging role of tumor-related functional peptides encoded by lncRNA and circRNA. *Mol. Cancer* *19*, 22.
- Xu, M., Luo, W., Elzi, D.J., Grandori, C., and Galloway, D.A. (2008). NFX1 Interacts with mSin3A/Histone Deacetylase To Repress hTERT Transcription in Keratinocytes. *Mol. Cell. Biol.* *28*, 4819–4828.
- Yang, E.J., Quick, M.C., Hanamornroongruang, S., Lai, K., Doyle, L.A., McKeon, F.D., Xian, W., Crum, C.P., and Herfs, M. (2015). Microanatomy of the cervical and anorectal squamocolumnar junctions: a proposed model for anatomical differences in HPV-related cancer risk. *Mod. Pathol.* *28*, 994–1000.
- Yang, G., Lu, X., and Yuan, L. (2014). LncRNA: A link between RNA and cancer. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* *1839*, 1097–1109.
- Yasumoto, S., Burkhardt, A.L., Doniger, J., and DiPaolo, J.A. (1986). Human papillomavirus type 16 DNA-induced malignant transformation of NIH 3T3 cells. *J. Virol.* *57*, 572–577.
- Yoshimatsu, Y., Nakahara, T., Tanaka, K., Inagawa, Y., Narisawa-Saito, M., Yugawa, T., Ohno, S., Fujita, M., Nakagama, H., and Kiyono, T. (2017). Roles of the PDZ-binding motif of HPV 16 E6 protein in oncogenic transformation of human cervical keratinocytes. *Cancer Sci.* *108*, 1303–1309.
- Yu, C.-H., Dang, Y., Zhou, Z., Wu, C., Zhao, F., Sachs, M.S., and Liu, Y. (2015). Codon Usage Influences the Local Rate of Translation Elongation to Regulate Co-translational Protein Folding. *Mol. Cell* *59*, 744–754.
- Zanier, K., Sidi, A. ould M. ould, Boulade-Ladame, C., Rybin, V., Chappelle, A., Atkinson, A., Kieffer, B., and Travé, G. (2012). Solution structure analysis of the HPV16 E6

oncoprotein reveals a self-association mechanism required for E6-mediated degradation of p53. *Struct. England* 1993 20, 604–617.

Zerfass-Thome, K., Zwerschke, W., Mannhardt, B., Tindle, R., Botz, J.W., and Jansen-Dürr, P. (1996). Inactivation of the cdk inhibitor p27KIP1 by the human papillomavirus type 16 E7 oncoprotein. *Oncogene* 13, 2323–2330.

Zhang, Y., Fan, S., Meng, Q., Ma, Y., Katiyar, P., Schlegel, R., and Rosen, E.M. (2005). BRCA1 Interaction with Human Papillomavirus Oncoproteins. *J. Biol. Chem.* 280, 33165–33177.

Zhang, Y., Koneva, L.A., Virani, S., Arthur, A.E., Virani, A., Hall, P.B., Warden, C.D., Carey, T.E., Chepeha, D.B., Prince, M.E., et al. (2016). Subtypes of HPV-positive head and neck cancers are associated with HPV characteristics, copy number alterations, PIK3CA mutation, and pathway signatures. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 22, 4735–4745.

Zhang, Y.-L., Wang, R.-C., Cheng, K., Ring, B.Z., and Su, L. (2017). Roles of Rap1 signaling in tumor cell migration and invasion. *Cancer Biol. Med.* 14, 90–99.

Zhao, K.-N., and Chen, J. (2011). Codon usage roles in human papillomavirus. *Rev. Med. Virol.* 21, 397–411.

Zhao, L., Wang, L., Zhang, C., Liu, Z., Piao, Y., Yan, J., Xiang, R., Yao, Y., and Shi, Y. (2019). E6-induced selective translation of WNT4 and JIP2 promotes the progression of cervical cancer via a noncanonical WNT signaling pathway. *Signal Transduct. Target. Ther.* 4, 1–10.

Zheng, Z.-M., and Baker, C.C. (2006). PAPILOMAVIRUS GENOME STRUCTURE, EXPRESSION, AND POST-TRANSCRIPTIONAL REGULATION. *Front. Biosci. J. Virtual Libr.* 11, 2286–2302.

Zheng, L., Ding, H., Lu, Z., Li, Y., Pan, Y., Ning, T., and Ke, Y. (2008). E3 ubiquitin ligase E6AP-mediated TSC2 turnover in the presence and absence of HPV16 E6. *Genes Cells* 13, 285–294.

Zheng, Y., Jönsson, J., Hao, C., Shoja Chaghervand, S., Cui, X., Kajitani, N., Gong, L., Wu, C., and Schwartz, S. (2020). Heterogeneous Nuclear Ribonucleoprotein A1 (hnRNP A1) and hnRNP A2 Inhibit Splicing to Human Papillomavirus 16 Splice Site SA409 through a UAG-Containing Sequence in the E7 Coding Region. *J. Virol.* 94.

Zheng, Z.-M., Tao, M., Yamanegi, K., Bodaghi, S., and Xiao, W. (2004). Splicing of a Cap-proximal Human Papillomavirus 16 E6E7 Intron Promotes E7 Expression, but can be Restrained by Distance of the Intron from its RNA 5' Cap. *J. Mol. Biol.* 337, 1091–1108.

(2006). Carcinoma of the cervix and tobacco smoking: Collaborative reanalysis of individual data on 13,541 women with carcinoma of the cervix and 23,017 women without carcinoma of the cervix from 23 epidemiological studies. *Int. J. Cancer* 118, 1481–1495.

Annexes

1 Données complémentaires (Paget-Bailly et al., 2019)

Tableau 3 : Differentially expressed genes U-CTL vs. U-E6All.

	Gene id	Gene name	Locus	FPKM U-CTL	FPKM U-E6All	Log2 fold change
1	ENSG00000171483	SSX6	X:48107991-48120210	0,00	0,25	9999
2	ENSG00000223495	RP11-763B22.7	1:143498784-143499102	0,00	1,02	9999
3	ENSG00000068985	PAGE1	X:49687449-49695993	0,42	82,34	7,63
4	ENSG00000007314	SCN4A	17:63938553-63972918	0,03	0,83	4,60
5	ENSG00000272636	DOC2B	17:142788-181636	0,02	0,51	4,43
6	ENSG00000279998	CH507-210P18.1	21:6789591-6812297	0,12	2,14	4,11
7	ENSG00000198732	SMOC1	14:69854130-70032366	1,97	28,85	3,87
8	ENSG00000123610	TNFAIP6	2:151357591-151380048	0,22	2,92	3,74
9	ENSG00000155962	CLIC2	X:155276210-155334657	0,43	5,61	3,72
10	ENSG00000255026	RP11-326C3.2	11:287304-288987	0,96	11,53	3,59
11	ENSG00000170367	CST5	20:23875934-23879750	0,67	7,83	3,54
12	ENSG00000184564	SLITRK6	13:85792789-85799488	0,17	1,77	3,35
13	ENSG00000196074	SYCP2	20:59863562-59933655	0,50	4,83	3,28
14	ENSG00000167723	TRPV3	17:3439922-3557995	0,10	0,92	3,26
15	ENSG00000233922	AL133493.2	21:45593653-45603056	0,11	1,02	3,22
16	ENSG00000283293	RN7SK	6:52995619-52995950	9,78	89,68	3,20
17	ENSG00000198732	SMOC1	14:69854130-70032366	1,71	15,57	3,18
18	ENSG00000180730	SHISA2	13:26044596-26051031	0,42	3,33	2,99
19	ENSG00000112769	LAMA4	6:112108759-112351296	0,32	2,27	2,83
20	ENSG00000257743	MGAM2	7:142111748-142222324	0,07	0,51	2,82
21	ENSG00000277027	RMRP	9:35657750-35658018	10,09	69,91	2,79
22	ENSG00000258585	ZNF519P3	15:21911558-21912590	0,21	1,43	2,75
23	ENSG00000018236	CNTN1	12:40692441-41072418	0,89	5,87	2,72
24	ENSG00000116329	OPRD1	1:28812141-28877336	0,06	0,38	2,62
25	ENSG00000283431	RP11-1217F2.22	7:57163167-57166024	0,42	2,57	2,61
26	ENSG00000134323	MYCN	2:15920398-15947007	0,49	2,96	2,59
27	ENSG00000185483	ROR1	1:63774021-64181498	1,88	11,16	2,57
28	ENSG00000197705	KLHL14	18:32672670-32774413	0,13	0,69	2,45
29	ENSG00000197020	ZNF100	19:21722765-21767628	0,18	0,95	2,38
30	ENSG00000049323	LTBP1	2:32946971-33399509	1,92	9,97	2,37
31	ENSG00000147408	CSGALNACT1	8:19404160-19758029	0,53	2,71	2,36
32	ENSG00000125872	LRRN4	20:6040777-6054049	0,21	1,05	2,35
33	ENSG00000196604	POTEF	2:130074029-130129222	0,57	2,88	2,33
34	ENSG00000162618	ADGRL4	1:78889763-79006718	1,45	7,04	2,28
35	ENSG00000134758	RNF138	18:32018371-32470484	3,04	14,21	2,22
36	ENSG00000134532	SOX5	12:23529499-23951032	0,31	1,35	2,14
37	ENSG00000261514	RP11-527L4.2	17:43938362-43938959	1,91	8,32	2,13
38	ENSG00000115419	GLS	2:190880796-191020960	20,62	88,44	2,10
39	ENSG00000175899	A2M	12:9065176-9116229	0,74	3,13	2,08

40	ENSG00000170454	KRT75	12:52407579-52434525	4,82	20,33	2,08
41	ENSG00000171385	KCND3	1:111770661-111989155	0,18	0,74	2,06
42	ENSG00000144227	NXPH2	2:138670771-138780348	0,39	1,61	2,04
43	ENSG00000120159	CAAP1	9:26840684-26892804	3,29	13,37	2,02
44	ENSG00000198205	ZXDA	X:57906707-57910820	0,91	3,59	1,98
45	ENSG00000253288	RP11-238K6.1	8:137809443-138083570	0,68	2,65	1,96
46	ENSG00000270392	PFN1P2	1:120432203-120434052	1,52	5,90	1,96
47	ENSG00000000971	CFH	1:196651877-196747504	1,00	3,81	1,93
48	ENSG00000255099	RP11-25D10.2	8:109777911-109782288	0,14	0,54	1,93
49	ENSG00000153956	CACNA2D1	7:81946443-82443798	2,16	8,20	1,92
50	ENSG00000260073	RP11-488I20.9	16:35192686-35195661	0,68	2,56	1,91
51	ENSG00000187068	C3orf70	3:185078049-185153014	0,47	1,78	1,91
52	ENSG00000279795	RP11-586K12.1	16:32787961-32788944	0,81	3,01	1,89
53	ENSG00000279561	RP11-4L24.4	9:41269915-41273997	0,86	3,18	1,88
54	ENSG00000091039	OSBPL8	12:76351796-76559809	1,98	7,26	1,87
55	ENSG00000189057	FAM111B	11:59107184-59127410	10,75	39,37	1,87
56	ENSG00000104321	TRPA1	8:71828166-72118393	0,95	3,49	1,87
57	ENSG00000143816	WNT9A	1:227918655-227947898	0,24	0,87	1,86
58	ENSG00000175445	LPL	8:19901716-19967258	2,88	10,42	1,86
59	ENSG00000120162	MOB3B	9:27325208-27529781	0,35	1,26	1,85
60	ENSG00000183206	POTEC	18:14507338-14543585	0,49	1,74	1,84
61	ENSG00000172061	LRRC15	3:194355246-194369743	1,23	4,39	1,84
62	ENSG00000196834	POTEI	2:130459454-130509666	2,19	7,78	1,83
63	ENSG00000119888	EPCAM	2:47345157-47387601	2,67	9,36	1,81
64	ENSG00000138336	TET1	10:68560655-68694482	0,37	1,29	1,81
65	ENSG00000135378	PRRG4	11:32829942-32858123	0,20	0,71	1,80
66	ENSG00000135549	PKIB	6:122471916-122726373	17,28	59,47	1,78
67	ENSG00000128791	TWSG1	18:9334766-9402420	7,61	25,87	1,76
68	ENSG00000162980	ARL5A	2:151788983-152109202	2,05	6,92	1,76
69	ENSG00000176597	B3GNT5	3:183178042-183428778	2,51	8,44	1,75
70	ENSG00000181827	RFX7	15:56087279-56243266	1,75	5,80	1,73
71	ENSG00000133731	IMPA1	8:81657960-81686693	4,64	15,41	1,73
72	ENSG00000164983	TMEM65	8:124306188-124372692	3,09	10,22	1,73
73	ENSG00000150764	DIXDC1	11:111927143-112022584	1,09	3,59	1,72
74	ENSG00000180354	MTURN	7:30134809-30162762	2,64	8,57	1,70
75	ENSG00000036054	TBC1D23	3:100260999-100325251	12,34	39,95	1,69
76	ENSG00000197410	DCHS2	4:154142121-154491716	1,01	3,21	1,67
77	ENSG00000157625	TAB3	X:30827441-30975084	2,16	6,82	1,66
78	ENSG00000164023	SGMS2	4:107824562-107978799	2,78	8,76	1,65
79	ENSG00000177888	ZBTB41	1:197153679-197200542	1,29	4,04	1,64
80	ENSG00000241860	RP11-34P13.13	1:141473-173862	1,12	3,48	1,64
81	ENSG00000164761	TNFRSF11B	8:118923556-118952200	6,21	19,31	1,64
82	ENSG00000179008	C14orf39	14:60396468-60554916	1,96	6,07	1,63
83	ENSG00000170802	FOXN2	2:48314636-48379294	3,34	10,36	1,63
84	ENSG00000136603	SKIL	3:170357677-170396835	4,06	12,52	1,62
85	ENSG00000145390	USP53	4:119192772-119295517	1,95	6,00	1,62

86	ENSG00000253741	CTD-2292P10.4	8:142700110-142726973	1,80	5,49	1,61
87	ENSG00000100697	DICER1	14:95086227-95179933	2,59	7,85	1,60
88	ENSG00000153404	PLEKHG4B	5:92150-189972	0,68	2,04	1,59
89	ENSG00000197329	PEL1	2:64092651-64252859	1,55	4,65	1,58
90	ENSG00000102780	DGKH	13:42040035-42256578	0,37	1,11	1,58
91	ENSG00000166250	CLMP	11:123069864-123228277	1,58	4,70	1,58
92	ENSG00000172197	MBOAT1	6:20102144-20317739	1,68	4,96	1,56
93	ENSG00000179869	ABCA13	7:48171457-48647496	1,35	3,95	1,55
94	ENSG00000178202	KDEL2	11:108472104-108498432	9,59	28,11	1,55
95	ENSG00000172059	KLF11	2:10042848-10054866	2,96	8,68	1,55
96	ENSG00000101888	NXT2	X:109535780-109544690	2,04	5,91	1,54
97	ENSG00000172461	FUT9	6:96015983-96215612	0,33	0,97	1,54
98	ENSG00000019186	CYP24A1	20:54153448-54173973	2,63	7,61	1,53
99	ENSG00000198455	ZXDB	X:57591651-57597545	3,01	8,70	1,53
100	ENSG00000139687	RB1	13:48303750-48481986	4,84	13,95	1,53
101	ENSG00000224559	LINC01087	2:131637024-131649615	1,58	4,52	1,52
102	ENSG00000170340	B3GNT2	2:62196112-62224731	3,77	10,77	1,51
103	ENSG00000099256	PRTFDC1	10:24848606-24952604	3,81	10,86	1,51
104	ENSG00000130303	BST2	19:17402938-17405648	25,30	71,40	1,50
105	ENSG00000075223	SEMA3C	7:80742537-80922359	7,24	20,43	1,50
106	ENSG00000113504	SLC12A7	5:1050375-1112035	26,41	74,14	1,49
107	ENSG00000152208	GRID2	4:92297250-93810157	2,82	7,86	1,48
108	ENSG00000139618	BRCA2	13:32315473-32400266	2,03	5,66	1,48
109	ENSG00000152332	UHMK1	1:162497250-162529629	4,63	12,89	1,48
110	ENSG00000163848	ZNF148	3:125225560-125375354	2,28	6,34	1,47
111	ENSG00000126821	SGPP1	14:63684213-63728039	7,14	19,79	1,47
112	ENSG00000175548	ALG10B	12:38316577-38329728	1,85	5,10	1,46
113	ENSG00000259976	MIR568	3:114314500-115147271	1,81	5,00	1,46
114	ENSG00000146247	PHIP	6:78867471-79078236	3,57	9,72	1,45
115	ENSG00000095203	EPB41L4B	9:109171974-109320964	2,01	5,46	1,44
116	ENSG00000152818	UTRN	6:144285700-144853034	4,78	12,89	1,43
117	ENSG00000164031	DNAJB14	4:99896247-99946726	3,45	9,30	1,43
118	ENSG00000150893	FREM2	13:38687128-38887131	0,33	0,89	1,43
119	ENSG00000204186	ZDBF2	2:206274662-206314428	1,20	3,22	1,42
120	ENSG00000132356	PRKAA1	5:40759378-40798374	7,35	19,67	1,42
121	ENSG00000141429	GALNT1	18:35581116-35711834	11,42	30,51	1,42
122	ENSG00000105855	ITGB8	7:20328298-20415754	2,74	7,31	1,42
123	ENSG00000254535	PABPC4L	4:134196332-134201748	1,29	3,45	1,41
124	ENSG00000131023	LATS1	6:149658152-149718888	2,53	6,71	1,41
125	ENSG00000023516	AKAP11	13:42272152-42323267	2,62	6,95	1,41
126	ENSG00000151151	IPMK	10:58191516-58267934	1,81	4,80	1,41
127	ENSG00000122335	SERAC1	6:158109514-158168270	1,50	3,96	1,40
128	ENSG00000101596	SMCHD1	18:2655737-2833065	7,91	20,86	1,40
129	ENSG00000102531	FNDC3A	13:48974966-49209779	3,09	8,13	1,40
130	ENSG00000074695	LMAN1	18:59327822-59359962	27,74	73,08	1,40
131	ENSG00000148943	LIN7C	11:27494575-27506773	3,86	10,16	1,40

132	ENSG00000056277	ZNF280C	X:130202710-130268899	1,03	2,71	1,40
133	ENSG00000198774	RASSF9	12:85800696-85836570	0,72	1,89	1,40
134	ENSG00000140396	NCOA2	8:70109761-70403805	2,28	6,00	1,39
135	ENSG00000198300	PEG3	19:56690685-56840729	3,30	8,66	1,39
136	ENSG00000163625	WDFY3	4:84669609-85011277	2,63	6,91	1,39
137	ENSG00000114450	GNB4	3:179396088-179451590	3,67	9,61	1,39
138	ENSG00000082258	CCNT2	2:134735463-134959342	4,42	11,57	1,39
139	ENSG00000111885	MAN1A1	6:119177208-119349761	11,55	30,00	1,38
140	ENSG00000166341	DCHS1	11:6621322-6655854	3,12	8,07	1,37
141	ENSG00000151458	ANKRD50	4:124664051-124712732	2,69	6,96	1,37
142	ENSG00000165105	RASEF	9:82979584-83063177	1,61	4,15	1,37
143	ENSG00000186130	ZBTB6	9:122908055-122913330	1,31	3,35	1,36
144	ENSG00000165244	ZNF367	9:96385940-96418329	7,78	19,89	1,35
145	ENSG00000174607	UGT8	4:114598454-114678224	5,00	12,75	1,35
146	ENSG00000172469	MANEA	6:95577542-95609457	2,08	5,27	1,34
147	ENSG00000182010	RTKN2	10:62183034-62268707	3,52	8,90	1,34
148	ENSG00000136536	MARCH7	2:159712456-159904749	9,06	22,72	1,33
149	ENSG00000122641	INHBA	7:41667167-41779388	7,59	18,98	1,32
150	ENSG00000156011	PSD3	8:18527300-19084730	1,69	4,22	1,32
151	ENSG00000085449	WDFY1	2:223855715-223945387	7,58	18,89	1,32
152	ENSG00000196159	FAT4	4:125316398-125492932	0,66	1,63	1,31
153	ENSG00000112972	HMGCS1	5:43287600-43313512	25,45	62,92	1,31
154	ENSG00000100934	SEC23A	14:39031918-39109646	9,17	22,60	1,30
155	ENSG00000198887	SMC5	9:70258961-70354888	7,09	17,36	1,29
156	ENSG00000176406	RIMS2	8:103483397-104256094	1,98	4,85	1,29
157	ENSG00000132031	MATN3	2:19990216-20012694	4,73	11,56	1,29
158	ENSG00000188647	PTAR1	9:69709521-69759959	3,41	8,30	1,28
159	ENSG00000112531	QKI	6:163413999-163578596	12,60	30,64	1,28
160	ENSG00000229373	LINC00452	13:113888305-113923512	8,74	21,21	1,28
161	ENSG00000175895	PLEKHF2	8:95133803-95156684	2,22	5,35	1,27
162	ENSG00000014123	UFL1	6:96358373-96555276	4,32	10,43	1,27
163	ENSG00000156869	FRRS1	1:99708702-99766631	3,43	8,29	1,27
164	ENSG00000118596	SLC16A7	12:59596066-59789855	0,73	1,75	1,27
165	ENSG00000184845	DRD1	5:175440038-175444208	1,83	4,41	1,27
166	ENSG00000143333	RGS16	1:182598622-182604408	3,40	8,19	1,27
167	ENSG00000115159	GPD2	2:156435289-156613735	6,36	15,33	1,27
168	ENSG00000162616	DNAJB4	1:77979174-78138449	4,53	10,89	1,26
169	ENSG00000139324	TMTC3	12:88142295-88199887	3,65	8,76	1,26
170	ENSG00000182263	FIGN	2:163593395-163736012	2,20	5,29	1,26
171	ENSG00000204177	BMS1P1	10:46753604-46811989	2,09	4,97	1,25
172	ENSG00000188994	ZNF292	6:87151158-87264196	2,24	5,32	1,25
173	ENSG00000136492	BRIP1	17:61681265-61863521	7,08	16,75	1,24
174	ENSG00000076716	GPC4	X:133300102-133415490	5,61	13,26	1,24
175	ENSG00000169967	MAP3K2	2:127298729-127388465	3,60	8,48	1,24
176	ENSG00000085224	ATRX	X:77504877-77786269	5,81	13,67	1,24
177	ENSG00000115461	IGFBP5	2:216672104-216994079	13,08	30,71	1,23

178	ENSG00000173068	BNC2	9:16409502-16870843	1,12	2,63	1,23
179	ENSG00000046604	DSG2	18:31498042-31556911	14,49	34,00	1,23
180	ENSG00000196584	XRCC2	7:152644778-152676165	2,32	5,43	1,23
181	ENSG00000162722	TRIM58	1:247857198-247878205	3,88	9,09	1,23
182	ENSG00000001561	ENPP4	6:46129992-46146699	4,18	9,80	1,23
183	ENSG00000105810	CDK6	7:92604920-92917187	7,26	16,94	1,22
184	ENSG00000163428	LRRC58	3:120324508-120349339	9,71	22,66	1,22
185	ENSG00000240583	AQP1	7:30651942-30925516	5,38	12,48	1,21
186	ENSG00000123066	MED13L	12:115953871-116277338	2,44	5,61	1,20
187	ENSG00000155903	RASA2	3:141487046-141615342	3,21	7,39	1,20
188	ENSG00000166479	TMX3	18:68673687-69055189	10,89	25,03	1,20
189	ENSG00000134755	DSC2	18:31058839-31162856	1,76	4,03	1,20
190	ENSG00000203668	CHML	1:241532133-241802133	10,16	23,28	1,20
191	ENSG00000156140	ADAMTS3	4:72280968-72569386	3,80	8,69	1,19
192	ENSG00000164176	EDIL3	5:83940553-84490765	32,00	72,90	1,19
193	ENSG00000156802	ATAD2	8:123319849-123416350	40,50	91,83	1,18
194	ENSG00000172845	SP3	2:173906458-173965702	13,88	31,43	1,18
195	ENSG00000091656	ZFHX4	8:76403997-76867285	1,27	2,87	1,18
196	ENSG00000060749	QSER1	11:32893177-32993316	7,70	17,38	1,17
197	ENSG00000136051	KIAA1033	12:105107323-105169134	3,79	8,54	1,17
198	ENSG00000174799	CEP135	4:55948870-56033363	5,94	13,33	1,17
199	ENSG00000198087	CD2AP	6:47477788-47627263	7,91	17,74	1,16
200	ENSG00000119541	VPS4B	18:63389189-63422483	6,84	15,26	1,16
201	ENSG00000135968	GCC2	2:108448560-108534196	5,67	12,64	1,16
202	ENSG00000166128	RAB8B	15:63189468-63267782	6,46	14,35	1,15
203	ENSG00000165997	ARL5B	10:18659404-18681639	4,36	9,66	1,15
204	ENSG00000152749	GPR180	13:94601902-94634645	2,33	5,15	1,15
205	ENSG00000132294	EFR3A	8:131904087-132013642	14,66	32,31	1,14
206	ENSG00000135913	USP37	2:218419091-218568361	3,45	7,60	1,14
207	ENSG00000156273	BACH1	21:29193479-29940033	5,72	12,44	1,12
208	ENSG00000198862	LTN1	21:28928143-28992956	7,98	17,32	1,12
209	ENSG00000175893	ZDHHC21	9:14611070-14693471	1,66	3,60	1,12
210	ENSG00000142149	HUNK	21:31873314-32044633	1,94	4,20	1,12
211	ENSG00000173705	SUSD5	3:33150041-33219215	2,98	6,45	1,11
212	ENSG00000123091	RNF11	1:51236270-51273455	14,10	30,36	1,11
213	ENSG00000104290	FZD3	8:28494204-28574268	1,36	2,94	1,11
214	ENSG00000145604	SKP2	5:36151988-36184319	12,55	26,98	1,10
215	ENSG00000125398	SOX9	17:72021850-72237203	16,66	35,74	1,10
216	ENSG00000114331	ACAP2	3:195274735-195443078	4,62	9,90	1,10
217	ENSG00000087338	GMCL1	2:69829641-69881396	6,98	14,90	1,09
218	ENSG00000157106	SMG1	16:18803082-18926454	6,64	14,16	1,09
219	ENSG00000213160	KLHL23	2:169694453-169776989	5,84	12,42	1,09
220	ENSG00000187325	TAF9B	X:78129747-78139706	10,75	22,83	1,09
221	ENSG00000157500	APPL1	3:57227736-57292682	4,68	9,93	1,08
222	ENSG00000124789	NUP153	6:17615034-17707344	7,37	15,61	1,08
223	ENSG00000145780	FEM1C	5:115520907-115544894	2,90	6,12	1,08

224	ENSG00000279184	RP3-323A16.1	22:35119823-35231056	0,62	1,32	1,08
225	ENSG00000197323	TRIM33	1:114392776-114511160	5,15	10,87	1,08
226	ENSG00000080839	RBL1	20:36996348-37097178	3,93	8,28	1,07
227	ENSG00000031081	ARHGAP31	3:119294372-119420714	1,67	3,52	1,07
228	ENSG00000132182	NUP210	3:13316234-13420309	9,85	20,70	1,07
229	ENSG00000132953	XPO4	13:20777328-20903048	2,21	4,65	1,07
230	ENSG00000064313	TAF2	8:119730774-119832863	8,57	18,01	1,07
231	ENSG00000168615	ADAM9	8:38996868-39105144	58,66	123,07	1,07
232	ENSG00000204217	BMPR2	2:202376935-202567751	4,58	9,59	1,07
233	ENSG00000143815	LBR	1:225401501-225428925	36,32	75,89	1,06
234	ENSG00000163781	TOPBP1	3:133598174-133661893	18,56	38,53	1,05
235	ENSG00000109586	GALNT7	4:173131927-173329694	14,82	30,72	1,05
236	ENSG00000116774	OLFML3	1:113979390-114035572	24,95	51,61	1,05
237	ENSG00000061676	NCKAP1	2:182909114-183038858	31,13	64,36	1,05
238	ENSG00000108510	MED13	17:61942604-62065282	8,75	18,10	1,05
239	ENSG00000104219	ZDHC2	8:17156028-17224799	7,30	15,07	1,05
240	ENSG00000181222	POLR2A	17:7484365-7514616	7,99	16,48	1,05
241	ENSG00000170759	KIF5B	10:32009009-32056431	41,15	84,86	1,04
242	ENSG00000152661	GJA1	6:121435691-121449727	17,33	35,41	1,03
243	ENSG00000174106	LEMD3	12:65169570-65248327	5,97	12,15	1,03
244	ENSG00000165417	GTF2A1	14:81175451-81222460	3,70	7,52	1,02
245	ENSG00000115808	STRN	2:36843639-36966472	5,89	11,96	1,02
246	ENSG00000065243	PKN2	1:88537512-88836255	9,53	19,34	1,02
247	ENSG00000198315	ZKSCAN8	6:28141909-28159472	3,85	7,80	1,02
248	ENSG00000173517	PEAK1	15:77100655-77420144	3,85	7,77	1,01
249	ENSG00000137962	ARHGAP29	1:94148987-94334848	26,94	54,35	1,01
250	ENSG00000176624	MEX3C	18:51174549-51218304	11,12	22,30	1,00
251	ENSG00000103657	HERC1	15:63608617-63833942	2,58	5,16	1,00
252	ENSG00000143195	ILDR2	1:166895710-166975482	2,97	5,92	1,00
253	ENSG00000155304	HSPA13	21:14371114-14383484	21,78	43,09	0,98
254	ENSG00000253729	PRKDC	8:47773107-47960183	26,27	51,85	0,98
255	ENSG00000120800	UTP20	12:101280108-101386616	3,72	7,32	0,98
256	ENSG00000001629	ANKIB1	7:92246233-92401384	7,19	14,04	0,97
257	ENSG00000054983	GALC	14:87710418-87993665	11,49	22,39	0,96
258	ENSG00000131711	MAP1B	5:72107233-72209570	9,88	19,11	0,95
259	ENSG00000138829	FBN2	5:128257908-129033642	15,33	29,31	0,93
260	ENSG00000274827	LINC01297	14:19344577-19384587	15,17	28,98	0,93
261	ENSG00000134250	NOTCH2	1:119911552-120069626	15,12	28,18	0,90
262	ENSG00000197601	FAR1	11:13668669-13732346	13,39	24,89	0,89
263	ENSG00000169504	CLIC4	1:24745356-24844324	61,24	113,50	0,89
264	ENSG00000074590	NUAK1	12:106063339-106140033	7,81	14,08	0,85
265	ENSG00000092621	PHGDH	1:119659797-119744215	252,20	142,86	-0,82
266	ENSG00000177954	RPS27	1:153990758-153992150	4592,41	2417,38	-0,93
267	ENSG00000126432	PRDX5	11:64318087-64321811	475,98	249,37	-0,93
268	ENSG00000129562	DAD1	14:22564904-22589269	205,97	107,77	-0,93
269	ENSG00000112306	RPS12	6:132814440-132817564	2834,93	1457,82	-0,96

270	ENSG00000136444	RSAD1	17:50478799-50485975	36,63	18,70	-0,97
271	ENSG00000197756	RPL37A	2:216385287-216579180	4458,80	2240,93	-0,99
272	ENSG00000182636	NDN	15:23685399-23687330	70,28	34,82	-1,01
273	ENSG00000164405	UQCRC	5:132861180-132868031	1070,61	529,25	-1,02
274	ENSG00000143947	RPS27A	2:55174790-55235853	2072,73	1019,05	-1,02
275	ENSG00000106366	SERPINE1	7:101127088-101139266	58,01	28,38	-1,03
276	ENSG00000171314	PGAM1	10:97426159-97433441	352,05	170,55	-1,05
277	ENSG00000090013	BLVRB	19:40447788-40465840	220,70	106,88	-1,05
278	ENSG00000198355	PIM3	22:49960512-49964080	44,96	21,76	-1,05
279	ENSG00000128272	ATF4	22:39519694-39522685	282,52	134,83	-1,07
280	ENSG00000196683	TOMM7	7:22812627-22822852	214,42	102,25	-1,07
281	ENSG00000169100	SLC25A6	X:1386151-1392724	100,26	47,64	-1,07
282	ENSG00000169189	NSMCE1	16:27224990-27290492	59,19	27,93	-1,08
283	ENSG00000167863	ATP5H	17:75032574-75065889	578,41	272,55	-1,09
284	ENSG00000142534	RPS11	19:49496364-49499689	3339,98	1572,18	-1,09
285	ENSG00000166189	HPS6	10:102065389-102068038	10,45	4,88	-1,10
286	ENSG00000205426	KRT81	12:52232519-52309163	535,59	250,07	-1,10
287	ENSG00000008988	RPS20	8:56067294-56074581	2361,34	1102,28	-1,10
288	ENSG00000083845	RPS5	19:58386399-58394806	1951,10	908,20	-1,10
289	ENSG00000169020	ATP5I	4:672435-689441	343,57	159,16	-1,11
290	ENSG00000182899	RPL35A	3:197889074-197960142	1391,35	641,55	-1,12
291	ENSG00000198918	RPL39	X:119786503-119791643	1666,44	765,21	-1,12
292	ENSG00000171858	RPS21	20:62387115-62388520	3596,88	1646,81	-1,13
293	ENSG00000171206	TRIM8	10:102644495-102658407	29,05	13,26	-1,13
294	ENSG00000168685	IL7R	5:35852694-35879603	34,17	15,58	-1,13
295	ENSG00000107816	LZTS2	10:100996617-101031157	43,58	19,84	-1,14
296	ENSG00000087074	PPP1R15A	19:48872391-48876057	20,72	9,39	-1,14
297	ENSG00000187051	RPS19BP1	22:39529092-39532855	168,10	75,04	-1,16
298	ENSG00000197956	S100A6	1:153533602-153536244	637,01	282,79	-1,17
299	ENSG00000170889	RPS9	19:54200741-54249003	1336,77	591,21	-1,18
300	ENSG00000145592	RPL37	5:40825261-40835335	2494,05	1101,82	-1,18
301	ENSG00000105193	RPS16	19:39433206-39435948	2165,20	954,64	-1,18
302	ENSG00000163874	ZC3H12A	1:37474551-37484379	10,94	4,82	-1,18
303	ENSG00000100316	RPL3	22:39312881-39320389	1669,73	735,25	-1,18
304	ENSG00000091947	TMEM101	17:44011187-44023946	49,70	21,72	-1,19
305	ENSG00000165886	UBTD1	10:97498867-97571209	22,51	9,80	-1,20
306	ENSG00000182154	MRPL41	9:137551198-137552555	65,76	28,59	-1,20
307	ENSG00000128165	ADM2	22:50481555-50486440	6,99	3,03	-1,21
308	ENSG00000260708	CTA-29F11.1	22:46761893-46762563	38,47	16,50	-1,22
309	ENSG00000105379	ETFB	19:51345168-51369003	100,03	42,58	-1,23
310	ENSG00000128185	DGCR6L	22:20314275-20320080	70,40	29,90	-1,24
311	ENSG00000233276	GPX1	3:49357175-49358600	332,74	140,04	-1,25
312	ENSG00000131094	C1QL1	17:44959692-44968071	148,83	62,35	-1,26
313	ENSG00000149541	B3GAT3	11:62615295-62622175	36,83	15,28	-1,27
314	ENSG00000137331	IER3	6:30742928-30744554	119,75	49,53	-1,27
315	ENSG00000233927	RPS28	19:8302126-8343262	1473,69	608,67	-1,28

316	ENSG00000124145	SDC4	20:45325287-45348424	125,75	51,70	-1,28
317	ENSG00000125691	RPL23	17:38847864-38853843	1695,31	695,17	-1,29
318	ENSG00000237550	RPL9P9	15:82355141-82439153	415,29	169,61	-1,29
319	ENSG00000233913	RPL10P9	5:168616351-168616996	132,64	53,66	-1,31
320	ENSG00000105640	RPL18A	19:17859875-17864153	1747,96	706,56	-1,31
321	ENSG00000131469	RPL27	17:42998272-43002959	3107,87	1255,06	-1,31
322	ENSG00000136048	DRAM1	12:101877350-102012130	7,23	2,91	-1,32
323	ENSG00000163682	RPL9	4:39454123-39458948	1245,35	497,47	-1,32
324	ENSG00000215030	RPL13P12	17:17383376-17384012	110,37	44,08	-1,32
325	ENSG00000112667	DNPH1	6:43225628-43229484	56,21	22,28	-1,34
326	ENSG00000114268	PFKFB4	3:48517683-48563773	14,20	5,59	-1,35
327	ENSG00000108298	RPL19	17:39200282-39204727	4866,34	1906,62	-1,35
328	ENSG00000130193	THEM6	8:142727202-142736927	11,47	4,44	-1,37
329	ENSG00000142546	NOSIP	19:49555710-49591015	85,19	32,86	-1,37
330	ENSG00000105372	RPS19	19:41859917-41872926	5160,69	1973,70	-1,39
331	ENSG00000159388	BTG2	1:203305490-203309602	12,86	4,91	-1,39
332	ENSG00000162244	RPL29	3:51993599-51995942	925,22	351,93	-1,39
333	ENSG00000155961	RAB39B	X:155258240-155264589	3,94	1,49	-1,41
334	ENSG00000167925	GHDC	17:42188798-42194532	13,60	5,07	-1,42
335	ENSG00000172216	CEBPB	20:50184597-50192689	45,14	16,82	-1,42
336	ENSG00000270069	MIR222HG	X:45745210-45770274	29,74	11,00	-1,43
337	ENSG00000240476	LINC00973	3:98902423-99018562	78,06	28,88	-1,43
338	ENSG00000128422	KRT17	17:41619436-41624842	267,00	98,68	-1,44
339	ENSG00000123933	MXD4	4:2247431-2262294	5,90	2,17	-1,45
340	ENSG00000187837	HIST1H1C	6:26055786-26056428	448,01	163,66	-1,45
341	ENSG00000108691	CCL2	17:34255217-34257203	70,91	25,88	-1,45
342	ENSG00000102109	PCSK1N	X:48831093-48835633	27,24	9,93	-1,46
343	ENSG00000136997	MYC	8:127735433-127741434	49,13	17,81	-1,46
344	ENSG00000172809	RPL38	17:74203581-74213342	1407,76	509,55	-1,47
345	ENSG00000107874	CUEDC2	10:102423244-102432661	71,72	25,60	-1,49
346	ENSG00000132563	REEP2	5:138439016-138446969	11,40	4,02	-1,50
347	ENSG00000130766	SESN2	1:28259526-28282491	18,07	6,28	-1,52
348	ENSG00000221821	C6orf226	6:42890264-42890816	29,08	10,11	-1,52
349	ENSG00000240972	MIF	22:23856426-23898930	563,80	195,91	-1,53
350	ENSG00000273576	RP11-390P24.1	17:39566914-39567559	12,08	4,16	-1,54
351	ENSG00000148834	GSTO1	10:104235355-104267459	89,60	30,06	-1,58
352	ENSG00000167747	C19orf48	19:50797703-50804929	77,32	25,87	-1,58
353	ENSG00000164587	RPS14	5:150443189-150449756	1632,22	533,34	-1,61
354	ENSG00000163584	RPL22L1	3:170864874-170870483	576,72	186,97	-1,63
355	ENSG00000213741	RPS29	14:49570983-49614672	4649,18	1482,79	-1,65
356	ENSG00000142541	RPL13A	19:49446297-49492308	1492,38	473,69	-1,66
357	ENSG00000185262	UBALD2	17:76265201-76271299	42,36	13,34	-1,67
358	ENSG00000130748	TMEM160	19:47045906-47048630	22,44	7,00	-1,68
359	ENSG00000276851	RP5-875H18.9	17:50094064-50094647	7,49	2,30	-1,70
360	ENSG00000273619	RP5-908M14.9	20:62386302-62386970	6,73	2,04	-1,72
361	ENSG00000167644	C19orf33	19:38244034-38305009	162,23	48,93	-1,73

362	ENSG00000260025	RP11-490M8.1	2:36354748-36355114	40,72	12,16	-1,74
363	ENSG00000274213	RP11-670E13.6	17:56914185-56914533	23,68	7,06	-1,75
364	ENSG00000116183	PAPPA2	1:176463170-176845605	0,91	0,27	-1,76
365	ENSG00000174307	PHLDA3	1:201464382-201469237	25,53	7,54	-1,76
366	ENSG00000248367	CTB-129O4.1	5:180293244-180295253	12,30	3,60	-1,77
367	ENSG00000260563	RP13-516M14.1	17:82293715-82294910	12,37	3,60	-1,78
368	ENSG00000126709	IFI6	1:27666060-27703063	41,67	12,08	-1,79
369	ENSG00000163347	CLDN1	3:190305700-190322475	12,44	3,54	-1,81
370	ENSG00000141753	IGFBP4	17:40443460-40457731	54,23	15,24	-1,83
371	ENSG00000137673	MMP7	11:102520507-102530753	2,65	0,74	-1,85
372	ENSG00000260260	SNHG19	16:2154796-2155358	451,71	124,39	-1,86
373	ENSG00000221869	CEBPD	8:47736908-47739086	8,74	2,39	-1,87
374	ENSG00000101255	TRIB3	20:362834-397559	38,20	10,29	-1,89
375	ENSG00000138395	CDK15	2:201790460-201895550	1,16	0,31	-1,89
376	ENSG00000175602	CCDC85B	11:65890111-65891635	78,59	20,64	-1,93
377	ENSG00000105373	GLTSCR2	19:47745521-47768840	177,93	46,51	-1,94
378	ENSG00000145287	PLAC8	4:83090047-83137075	53,58	13,95	-1,94
379	ENSG00000114854	TNNC1	3:52451101-52454070	92,87	24,13	-1,94
380	ENSG00000187608	ISG15	1:1001137-1014541	50,09	13,00	-1,95
381	ENSG00000246705	H2AFJ	12:14774382-14778002	106,60	27,20	-1,97
382	ENSG00000125726	CD70	19:6583182-6604103	50,66	12,74	-1,99
383	ENSG00000131738	KRT33B	17:41363493-41369800	1,63	0,39	-2,04
384	ENSG00000272449	RP3-395M20.12	1:2546464-2547460	1,80	0,42	-2,09
385	ENSG00000105246	EBI3	19:4229497-4237531	8,21	1,88	-2,12
386	ENSG00000111245	MYL2	12:110910818-110920722	12,94	2,95	-2,13
387	ENSG00000136167	LCP1	13:46125919-46211871	4,49	1,02	-2,14
388	ENSG00000266947	RP11-799D4.4	17:35231449-35242963	1,02	0,23	-2,16
389	ENSG00000128965	CHAC1	15:40952961-40956519	20,37	4,50	-2,18
390	ENSG00000174951	FUT1	19:48748010-48755390	2,96	0,65	-2,18
391	ENSG00000225663	MCRIP1	17:81822360-81833302	56,70	12,39	-2,19
392	ENSG00000163661	PTX3	3:157175222-157533619	29,40	6,41	-2,20
393	ENSG00000161513	FDXR	17:74862496-74873031	70,87	14,90	-2,25
394	ENSG00000250420	AACSP1	5:178764860-178818435	3,02	0,62	-2,27
395	ENSG00000110195	FOLR1	11:72163321-72209241	14,78	2,93	-2,33
396	ENSG00000265972	TXNIP	1:145992434-145996600	7,15	1,41	-2,34
397	ENSG00000269926	RP11-442H21.2	10:72273919-72276036	133,49	22,04	-2,60
398	ENSG00000180211	RP1-278E11.3	6:39958413-39958833	3,50	0,55	-2,66
399	ENSG00000168209	DDIT4	10:72273919-72276036	53,83	7,44	-2,85
400	ENSG00000115008	IL1A	2:112773914-112784590	1,58	0,21	-2,89
401	ENSG00000197614	MFAP5	12:8634366-8662888	3,57	0,47	-2,92
402	ENSG00000131737	KRT34	17:41377649-41382403	18,33	1,99	-3,20
403	ENSG00000228208	C1orf143	1:218510095-218525978	3,80	0,40	-3,24
404	ENSG00000189275	LINC01164	10:131772397-131790199	0,54	0,05	-3,30

405	ENSG00000105550	FGF21	19:48755558-48758330	8,70	0,70	-3,63
406	ENSG00000253522	MIR3142HG	5:160468267-160487426	3,57	0,27	-3,75
407	ENSG00000143217	NECTIN4	1:161070994-161089599	1,19	0,09	-3,79
408	ENSG00000221867	MAGEA3	X:152698751-152702347	7,42	0,41	-4,18
409	ENSG00000164400	CSF2	5:132073789-132076170	25,29	0,98	-4,69
410	ENSG00000130513	GDF15	19:18374730-18389176	26,86	0,82	-5,03
411	ENSG00000170498	KISS1	1:204190340-204196486	18,65	0,34	-5,79
412	ENSG00000169429	CXCL8	4:73740505-73743716	20,28	0,36	-5,83
413	ENSG00000201085	RNU6-173P	4:188121530-188121634	37,58	0,00	-9999
414	ENSG00000204065	TCEAL5	X:103273690-103276872	0,35	0,00	-9999
415	ENSG00000095970	TREM2	6:41158505-41163186	0,52	0,00	-9999
416	ENSG00000226965	AC003088.1	7:110432238-110534754	0,29	0,00	-9999
417	ENSG00000231715	COX6CP2	20:50479766-50479991	2,04	0,00	-9999
418	ENSG00000243635	RP11-281P11.1	3:104502699-104503266	0,47	0,00	-9999
419	ENSG00000278626	RP11-467L19.16	15:20642798-20643448	0,46	0,00	-9999

Tableau 4 : Differentially expressed genes U-CTL vs. U-E6*I.

	Gene id	Gene name	Locus	FPKM U-CTL	FPKM U-E6*I	Log2 fold change
1	ENSG00000224659	GAGE12J	X:49322056-49329384	0,00	5,88	9999
2	ENSG00000231412	CTC-490G23.2	19:43329294-43331430	0,00	0,64	9999
3	ENSG00000068985	PAGE1	X:49687449-49695993	0,42	29,99	6,17
4	ENSG00000158639	PAGE5	X:55220354-55224108	0,47	19,59	5,37
5	ENSG00000123610	TNFAIP6	2:151357591-151380048	0,22	2,86	3,71
6	ENSG00000279998	CH507-210P18.1	21:6789591-6812297	0,12	1,44	3,54
7	ENSG00000162892	IL24	1:206897442-206904139	1,12	10,27	3,20
8	ENSG00000273415	RP11-702B10.2	11:128180426-128183554	0,15	1,38	3,16
9	ENSG00000244128	LINC01322	3:165206959-165655423	0,68	5,92	3,12
10	ENSG00000170454	KRT75	12:52407579-52434525	4,82	40,32	3,06
11	ENSG00000198835	GJC2	1:228149851-228159826	0,78	6,50	3,05
12	ENSG00000066056	TIE1	1:43300992-43323108	0,16	1,30	2,99
13	ENSG00000172399	MYOZ2	4:119135783-119187789	0,07	0,54	2,94
14	ENSG00000175302	ANKRD30BP1	21:13384248-13427773	0,10	0,73	2,84
15	ENSG00000155622	XAGE2	X:52369029-52375680	0,73	4,83	2,72
16	ENSG00000198732	SMOC1	14:69854130-70032366	1,71	10,48	2,61
17	ENSG00000181634	TNFSF15	9:114784634-114806126	0,20	1,15	2,54
18	ENSG00000244155	CYP4F34P	13:18741729-18742718	0,77	4,08	2,40
19	ENSG00000137809	ITGA11	15:68296532-68432162	0,08	0,41	2,38
20	ENSG00000175899	A2M	12:9065176-9116229	0,74	3,85	2,38
21	ENSG00000179772	FOXS1	20:31844300-31845619	1,72	8,75	2,35
22	ENSG00000120937	NPPB	1:11857463-11858931	0,94	4,58	2,29
23	ENSG00000143851	PTPN7	1:202147012-202161588	0,71	3,38	2,26
24	ENSG00000167487	KLHL26	19:18636964-18671714	0,32	1,27	2,01

25	ENSG00000130176	CNN1	19:11538716-11550323	2,61	10,26	1,97
26	ENSG00000128340	RAC2	22:37225260-37244448	4,62	18,10	1,97
27	ENSG00000249550	LINC01234	12:113679458-113773683	1,03	3,87	1,91
28	ENSG00000133048	CHI3L1	1:203178930-203186749	2,69	9,95	1,89
29	ENSG00000109846	CRYAB	11:111908564-111926872	14,47	53,54	1,89
30	ENSG00000235500	SNX19P2	13:18892259-18895230	0,25	0,90	1,84
31	ENSG00000185972	CCIN	9:36169393-36171334	0,52	1,81	1,80
32	ENSG00000153237	CCDC148	2:158166649-158456753	2,59	8,78	1,76
33	ENSG00000162618	ADGRL4	1:78889763-79006718	1,45	4,88	1,75
34	ENSG00000108691	CCL2	17:34255217-34257203	70,91	229,65	1,70
35	ENSG00000149968	MMP3	11:102746967-102843803	7,12	22,86	1,68
36	ENSG00000100311	PDGFB	22:39223358-39244751	3,72	10,85	1,55
37	ENSG00000120051	CFAP58	10:104353763-104455090	2,76	7,94	1,53
38	ENSG00000058335	RASGRF1	15:78959946-79090773	4,28	12,12	1,50
39	ENSG00000169436	COL22A1	8:138588234-138914006	3,52	9,16	1,38
40	ENSG00000275874	PICRAR	21:44999207-45004727	10,48	26,98	1,36
41	ENSG00000102003	SYP	X:49187803-49202454	2,73	6,76	1,31
42	ENSG00000172379	ARNT2	15:80404349-80597937	1,91	4,50	1,23
43	ENSG00000120129	DUSP1	5:172758225-172777774	19,69	42,87	1,12
44	ENSG00000114854	TNNC1	3:52451101-52454070	92,87	193,50	1,06
45	ENSG00000118523	CTGF	6:131948175-132077393	118,54	236,48	1,00
46	ENSG00000114268	PFKFB4	3:48517683-48563773	14,20	6,08	-1,22
47	ENSG00000141753	IGFBP4	17:40443460-40457731	54,23	22,45	-1,27
48	ENSG00000122884	P4HA1	10:73007216-73096974	38,78	14,48	-1,42
49	ENSG00000205922	ONECUT3	19:1752372-1780988	0,39	0,11	-1,89
50	ENSG00000159167	STC1	8:23841914-23854807	5,52	0,92	-2,59
51	ENSG00000137745	MMP13	11:102942994-102955734	0,78	0,11	-2,85
52	ENSG00000083782	EPYC	12:90963678-91005026	0,33	0,00	-9999
53	ENSG00000230109	RP11-275N1.1	10:21340232-21372950	0,45	0,00	-9999

Tableau 5 : PCR and qPCR primers.

Name	Also known as	Gene id	Target	Sequence (5' > 3')
A2M		ENSG00000175899	A2M_ex19_F	TGCCTTACTCTGTGATTCGTG
			A2M_ex20_R	CGCTTGTTCTCTCCACTG
ADGRL4	ELTD1	ENSG00000162618	ADGRL4_ex10_F	CTTTGCTGTAGCCTATTTCTTGC
			ADGRL4_ex11_R	GCCTTCAATGCACATCCATG
CH507-210P18.1		ENSG00000279998	CH507_ex1/2_F	GGATAGAAAATGCAGTTGAAAGATATTTTC
			CH507_ex3_R	CTCTCGCTCCAGGTCCTAG
IGFBP4		ENSG00000141753	IGFBP4_ex2_F	GCACTTCGCCAAAATTCGAG
			IGFBP4_ex3_R	CTCTGTGAAGCGGCCAG
KRT75	K6HF	ENSG00000170454	KRT75_ex2_F	GCCCCTCTTTGATTCTATACC
			KRT75_ex3_R	TGCGCTTGTTAATTTTCATCTTCG
PAGE1	AL5	ENSG00000068985	PAGE1_ex4_F	AAGACTGGGTGTGAGCTTG
			PAGE1_ex5_R	GTGAACCTGTTCTGGCTATC
PFKFB4		ENSG00000114268	PFKFB4_ex8_F	GTGCTATGAGAACTCTACGAG

			PFKFB4_ex9_R	GAGGTAATATACGATGCGGCTC
SMOC1	OAS	ENSG00000198732	SMOC1_ex9_F	TGATGCCCAGTTGTGAGAG
			SMOC1_ex10_R	CAGTAGGCTGGTGATAAACTCC
TNFAIP6	TSG6	ENSG00000123610	TNFAIP6_ex2_F	AATACAAGCTCACCTACGCAG
			TNFAIP6_ex3_R	GGTATCCAACCTCTGCCCTTAG
CCL2	MCP-1	ENSG00000108691	CCL2_ex2_F	TGTCCCAAAGAAGCTGTGATC
			CCL2_ex3_R	ATTCTGGGTTGTGGAGTGAG
TNNC1		ENSG00000114854	TNNC1_ex4_F	TGACCTCTCCGCATGTTTG
			TNNC1_ex5_R	TTGTTCTGTCTCCGTCCTTC
CFAP58	CCDC147	ENSG00000120051	CFAP58_ex8_F	ACAGAAAGGCAATGGACGAG
			CFAP58_ex9_R	TCCTCTGGCTTGTTTCATGG
CNN1		ENSG00000130176	CNN1_ex4_F	AACCATACACAGGTGCAGTC
			CNN1_ex5_R	GATGTTCCGCCCTTCTCTTAG
CRYAB	HSPB5, CRYA2	ENSG00000109846	CRYAB_ex2_F	TGTTGGGAGATGTGATTGAGG
			CRYAB_ex3_R	GGGATGAAGTAATGGTGAGAGG
CTGF	IGFBP8, CCN2	ENSG00000118523	CTGF_ex4_F	ACCAATGACAACGCCTCC
			CTGF_ex5_R	TTGGAGATTTTGGGAGTACGG
MMP3	Stromelysin 1	ENSG00000149968	MMP3_ex4_F	CCAGGGATTAATGGAGATGCC
			MMP3_ex5_R	AGTGTGGCTGAGTGAAAGAG
P4HA1		ENSG00000122884	P4HA1_ex8_F	AGGATGAATGGGACAAGCC
			P4HA1_ex9_R	TTTTCCAGTCTCAGGGTCATG
PICSAR	NLC1-C	ENSG00000275874	PICSAR_F	TCAAGAGGTAAATCCGAGCTC
			PICSAR_R	GATGTCTGAGGAAGAGGCAC
SYP		ENSG00000102003	SYP_ex5_F	AGACAGGGAACACATGCAAG
			SYP_ex6_R	TCTCCTAAACACGAACCCACAG
PDGFB	becaplermin	ENSG00000100311	PDGFB_ex2_F	ATGATCTCCAACGCCTGC
			PDGFB_ex4_R	TCAGCAATGGTCAGGGAAC
RAC2	EN-7	ENSG00000128340	RAC2_ex2_F	CCGTGTTTGACAACTATTGAGC
			RAC2_ex4_R	GGAGAAGCAGATGAGGAAGAC
ARNT2	BHLHe1	ENSG00000172379	ARNT2_ex5_F	GCACCTCAGAAACTCAATGAC
			ARNT2_ex6/7_R	AGCATTTCCACACCTCATCC
MMP13	Collagenase 3	ENSG00000137745	MMP13_ex5_F	GATGACGATGTACAAGGGATCC
			MMP13_ex6_R	ACTGGTAATGGCATCAAGGG
STC1		ENSG00000159167	STC1_ex3_F	AACCCTGAAGCCATCACTG
			STC1_ex4_R	AGGCCCAATTTTCTCCATCAG
SYCP2	SCP-2	ENSG00000196074	SYCP2_ex34_F	GTAACAGTGGAGTTGGAGGTAC
			SYCP2_ex34_R	CTCTTAGTTGGAGTATGTCTGGG
NUP210		ENSG00000132182	NUP210_ex23_F	AAGCTGGACAGAGAATCAACTC
			NUP210_ex25_R	AAAGGATGTTGGACTGAGGC
DICER1		ENSG00000100697	DICER1_F	GTCTCCCAGCATACTTTATCG
			DICER1_R	TTTCCACCCGAAGTCTAAGTTAG
E6			E6_F	TGACTTTGCTTTTCGGGATTT
			E6_R	ACAGCATATGGATTCCCATCTC
E6*I			E6*I_F	CTGCGACGTGAGGTGTATTA
			E6*I_R	TGTCCAGGTGTCTTTGCTTT
E6All			E6_orf_F	ATGCACCAAAGAGAACTGC

		E6_orf_R	TTACAGCTGGGTTTCTCTACGTGT
PXJ40		PXJ40-Vector_F	ATCCACGCTGTTTTGACCTC
		PXJ40-Vector_R	GCACAATAACCAGCACGTTG
RPLP0	ENSG00000089157	RPLP0_F	GAAGGCTGTGGTGCTGATGG
		RPLP0_R	CCGGATATGAGGCAGCAGTT

2 Autres publications

2.1 Etude du métabolisme et de la régulation des ARN viraux

La régulation post-transcriptionnelle des transcrits HPV influe grandement sur l'expression des protéines virales. Mais le devenir de ces transcrits après épissage est peu connu que ce soit au niveau de leur export nucléaire, de leur traduction ou de la régulation de leur stabilité. L'épissage massif des transcrits E6E7 d'HPV16 et des données d'interactomique ont montré que eIF4AIII, un facteur de l'EJC, interagit avec les transcrits E6E7. Les complexes EJC sont déposés en amont des jonctions exon-exon par la machinerie d'épissage et servent de plateforme protéique dans la régulation des processus cités plus haut. La nature polycistronique des transcrits E6E7 épissés avec des ORF et codons STOP multiples en font des cibles potentielles pour le mécanisme de surveillance du NMD. J'ai donc participé à l'étude de l'influence du composant de l'EJC eIF4AIII dans la régulation du devenir des transcrits bicistroniques E6E7 et de l'expression des oncoprotéines E6 et E7.

Nous avons dans un premier temps montré que la déplétion de eIF4AIII entraîne une augmentation de l'expression de l'oncoprotéine E7 mais pas E6. De manière intéressante, l'étude du schéma d'épissage des différents transcrits E6E7 épissés ou non n'a pas révélé de changement d'épissage, indiquant que l'effet inhibiteur d'eIF4AIII sur E7 est médié par un autre mécanisme. Les niveaux de transcrits E6E7 détectés par RNA-FISH suite à une déplétion ou surexpression de eIF4AIII n'ont ensuite pas révélé de changement dans leur localisation subcellulaire. En revanche, ils révèlent que l'expression de eIF4AIII corrèle inversement avec l'expression des transcrits E6E7, confirmant que 4AIII régule négativement les taux de transcrits E6E7. En complément, nous avons montré que 4AIII interagit aussi bien avec les transcrits E6E7 qu'avec les transcrits E6*IE7 et E6*IIE7. La prise en charge de tous les transcrits E6E7 par 4AIII avec la diminution protéique uniquement observé pour E7 et pas pour E6 nous a dirigé vers l'étude du mécanisme de dégradation des ARNm du NMD. L'inhibition chimique du NMD à l'aide des drogues CHX et 5-azacytidine provoque une augmentation du taux des transcrits épissés mais pas du transcrit non-épissé codant E6. L'augmentation du taux de transcrits épissés s'accompagne d'une surexpression de E7.

Prises dans leur ensemble, ces données indiquent que eIF4AIII, en se liant aux transcrits E6*IE7, semble capable de déclencher leur dégradation par le NMD. Le NMD représenterait un nouveau mécanisme de régulation des ARNm E6E7.

The Exon Junction Complex core factor eIF4A3 is a key regulator of HPV16 gene expression (under review)

Koceila Meznad^{a,b,§}, Philippe Paget-Bailly^{a,b,d,§}, Elise Jacquin^{a,c}, Anne Peigney^{a,b,d}, François Aubin^{a,d,e}, Michaël Guittaut^{a,d,g}, Christiane Mougina^{a,b,f}, Jean-Luc Prétet^{a,b,e,f}, Aurélie Baguet^{a,d,g}*

^a Université Bourgogne Franche Comté, France

^b EA3181, UFR Santé, F-25000, Besançon, France.

^c INSERM, U1231, Université Bourgogne Franche Comté, Dijon, France.

^d INSERM, UMR1098, Interactions Hôte-Greffon-Tumeur/Ingénierie Cellulaire et Génique, F-25000, Besançon, France.

^e Centre Hospitalier Régional Universitaire, 3 Bvd Alexandre Fleming, Besançon, France.

^f Centre National de Référence Papillomavirus.

^g DimaCell platform, Université Bourgogne Franche-Comté, F-25000 Besançon, France

§ These authors contributed equally to this work.

*Corresponding author: Aurélie Baguet, INSERM, EFS BFC, UMR1098, Interactions Hôte-Greffon-Tumeur/Ingénierie Cellulaire et Génique, F-25000, Besançon, France.

Tel: +33 3 63 08 22 65

E-mail: aurelie.baguet@univ-fcomte.fr

Keywords: HPV; Exon Junction Complex (EJC); NMD (Nonsense mediated mRNA Decay), Cervical cancer

Abbreviations: human papillomavirus, HPV; Exon Junction Complex, EJC; Nonsense mediated mRNA Decay, NMD

Abstract

High-risk human papillomavirus (hrHPVs), particularly HPV16 and HPV18, are the etiologic factors of ano-genital cancers and some head and neck squamous cell carcinomas. Viral E6 and E7 oncoproteins, controlled at both transcriptional and post-transcriptional levels, drive hrHPVs-induced carcinogenesis. In this study, we investigated the implication of the DEAD-box helicase eIF4A3, an Exon Junction Complex factor, in the regulation of HPV16 gene expression. Our data revealed that the depletion of the factor eIF4A3 upregulated E7 oncoprotein levels. We also showed that the inhibition of the nonsense-mediated RNA decay (NMD) pathway, resulted in the upregulation of E7 at both RNA and protein levels. We therefore proposed that HPV16 transcripts might present different susceptibilities to NMD and that this pathway could play a key role in the levels of expression of these viral oncoproteins during the development of HPV-related cancers.

Introduction

Human papillomaviruses (HPVs) are small non-enveloped viruses with circular double-stranded DNA genomes. More than 200 types of HPVs infect humans ¹ and are classified according to their carcinogenic potential. High-risk HPV types (hrHPVs) (including HPV16 and HPV18) are the etiological agents of cervical and anal carcinomas and frequently found in head and neck squamous cell carcinomas (HNSCC), and low-risk HPV types (lrHPVs) (including HPV6 and HPV11) are associated with benign lesions such as genital warts ^{2,3}. In most cases, HPV is cleared by the immune system in the first year after infection. However, persistence of hrHPV infection is associated with an increased risk of developing high-grade squamous intra-epithelial lesions (HSIL), corresponding to the histological grade of Cervical Intraepithelial Neoplasia 2/3 (CIN). In this case, HPV infection is most often abortive and associated with a deregulation of viral gene expression and a high risk of progression to cervical cancer ⁴⁻⁷.

The HPV16 genome comprises a long control region (LCR) and 8 open reading frames (ORFs) encoding early (E1, E2, E4, E5, E6 and E7) and late (L1 and L2) viral proteins. The early promoter p97 located in the LCR drives the expression of genes coding early proteins whereas late proteins are produced from transcripts regulated by the late promoter p670. hrHPV-driven carcinogenesis is most often related to viral genome integration into the host genome, leading to E6 and E7 oncoproteins overexpression^{5,8,9,10}. Both HPV16 E6 and E7 proteins present pleiotropic effects by interacting with multiple host cell proteins, such as p53 and pRb tumor suppressors, respectively¹¹⁻¹⁴. Continuous expression of HPV16 E6 and E7 proteins has been demonstrated to be essential for the initiation and the maintenance of cancer¹⁵.

In addition to transcriptional regulation, HPV16 RNA processing largely contributes to the regulation of viral protein expression. For instance, HPVs take advantage of alternative splicing regulated by cellular splicing factors to allow for the production of multiple viral protein-encoding mRNAs^{16,17}. Indeed, p97-derived HPV16 transcripts are submitted to alternative splicing which generates more than a dozen different polycistronic mRNAs encoding early proteins^{16,18}. At least four cryptic splicing sites have been identified in HPV16 E6 and E7 ORFs: one splicing donor site, SD226, and three splicing acceptor sites, SA409, SA526 and SA742, leading to the putative expression of truncated proteins, respectively named E6*I, E6*II and E6^E7¹⁹⁻²¹.

The eukaryotic translation Initiation Factor 4A3 (eIF4A3, also known as DDX48) is a DEAD-box RNA helicase implicated in the formation of the Exon Junction Complex (EJC), unlike its homologous eIF4AI and eIF4AII which are required for translation initiation. The EJC is a multiprotein complex assembled onto the pre-messenger RNA (pre-mRNA) during splicing. The complex is recruited 20 to 24 nucleotides upstream of the exon-exon junction, in a sequence-independent manner²². The EJC core is composed of four proteins, named eIF4A3, Y14 (RBM8A), MAGOH and MLN51 (CASC3), which acts as a binding platform to recruit other peripheral EJC factors²³⁻²⁷. Consequently, the composition of the EJC evolves all along the nuclear and cytoplasmic mRNA life^{28,29}.

Ultimately, EJC factors are removed by the translation machinery during the first round of translation^{30,31}. The EJC greatly influences mRNA fate by affecting RNA processing: *i)* the EJC regulates the splicing process^{25,32-34} and mRNA transport^{35,36}, *ii)* it promotes nuclear export³⁷⁻³⁹ and translation⁴⁰⁻⁴² and, *iii)* it plays a crucial role in the mRNA surveillance pathway⁴³. Indeed, it has been shown that the EJC allows for the recognition of Premature Termination Codons (PTCs) on transcripts and their targeting for degradation by the nonsense-mediated mRNA decay (NMD). So, even if it is clear that EJC factors are key players in post-transcriptional regulation of gene expression, the role of eIF4A3 in HPV RNA processing has never been studied. In this study, we therefore investigated, for the first time, the implication of eIF4A3 in the regulation of HPV16 E6 and E7 oncoprotein expression.

Materials and Methods

Overall survival analysis based on TCGA data

Correlation analysis between eIF4A3 expression and overall survival of patients presenting a cervical cancer were based on TCGA Research Network (<http://www.cancer.gov/tcga>). The graph and log-rank analysis were generated from the Human Protein Atlas (<http://www.protein-atlas.org>)⁴⁴. The 291 cervical cancer patients were stratified by the median of eIF4A3 mRNA expression.

Cell culture

Human cervical cancer-derived cell lines, C-33 A (HPV-negative, ATCC® HTB-31), SiHa (HPV16-positive, ATCC® HTB-35) and Ca Ski (HPV16-positive, ATCC® CRL-1550) were obtained from the American Type Culture Collection. C-33 A and SiHa cells were cultured in DMEM medium and Ca Ski in RPMI-1640 medium (Lonza, Verviers, Belgium) supplemented with 10 % Fetal bovine serum (Eurobio, Courtaboeuf, France). Cells were incubated at 37°C in a humidified atmosphere with 5 % of CO₂. All cell lines were routinely tested for mycoplasma contamination.

siRNAs, expression vectors and chemical reagents

Negative control (SR-CL000-005) and eIF4A3 (target sequence: 5'-AGACATGACTAAAGTGGAA-3') targeting siRNA duplexes were purchased from Eurogentec (Liege, Belgium). The peGFP-eIF4A3, peGFP-MLN51 and peGFP-Dcp1a vectors were a kind gift from Dr Catherine Tomasetto⁴⁵ and Dr Fabienne Mauxion⁴⁶. The translation inhibitor cycloheximide (CHX) was purchased from Sigma and used at the concentration of 100 µg/mL. The nucleotide analogues 5-aza-2'-deoxycytidine and 5-azacytidine were purchased from Epigentek and StressMarq Biosciences, respectively, and used at a concentration range from 0.5 to 10 µM.

Antibodies

Mouse antibodies anti-E7 (1/1000), anti-E6 2E-3F8 (1/1000) and anti-βactin AC-15 (1/10000) were purchased from Santa Cruz, Euromedex and Sigma, respectively. Mouse and rabbit anti-eIF4A3 antibodies were generously provided by Catherine Tomasetto (I.G.B.M.C., Illkirch, Strasbourg). HRP-conjugated goat anti-mouse and anti-rabbit secondary antibodies were purchased from BD Pharmingen. Alexa488-labelled goat anti-rabbit and Cy3-labelled donkey anti-mouse secondary antibodies used for immunofluorescence were purchased from Jackson ImmunoResearch.

Transfection

Cells were plated in 6-well plates one day before transfection so that they were 60-75% confluent the day of the transfection. siRNA transfections were performed using the Lipofectamine 2000 reagent (Invitrogen, Carlsbad, USA) following the supplier's recommendations. Cells were then cultured for 48 h in complete medium before harvesting. Plasmid transfections were performed using the JetPEI reagent (Polyplus-transfection, Illkirch, France) according to the manufacturer's instructions. Cells were incubated 24 h in complete medium before harvesting.

RNA-FISH

Cells were cultured on coverslips in 24-well plates. Plasmid or siRNA transfections were performed as described above. After transfection, cells were fixed in 4 % paraformaldehyde for 30 min at room temperature. RNA-FISH assays were performed using the RNAscope® Fluorescent Multiplex Assay (Advanced Cell Diagnostics Bio, Milan, Italy) following the supplier's recommendations. Briefly, cells were dehydrated with increasing concentrations of EtOH (50 %, 70 % and 100 % for 2 min) and

rehydrated with decreasing concentrations of EtOH (70 % and 100 % for 2 min). Cells were then incubated with Protease III for 10 min at room temperature and subjected to RNAscope probe hybridization and counterstaining with DAPI. Slides were mounted in Vectashield (Polysciences Inc.) and subjected to confocal microscopy (Zeiss LSM 800 AiryScan).

RT-PCR and RT-qPCR

Total RNAs were extracted using the Ribozol reagent (Amresco) and 500 ng total RNAs were retro-transcribed into cDNA using the Maxima first strand cDNA synthesis kit (ThermoScientific) according to the manufacturer's instructions. For PCR, cDNAs were amplified using the DreamTaq polymerase (Thermo Scientific) according to the manufacturer's instructions. PCR products were then analyzed on a 1 or 2 % (w/v) agarose gel prepared in 1X Tris-Borate-EDTA buffer. Quantitative PCR was performed using the Power SYBR Green PCR mix (Life Technologies) according to the manufacturer's instructions and analyzed in a StepOnePlus Real-Time PCR system (Applied Biosystems). Primer sequences are listed in the supplementary Table 1. 18S rRNA or β 2-microglobulin expression levels were used as internal normalization standards⁴⁷. The $2^{-\Delta\Delta Ct}$ method was used for relative mRNA quantification. Unilateral student t tests were used to analyze statistical significance of RT-qPCR experiments (*: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$).

Protein Extraction and Western blotting

Cell lysis was performed using RIPA buffer (50 mM Tris-HCl pH7.4, 150 mM NaCl, 1 % NP40, 0.5 % Na deoxycholate, 1 mM EDTA) supplemented with 30 μ g/mL of anti-protease mixture (Roche Diagnostics, Mannheim, Germany) followed by 5 sec sonication. Following centrifugation at 10 000g for 10 min at 4°C, protein concentrations were determined using the Bio-Rad Protein Assay (Bio-Rad, Munich, Germany) according to the manufacturer's instructions. Proteins were then separated by SDS-PAGE and transferred onto AmershamTMHybondTM-PVDF membranes (GE Healthcare, Munich, Germany). Membranes were blocked overnight at 4°C with 5 % nonfat milk, probed during 1 h with primary antibodies (see above) and then with HRP-conjugated secondary antibodies. Chemiluminescent signals were detected using the PierceECL 2 Western Blotting Substrate (Thermo Scientific). The signals were then analyzed and

quantified using the ChemiDoc XRS+ system (Biorad), and the Image Lab software (5.1, Biorad).

RNA-Immunoprecipitation

Cells were lysed in NET-2 buffer (50 mM Tris-HCl pH7.4, 150 mM NaCl, 10 mM MgOAc, 0.05 % (v/v) NP-40, 1 mM DTT, protease inhibitor (Roche Diagnostics), 80 U/mL RNase OUT (Invitrogen)), sonicated 5 times for 1 second, and centrifuged at 16,100 g for 10 min at 4°C. Then, the samples were washed with NET-2 buffer/300 mM NaCl followed by a centrifugation at 500 g for 5 min at 4°C. Then, NET-2 buffer/300 mM NaCl and 1 mg/ml heparin was added and incubated under agitation for 10 min at 4°C, followed by a centrifugation step at 500 g for 5 min at 4°C. After 8 washing steps using NET-2 buffer and a centrifugation step at 500 g for 5 min at 4°C, protein and antibodies complexes were eluted in 50 µl Elution buffer (0.1 M Tris-HCl pH 6.8, 20 % (v/v) glycerol, 4 % (w/v) SDS, 12 % (v/v) βMercapto-EtOH) followed by a centrifugation step at 500 g for 5 min at 4°C. Eluates were finally split into 2 fractions: one fraction was used to perform western blotting and the other was subjected to RNA extraction followed by RT-qPCR analysis.

Results

High levels of eIF4A3 is associated with a good prognosis in cervical cancers

First, we evaluated the prognostic value of eIF4A3 expression in cervical cancers using The Cancer Genome Atlas (TCGA) database. Amongst the 291 samples available, the 5-year survival of patients presenting high levels of eIF4A3 (above median of eIF4A3 level) was 76 %, compared to 59 % for patients presenting low levels of eIF4A3 (below median of eIF4A3 level) ($p < 0.05$) (**Figure 1**). These data demonstrated that a high expression of eIF4A3 was an indicator of good prognosis for cervical cancer patients.

The expression of the HPV16 E7 oncoprotein was increased in eIF4A3-depleted cells

To determine whether eIF4A3 could regulate HPV16 gene expression, we performed siRNA-mediated depletion of eIF4A3 in the HPV16-positive cervical carcinoma SiHa cell line. We observed that the depletion of eIF4A3 had no effect on E6 protein levels (**Figure 2A, left panel**). Interestingly, the depletion of eIF4A3 led to an increase of E7 protein levels in eIF4A3-depleted cells compared to control samples (**Figure 2A, right panel**). Several alternatively spliced transcripts produced E6 and E7 oncoproteins following the use of 1 donor splice site (SD226) and 3 acceptor splice sites (SA409, SA526 and SA742) (**Figure 2B**). The recognition by the spliceosome of the combinations of SD226/SA409 or SD226/SA526 splicing sites produced two different spliced transcripts encoding the E7 protein and potentially two truncated E6 proteins, E6*I and E6*II, respectively. A fusion protein called E6^{E7} is also potentially encoded by the spliced transcript SD226/SA742. Finally, the full length E6 protein could exclusively be encoded by the unspliced transcript. In SiHa cell line, spliced transcripts potentially encoding both E7 and E6-truncated proteins were mainly expressed (**Figure S1**). Next, we asked whether the deregulation of E7 expression mediated by eIF4A3 could lead to a modulation of the alternative splicing pattern of E6/E7 transcripts. To answer this question, we used specific primer sets developed for RT-PCR (F6 and R7 primers) or RT-qPCR (E6All, E6, E6*I and E7 primers) to simultaneously, or individually, detect HPV16 E6/E7 alternative transcripts (see **Figure 2B**). The representative RT-PCR results, obtained using the primers (F6-R7) targeting early HPV16 transcripts, are presented in **Figure 2C**. The KPNA1 transcript was used as a positive control since it has been previously reported that the depletion of eIF4A3 led to an increase in exon skipping of the constitutive exon 11 of KPNA1³⁴. As expected, the eIF4A3 knockdown indeed led to an increase of the exclusion of the exon 11 of KPNA1. However, the depletion of eIF4A3 did not alter the splicing profile of HPV16 early transcripts. Also, we performed RT-qPCR using E6All, E6, E6*I and E7 primers and observed a low but statistically significant elevation of HPV16 early transcripts (1.6-fold increase) in eIF4A3-depleted cells (**Figure 2D**). Finally, both E6/E6All ratios calculated in siRNA-CTL and siRNA-eIF4A3 treated cells were around 1,

confirming that the depletion of eIF4A3 did not affect the splicing pattern of HPV16 early transcripts (**Figure 2E**). Altogether, these data indicated that eIF4A3 induced an increase of E7 oncoprotein levels and that this increase was not linked to a differential alternative splicing.

HPV16 RNA localization in eIF4A3-depleted cells

To test whether eIF4A3 could affect the subcellular localization of HPV16 E6/E7 transcripts, we analyzed, using RNA-FISH, their subcellular localization in SiHa cells depleted for eIF4A3. We used two probe sets designed to recognize the E6 sequence (from nucleotide 2 to nucleotide 476) and the E7 sequence (from nucleotide 563 to nucleotide 846) (**Figure 3A**). In both conditions (siRNA-CTL and siRNA-eIF4A3), HPV16 E6 and E7 transcripts were detected as punctate dots in the cytoplasm, indicating that the eIF4A3 knockdown had no effect, neither on the nuclear export, nor on the subcellular localization of E6 or E7 mRNAs (**Figure 3B**). However, we observed a significant increase in punctate dots corresponding to E6 and E7 mRNAs when SiHa cells were depleted for eIF4A3. Indeed, about 10 dots per cell were counted in eIF4A3-depleted cells compared to 4 dots per cell in control cells (**Figure 3C**). These findings indicated that eIF4A3 did not influence the localization of E6 and E7 mRNAs but regulated their cellular levels. To confirm these observations, a plasmid coding the eIF4A3 protein fused to an N-terminal EYFP protein was transiently transfected in SiHa cells and E6/E7 mRNA levels were analyzed by RNA-FISH. Interestingly, less cytoplasmic dots corresponding to E6 and E7 RNAs were detected in eIF4A3-YFP transfected cells compared to non-transfected cells (**Figures 3D and 3E**). Finally, to examine whether HPV16 transcripts were recruited to cytoplasmic RNA granules such as stress granules or P-bodies, E6 RNA-FISH in SiHa cells transfected with either the peGFP-Dcp1 vector (**Figure S3**) or the peGFP-MLN51 vector (**Figure S4**) were performed. No colocalization between GFP-Dcp1 or GFP-MLN51 proteins and E6 RNAs were observed, indicating that E6 RNAs were not recruited in these RNA granules.

Altogether, these data confirmed that eIF4A3 act as a negative regulator of HPV16 mRNA levels without altering their intracellular localization.

The eIF4A3 protein associated with ribonucleoparticles containing HPV16 early transcripts

Next, we analyzed the interaction between the eIF4A3 protein and HPV16 mRNAs by performing RNA-IPs. Total RNA extracts from SiHa cells were incubated with anti-eIF4A3 (4A3-IP) or control (IgG-IP) antibodies, and the composition of the immunoprecipitated complexes were analyzed by western blotting and RT-PCR (**Figures 4A and 4B, respectively**). As expected, the intronless SF3B5 mRNAs (negative control) were not immunoprecipitated, while the KPNA1 mRNAs (positive control) were detected to bind to eIF4A3 (**Figure 4B**). Taken together, these data validated our protocol of eIF4A3-targeted RNA-IP (4A3-IP). Interestingly, we observed a significant enrichment of E6/E7, E6*I/E7 and E6*II/E7 transcripts in the eIF4A3-IP compared to the control IgG-IP. Our results therefore confirmed that eIF4A3 proteins were indeed present in messenger ribonucleoprotein particles (mRNPs) containing HPV16 transcripts.

Nonsense-mediated mRNA decay (NMD) inhibitors, cycloheximide and 5-Azacytidine, differently affected HPV16 E6 and E7 mRNA and protein levels

One possibility to explain how eIF4A3 depletion might lead to the upregulation of HPV16 transcripts levels is its role in the NMD pathway. The NMD is responsible for the elimination of abnormal transcripts but also natural mRNAs which harbor NMD-eliciting features such as PTCs, multiple ORFs or long 3'-UTRs. The presence of multiple ORFs in HPV16 mRNAs and their multiple alternative splicing possibilities could therefore represent NMD-degradation signals. In order to determine whether the NMD pathway indeed regulated the levels of HPV16 transcripts, we explored the impact of drug-induced NMD inhibition. Because NMD requires translation, SiHa cells were first treated

with cycloheximide (CHX), a protein synthesis inhibitor, to inhibit NMD. As expected, the levels of SC35 mRNAs, a well-described physiological NMD substrate, were increased in CHX-treated cells compared to non-treated cells (**Figure 5A**). Interestingly, the CHX treatment had no effect on unspliced E6 RNA levels but E6*I and E7 RNA levels were upregulated in CHX-treated cells compared to cells treated with DMSO alone (**Figure 5A**). Using RNA FISH experiments designed to detect HPV16 or c-JUN mRNAs (used as positive control) ⁴⁸, we detected, as expected, an increase in the number of fluorescent signals corresponding to the positive control c-JUN mRNAs in CHX-treated SiHa cells compared to DMSO-treated cells. Regarding HPV16 mRNAs, we observed the presence of few cytoplasmic dots corresponding to HPV16 mRNAs (about 2 to 4 dots per cell) in DMSO-treated cells (**Figure 5B**) while, in CHX-treated cells, the number of dots corresponding to E6/E7 mRNAs was significantly increased (about 5 to 7 dots per cell) (**Figure 5C**).

Recently, it has been reported that the inhibition of NMD in cells could also be obtained following a treatment using a nucleotide analog called 5-azacytidine (5-azaC) ⁴⁹. Endogenous E6 and E7 mRNA and protein levels were thus monitored in SiHa cells treated with the active compound 5-azaC or the control compound 5-aza-2'-deoxycytidine (5-azadC), used as a negative control. Two physiological NMD substrates, the SC35 and asparagine synthetase (ASNS) mRNAs, were used as positive controls of NMD targets and we indeed showed an increase of their respective levels in the 5-azaC-treated cells (from a 1.4 to a 20-fold induction) (**Figure 6A, upper panel**) compared to the control 5-azadC-treated cells (**Figure 6A, lower panel**). No significant differences in the HPV16 mRNA levels were observed in cells treated with the 5-azadC compound (**Figure 6A, lower panel**). But, when cells were treated with 5-azaC, we detected a dose-dependent increase of E6*I and E7 mRNA levels while the E6 mRNA levels were not affected (**Figure 6A upper panel**). Moreover, we observed that the 5-azaC treatment significantly upregulated the HPV16 E7 protein levels (2 fold change) while we did not detect any significant changes in the levels of the E6 protein (**Figures 6B and**

6C). As previously described, the demethylating agent 5-azacytosine (5-azaC) induces repression of E6 expression at both mRNA and protein levels⁵⁰. Altogether, these data suggested that the alternatively spliced HPV16 transcripts encoding the E7 protein could be targeted by the NMD pathway to induce its degradation whereas the unspliced HPV16 transcript encoding the E6 protein might escape NMD-induced degradation.

Discussion

The expression of the viral E6 and E7 oncoproteins is critical for the development and the maintenance of HPV-induced tumors. HPV16 transcripts generated from the p97 promoter allow for the production of the E6 and E7 proteins causing the inactivation of the tumor suppressor proteins p53 and pRb, respectively. High E6 and E7 levels were linked to tumor growth and poor patient survival^{51,52}. In contrast, low E6 and E7 levels restored antitumor mechanisms and induced tumor cell death^{53,54}. Also, high levels of HPV16 spliced transcripts encoding the protein E6*1 increased the levels of reactive oxygen species leading to tumor growth inhibition^{55,56}. The expression of HPV16 transcripts is under the control of both viral and host cell factors. The main step of their regulation occurs at the post-transcriptional level through alternative splicing, leading to viral transcriptome and proteome diversification¹⁶. Nevertheless, HPV16 mRNA processing and the influence of RNA binding proteins in the control of viral RNA expression during infection and in the development of HPV-related cancer remain elusive. According to the TCGA database, we showed that the levels of the EJC factor eIF4A3 could be considered as an indicator of good prognosis in cervical cancers. Therefore, we addressed the role of this DEAD-box RNA helicase in the regulation of the expression of HPV16 early transcripts encoding E6 and E7 oncoproteins. For the first time, in this study, we observed an upregulation of HPV16 E7 mRNA and protein levels following eIF4A3 depletion and therefore, that eIF4A3 acted as a negative regulator of E7 expression.

Our data suggested that eIF4A3 might become an interesting target to regulate viral oncoprotein levels in HPV16-infected cancer cells. Previous studies showed that the inhibition of eIF4A3 could induce cell death by modifying the ratio between pro- and anti-apoptotic proteins. Indeed, the inhibition of eIF4A3 deregulated the splicing of several cellular transcripts including those encoding apoptosis-related factors, such as Bcl-X^{25,33}. However, our current data indicated that the depletion of eIF4A3 did not alter the E6/E7 splicing pattern in the SiHa cell line leading us to the conclusion that the increased E7 levels, observed upon eIF4A3 depletion in the HPV16-positive SiHa cell line, were not a consequence of the alteration of their RNA splicing pattern.

During the course of RNA-FISH experiments, we detected more HPV16 transcripts in the cytoplasm of eIF4A3-depleted cells compared to the control ones. However, in our study, we did not observe any subcellular relocalization of HPV16 mRNAs in eIF4A3-depleted cells. In a previous publication, we observed that eIF4A3 accumulated in the nucleus, and more particularly in nuclear speckles, but, under stress conditions, eIF4A3 also partially relocalized in cytoplasmic stress granules (SGs) (Daguenet et al., 2012 and data not shown). SGs are cytoplasmic membrane-less aggregates of untranslated messenger ribonucleoproteins (mRNPs) implicated in cell survival under stress conditions. More recently, it has also been shown that the inhibition of eIF4A3 might affect the cellular stress response pathway by decreasing the formation of SGs⁵⁷. In our data, we did not observe a colocalization between HPV16 transcripts and SGs or P-bodies. It may therefore be interesting to analyze the subcellular localization of HPV16 transcripts during stress conditions to determine whether these transcripts could be recruited, or not, into SGs.

Then, we confirmed the presence of eIF4A3 in HPV16 E6/E7 mRNPs highlighting a direct impact of this protein on viral mRNA biology. These results are in agreement with those of Martínez-Salazar *et al.*, 2014, whom identified eIF4A3 and EJC factors as binding partners of the HPV16 transcripts. Amongst the various EJC-related functions of eIF4A3, its involvement in the NMD pathway might explain the differences in HPV16 transcript

levels observed following eIF4A3 depletion. HPV16 transcripts contain multiple ORFs, a feature that has been described to target mRNA for degradation by the NMD. Indeed, it has been suggested that SRSF2, a splicing factor, may regulate E6 and E7 oncoprotein expression by protecting HPV16 transcripts from NMD-induced mRNA degradation ⁵⁹. Moreover, several previous studies have already described viral mRNA degradation induced by the NMD ⁶⁰⁻⁶⁵. To assess the impact of the NMD on HPV16 E6/E7 mRNA and protein levels, we used cycloheximide and 5-azaC to inhibit the NMD pathway. We observed that a CHX, or a 5-azaC treatment, did increase E6*I/E7 RNA levels, but not E6 RNA levels. Interestingly, these results suggested that the spliced E6*I/E7 transcripts could be targeted by the NMD, whereas the unspliced E6/E7 mRNAs could escape NMD-targeted mRNA degradation. These differences could be explained by the distance separating the two ORFs. Indeed, the E6 and E7 ORFs are separated by 2 nucleotides whereas 153 nucleotides are present between the E6*I and E7 ORFs. It has been recently shown that translation re-initiation is one mechanism by which PTC-containing transcripts can escape NMD ⁶⁶. So, an efficient translation re-initiation process between the E6 and E7 ORFs could protect these transcripts from degradation. Other mechanisms have also been described to protect viral mRNA from NMD. This is for instance the case for the Rous Sarcoma Virus RNAs which possess a *cis*-acting element protecting them from NMD degradation ⁶³. Further studies will therefore be needed to unravel the mechanism by which eIF4A3 can modulate E6/E7 mRNA stability. Such work would also allow us to ameliorate our knowledge of HPV16 mRNA regulation but also the functions of EJC factors, and especially those of eIF4A3, in this regulation.

Altogether, we showed, for the first time in this study, that eIF4A3 is a key regulator of HPV16 gene expression, likely through its implication in the NMD pathway. Given the good prognosis linked to high levels of eIF4A3 in cervical cancers, it would be interesting in the near future to quantify E6 and E7 expression levels in patients and determine whether we can find a correlation with the levels of expression of eIF4A3. If this correlation can be detected, we could then hypothesize that a loss of eIF4A3

function, along the evolution of HPV-induced carcinogenesis, could favor an overexpression of these oncoproteins and cervical cancer development.

According to our results, we can propose a model in which HPV16 transcripts can differentially elicit the NMD pathway. Indeed, high eIF4A3 levels and a high NMD activity could lead to low levels of E6*1/E7 mRNAs causing a moderate E7 protein production. On the other hand, low eIF4A3 levels and a decrease of NMD efficiency would lead to an increase in E6*1/E7 mRNA levels inducing a high expression of the E7 protein (**Figure 7**). Our data are in agreement with a recent study which showed that the HPV18 E7 protein can induce the expression of ABC transporters, such like P-gp, MRP1, MRP2 and BCRP, through a decrease of NMD efficiency in HPV-positive oropharyngeal cancer cells ⁶⁷. In addition, E6 and/or E7 protein have been described to affect NMD efficiency in order to favor E7 production and/or other viral proteins. Nevertheless, further studies will be required to determine whether the modulation of NMD efficiency during HPV infection could contribute to HPV-related carcinogenesis.

In conclusion, our data highlight the importance of RNA-binding proteins on HPV16 gene expression during the evolution of HPV-positive tumors and the question about their importance in the promotion of HPV-induced carcinogenesis.

Acknowledgements

This work was supported by fundings from Ligue Contre le Cancer (Comité du Doubs and CCIR-GE), Cancéropôle Est (CE), Université de Franche-Comté (UFC) and Région de Franche-Comté. We thank Fabienne Mauxion for kindly providing the peGFP-Dcp1a and Catherine Tomasetto for the peYFP-eIF4A3 and peYFP-MLN51 vectors and anti-eIF4A3 antibodies, Julie Durel for their technical assistance and the Dimacell Imaging Facility. K.M. and P.P. were financed by fellowships from the French Ministry of Research and Technology.

Author contributions statement

A.B. conceived the project. K.M., P.P.B. and A.P. performed the biochemical assays and the functional assays, K.M., P.P.B., E.J., F.A., C.M., J.L.P. and A.B. analyzed the data. M.G. and A.B. wrote the paper.

Data availability statement

The authors confirm that all data underlying the findings are fully available without restriction.

Competing interests statement

The authors declare that they have no competing interests.

References

1. Schiffman, M. et al. Carcinogenic human papillomavirus infection. *Nat. Rev. Dis. Primer* 2, 16086 (2016).
2. De Vuyst, H., Clifford, G. M., Nascimento, M. C., Madeleine, M. M. & Franceschi, S. Prevalence and type distribution of human papillomavirus in carcinoma and intraepithelial neoplasia of the vulva, vagina and anus: A meta-analysis. *Int. J. Cancer* 124, 1626–1636 (2009).
3. Guan, P. et al. Human papillomavirus types in 115,789 HPV-positive women: A meta-analysis from cervical infection to cancer. *Int. J. Cancer* 131, 2349–2359 (2012).
4. Koutsky, L. A. et al. A Cohort Study of the Risk of Cervical Intraepithelial Neoplasia Grade 2 or 3 in Relation to Papillomavirus Infection. *N. Engl. J. Med.* 327, 1272–1278 (1992).
5. Shulzhenko, N., Lyng, H., Sanson, G. F. & Morgun, A. Ménage à trois: an evolutionary interplay between human papillomavirus, a tumor, and a woman. *Trends Microbiol.* 22, 345–353 (2014).
6. Wallin, K.-L. et al. Type-Specific Persistence of Human Papillomavirus DNA before the Development of Invasive Cervical Cancer. *N. Engl. J. Med.* 341, 1633–1638 (1999).

7. Ylitalo, N. et al. A prospective study showing long-term infection with human papillomavirus 16 before the development of cervical carcinoma in situ. *Cancer Res.* 60, 6027–6032 (2000).
8. Klaes, R. et al. Detection of high-risk cervical intraepithelial neoplasia and cervical cancer by amplification of transcripts derived from integrated papillomavirus oncogenes. *Cancer Res.* 59, 6132–6136 (1999).
9. Sathish, N. et al. Human papillomavirus 16 E6/E7 transcript and E2 gene status in patients with cervical neoplasia. *Mol. Diagn. J. Devoted Underst. Hum. Dis. Clin. Appl. Mol. Biol.* 8, 57–64 (2004).
10. McBride, A. A. & Warburton, A. The role of integration in oncogenic progression of HPV-associated cancers. *PLOS Pathog.* 13, e1006211 (2017).
11. Chellappan, S. et al. Adenovirus E1A, simian virus 40 tumor antigen, and human papillomavirus E7 protein share the capacity to disrupt the interaction between transcription factor E2F and the retinoblastoma gene product. *Proc. Natl. Acad. Sci.* 89, 4549–4553 (1992).
12. Roman, A. & Munger, K. The papillomavirus E7 proteins. *Virology* 445, 138–168 (2013).
13. Scheffner, M., Huibregtse, J. M., Vierstra, R. D. & Howley, P. M. The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. *Cell* 75, 495–505 (1993).
14. Vande Pol, S. B. & Klingelutz, A. J. Papillomavirus E6 oncoproteins. *Virology* 445, 115–137 (2013).
15. Münger, K. The role of human papillomaviruses in human cancers. *Front. Biosci. J. Virtual Libr.* 7, d641-649 (2002).
16. Graham, S. V. & Faizo, A. A. A. Control of human papillomavirus gene expression by alternative splicing. *Virus Res.* 231, 83–95 (2017).
17. Schwartz, S. Papillomavirus transcripts and posttranscriptional regulation. *Virology* 445, 187–196 (2013).
18. Chen, J. et al. Mapping of HPV transcripts in four human cervical lesions using RNAseq suggests quantitative rearrangements during carcinogenic progression. *Virology* 462–463, 14–24 (2014).

19. Schneider-Gädicke, A. & Schwarz, E. Different human cervical carcinoma cell lines show similar transcription patterns of human papillomavirus type 18 early genes. *EMBO J.* 5, 2285–2292 (1986).
20. Schneider-Gädicke, A. et al. Identification of the human papillomavirus type 18 E6 and E6 proteins in nuclear protein fractions from human cervical carcinoma cells grown in the nude mouse or in vitro. *Cancer Res.* 48, 2969–2974 (1988).
21. Zheng, Z.-M., Tao, M., Yamanegi, K., Bodaghi, S. & Xiao, W. Splicing of a Cap-proximal Human Papillomavirus 16 E6E7 Intron Promotes E7 Expression, but can be Restrained by Distance of the Intron from its RNA 5' Cap. *J. Mol. Biol.* 337, 1091–1108 (2004).
22. Le Hir, H., Moore, M. J. & Maquat, L. E. Pre-mRNA splicing alters mRNP composition: evidence for stable association of proteins at exon-exon junctions. *Genes Dev.* 14, 1098–1108 (2000).
23. Chan, C. C. eIF4A3 is a novel component of the exon junction complex. *RNA* 10, 200–209 (2004).
24. Degot, S. et al. Association of the Breast Cancer Protein MLN51 with the Exon Junction Complex via Its Speckle Localizer and RNA Binding Module. *J. Biol. Chem.* 279, 33702–33715 (2004).
25. Fukumura, K. et al. The Exon Junction Complex Controls the Efficient and Faithful Splicing of a Subset of Transcripts Involved in Mitotic Cell-Cycle Progression. *Int. J. Mol. Sci.* 17, 1153 (2016).
26. Hir, H. L., Saulière, J. & Wang, Z. The exon junction complex as a node of post-transcriptional networks. *Nat. Rev. Mol. Cell Biol.* 17, 41–54 (2016).
27. Kataoka, N. Magoh, a human homolog of *Drosophila mago nashi* protein, is a component of the splicing-dependent exon-exon junction complex. *EMBO J.* 20, 6424–6433 (2001).
28. Bono, F. & Gehring, N. H. Assembly, disassembly and recycling: The dynamics of exon junction complexes. *RNA Biol.* 8, 24–29 (2011).
29. Zhang, Z. & Krainer, A. R. Splicing remodels messenger ribonucleoprotein architecture via eIF4A3-dependent and -independent recruitment of exon junction complex components. *Proc. Natl. Acad. Sci.* 104, 11574–11579 (2007).
30. Dostie, J. & Dreyfuss, G. Translation Is Required to Remove Y14 from mRNAs in the Cytoplasm. *Curr. Biol.* 12, 1060–1067 (2002).

31. Lejeune, F. The exon junction complex is detected on CBP80-bound but not eIF4E-bound mRNA in mammalian cells: dynamics of mRNP remodeling. *EMBO J.* 21, 3536–3545 (2002).
32. Ashton-Beaucage, D. et al. The Exon Junction Complex Controls the Splicing of *mapk* and Other Long Intron-Containing Transcripts in *Drosophila*. *Cell* 143, 251–262 (2010).
33. Michelle, L. et al. Proteins Associated with the Exon Junction Complex Also Control the Alternative Splicing of Apoptotic Regulators. *Mol. Cell. Biol.* 32, 954–967 (2012).
34. Wang, Z., Murigneux, V. & Le Hir, H. Transcriptome-wide modulation of splicing by the exon junction complex. *Genome Biol.* 15, 551 (2014).
35. Giorgi, C. & Moore, M. J. The nuclear nurture and cytoplasmic nature of localized mRNPs. *Semin. Cell Dev. Biol.* 18, 186–193 (2007).
36. Martin, K. C. & Ephrussi, A. mRNA Localization: Gene Expression in the Spatial Dimension. *Cell* 136, 719–730 (2009).
37. Le Hir, H. The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *EMBO J.* 20, 4987–4997 (2001).
38. Valencia, P., Dias, A. P. & Reed, R. Splicing promotes rapid and efficient mRNA export in mammalian cells. *Proc. Natl. Acad. Sci.* 105, 3386–3391 (2008).
39. Zhou, Z. et al. The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. *Nature* 407, 401–405 (2000).
40. Chazal, P.-E. et al. EJC core component MLN51 interacts with eIF3 and activates translation. *Proc. Natl. Acad. Sci.* 110, 5903–5908 (2013).
41. Choe, J. et al. eIF4AIII enhances translation of nuclear cap-binding complex-bound mRNAs by promoting disruption of secondary structures in 5'UTR. *Proc. Natl. Acad. Sci.* 111, E4577–E4586 (2014).
42. Nott, A. Splicing enhances translation in mammalian cells: an additional function of the exon junction complex. *Genes Dev.* 18, 210–222 (2004).
43. Palacios, I. M., Gatfield, D., St Johnston, D. & Izaurralde, E. An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay. *Nature* 427, 753–757 (2004).

44. Uhlen, M. et al. A pathology atlas of the human cancer transcriptome. *Science* 357, (2017).
45. Dagueuet, E. et al. Perispeckles are major assembly sites for the exon junction core complex. *Mol. Biol. Cell* 23, 1765–1782 (2012).
46. Cougot, N., Babajko, S. & Séraphin, B. Cytoplasmic foci are sites of mRNA decay in human cells. *J. Cell Biol.* 165, 31–40 (2004).
47. Riemer, A. B., Keskin, D. B. & Reinherz, E. L. Identification and validation of reference genes for expression studies in human keratinocyte cell lines treated with and without interferon- γ - a method for qRT-PCR reference gene determination. *Exp. Dermatol.* 21, 625–629 (2012).
48. Sherman, M. L., Stone, R. M., Datta, R., Bernstein, S. H. & Kufe, D. W. Transcriptional and post-transcriptional regulation of c-jun expression during monocytic differentiation of human myeloid leukemic cells. *J. Biol. Chem.* 265, 3320–3323 (1990).
49. Bhuvanagiri, M. et al. 5-azacytidine inhibits nonsense-mediated decay in a MYC - dependent fashion. *EMBO Mol. Med.* 6, 1593–1609 (2014).
50. Morel, A. et al. 5azadC treatment upregulates miR-375 level and represses HPV16 E6 expression. *Oncotarget* 8, 46163–46176 (2017).
51. Khwaja, S. S. et al. High E6 Gene Expression Predicts for Distant Metastasis and Poor Survival in Patients With HPV-Positive Oropharyngeal Squamous Cell Carcinoma. *Int. J. Radiat. Oncol.* 95, 1132–1141 (2016).
52. Wichmann, G. et al. The role of HPV RNA transcription, immune response-related gene expression and disruptive TP53 mutations in diagnostic and prognostic profiling of head and neck cancer: Prognostic Profiling of Head and Neck Cancer. *Int. J. Cancer* 137, 2846–2857 (2015).
53. Jiang, M. & Milner, J. Selective silencing of viral gene expression in HPV-positive human cervical carcinoma cells treated with siRNA, a primer of RNA interference. *Oncogene* 21, 6041–6048 (2002).
54. Tan, S., G.E. de Vries, E., G.J. van der Zee, A. & de Jong, S. Anticancer Drugs Aimed at E6 and E7 Activity in HPV-Positive Cervical Cancer. *Curr. Cancer Drug Targets* 12, 170–184 (2012).
55. Filippova, M. et al. The small splice variant of HPV16 E6, E6*, reduces tumor formation in cervical carcinoma xenografts. *Virology* 450–451, 153–164 (2014).

56. Paget-Bailly, P. et al. Comparative RNA sequencing reveals that HPV16 E6 abrogates the effect of E6*I on ROS metabolism. *Sci. Rep.* 9, (2019).
57. Mazloomian, A. et al. Pharmacological systems analysis defines EIF4A3 functions in cell-cycle and RNA stress granule formation. *Commun. Biol.* 2, (2019).
58. Martínez-Salazar, M. et al. Biochemical and proteomic analysis of spliceosome factors interacting with intron-1 of human papillomavirus type-16. *J. Proteomics* 111, 184–197 (2014).
59. McFarlane, M., MacDonald, A. I., Stevenson, A. & Graham, S. V. Human Papillomavirus 16 Oncoprotein Expression Is Controlled by the Cellular Splicing Factor SRSF2 (SC35). *J. Virol.* 89, 5276–5287 (2015).
60. Balistreri, G. et al. The Host Nonsense-Mediated mRNA Decay Pathway Restricts Mammalian RNA Virus Replication. *Cell Host Microbe* 16, 403–411 (2014).
61. Fontaine, K. A. et al. The Cellular NMD Pathway Restricts Zika Virus Infection and Is Targeted by the Viral Capsid Protein. *mBio* 9, (2018).
62. Garcia, D., Garcia, S. & Voinnet, O. Nonsense-Mediated Decay Serves as a General Viral Restriction Mechanism in Plants. *Cell Host Microbe* 16, 391–402 (2014).
63. LeBlanc, J. J. & Beemon, K. L. Unspliced Rous Sarcoma Virus Genomic RNAs Are Translated and Subjected to Nonsense-Mediated mRNA Decay before Packaging. *J. Virol.* 78, 5139–5146 (2004).
64. Mocquet, V. et al. The Human T-Lymphotropic Virus Type 1 Tax Protein Inhibits Nonsense-Mediated mRNA Decay by Interacting with INT6/EIF3E and UPF1. *J. Virol.* 86, 7530–7543 (2012).
65. Wada, M., Lokugamage, K. G., Nakagawa, K., Narayanan, K. & Makino, S. Interplay between coronavirus, a cytoplasmic RNA virus, and nonsense-mediated mRNA decay pathway. *Proc. Natl. Acad. Sci.* 115, E10157–E10166 (2018).
66. Jagannathan, S. & Bradley, R. K. Translational plasticity facilitates the accumulation of nonsense genetic variants in the human population. *Genome Res.* 26, 1639–1650 (2016).
67. Rigalli, J. P. et al. Human papilloma virus (HPV) 18 proteins E6 and E7 up-regulate ABC transporters in oropharyngeal carcinoma. Involvement of the nonsense-mediated decay (NMD) pathway. *Cancer Lett.* 428, 69–76 (2018).

Figure 1: eIF4A3 expression in cervical precancerous and cancerous lesions. Kaplan-Meier overall survival of the 291 TCGA cervical cancer samples (CESC) stratified according to the median of eIF4A3 mRNA expression.

Figure 2: The depletion of eIF4A3 increases HPV16 E7 protein expression. (A) Western blotting analyses showing endogenous E6 (left panel), E7 (right panel) and eIF4A3 protein levels in CTL or eIF4A3-depleted SiHa cells. β -actin was used as a loading control. (B) Representative scheme of unspliced and alternatively spliced HPV16 early transcripts. Real time PCR primer sets are in dotted lines (E6All, E6, E6*I and E7) and RT-PCR primer sets are represented by arrows (F6 and R7). (C) RT-PCR showing alternative splicing patterns of HPV16 E6/E7 transcripts in CTL or eIF4A3-depleted SiHa cells. KPNA1 was used as a positive control of eIF4A3 depletion efficiency. B2M and 18S were both used as loading controls. (D) RT-qPCR analysis of HPV16 E6all, E6, E6*I and E7 expression in SiHa cells treated with siRNA-CTL or siRNA-eIF4A3 for 48h. 18S was used for normalization. Data are represented as means of seven independent experiments. Error bars = s.e.m.; ** $p < 0.01$; *** $p < 0.001$; two-tailed student's t-test. (E) E6/E6All ratio of mRNA levels quantified by real-time PCR in CTL or eIF4A3-depleted SiHa cells. Data are represented as means of seven independent experiments. Error bars = s.e.m.

Figure 3: Subcellular localization of HPV16 E6/E7 mRNA using RNAscope. (A) Representative scheme of E6 probe (green) and E7 probe (red) recognizing HPV16 transcripts. (B) Transcripts were detected as punctate dots in eIF4A3-depleted SiHa cells or control cells by confocal microscopy (Z-section). Nuclei were stained with DAPI (blue). (C) The number of dots per cell are represented as a box-plot graph. The number of HPV16 RNA dots per cell was counted in 80 and 62 cells for CTL and eIF4A3-depleted cells, respectively. One-tailed student's t-test; *** $p < 0.001$. (D) E6 RNA probe (red) and E7 RNA probe (grey) were detected as punctate dots in SiHa cells transfected with the pEYFP-eIF4A3 vector for 24h. Nuclei were stained with DAPI (blue). (E) The number of dots per cell are represented as a box-plot graph. The number of HPV16 RNA dots per

cell was counted in 101 cells for CTL and 86 cells for GFP-eIF4A3. One-tailed student's t-test; *** $p < 0.001$.

Figure 4: eIF4A3 associates with HPV16 E6/E7 transcripts. (A) Western blotting analysis showing the immunoprecipitation of eIF4A3 protein in SiHa cells. (B) RNA ImmunoPrecipitation (RIP) analysis of RNAs from SiHa cells associated with either rabbit IgG control (IgG-IP) or eIF4A3 antibody (eIF4A3-IP). KPNA1 and the intronless SF3B5 mRNAs were used as positive and negative controls, respectively.

Figure 5: The inhibition of protein synthesis using CHX upregulates HPV16 E6*1/E7 mRNA levels. (A) RT-qPCR analysis of SC35 and HPV16 E6all, E6, E6*1 and E7 mRNA levels in SiHa cells treated for 6 h with DMSO or CHX at 100 μ g/mL. 18S was used for normalization. SC35 was used as a positive control of NMD inhibition. Data are represented as means of three independent experiments. Error bars = s.e.m.; * $p < 0.05$; one-tailed student's t-test; n.s. non-significant. (B) HPV16 (green) and c-JUN (red) transcripts were detected as punctate dots in DMSO- or CHX-treated SiHa cells using RNAscope and confocal microscopy. Data from one representative experiment of three independent experiments. (C) The number of HPV16 mRNA dots per cell was counted in 50 and 52 cells for DMSO and CHX treated cells, respectively, and are presented as a bar graph. One-tailed student's t-test; *** $p < 0.001$.

Figure 6: Upregulation of HPV16 E7 expression in SiHa cells upon 5-azaC treatment. (A) RT-qPCR analysis of ASNS, SC35 and HPV16 E6all, E6, E6*1 and E7 expression in SiHa cells treated with increasing amount (0, 0.5, 1, 3, 5 and 10 μ M) of 5-azaC (upper panel) or 5-azadC (lower panel) for 18h. 18S was used for normalization. Data are represented as means of three independent experiments. Error bars = s.e.m.; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; one-tailed student's t-test. (B) Western blotting analysis of E6 and E7 protein levels in SiHa cells treated with 0, 5 or 10 μ M of 5-azaC or 5-azadC for 18h. β -actin was used as a loading control. A representative blot is shown. (C) Relative quantification of E6 and E7 expression normalized with β -actin. The data are presented as mean values from two independent experiment. Error bars = s.e.m.

Figure 7: Model for the regulation of HPV16 transcripts by the eIF4A3 factor. A decrease of NMD efficiency would increase E6*I/E7 mRNA levels leading to a high expression of the E7 protein. High eIF4A3 levels and high NMD activity would ensure low levels of E6*I/E7 mRNAs leading to a moderate E7 protein production. The color boxes represent ORFs and color ovals represent the proteins generated from the corresponding transcripts.

Figure S1: Splicing pattern of HPV16 early transcripts. RT-PCR showing alternative splicing patterns of HPV16 E6/E7 transcripts in SiHa and ca Ski cell lines. C-33A is an HPV-negative cell line derived from cervical carcinoma. B2M and 18S were used as loading controls. Agarose gel images are representative of two independent experiments.

Figure S2: Uncropped Western blotting used in figure 2. Analyses showing endogenous E6, E7, eIF4A3 and β -actin protein levels in CTL or eIF4A3-depleted SiHa cells. * correspond to the lane displayed in figure 2.

Figure S3: HPV16 E6 RNAs do not colocalize with stress granules. HPV16 E6 RNA-FISH were performed with SiHa cells transfected with the peGFP-MLN51 vector. MLN51 was recruited to stress granules. The image corresponds to a Z-section generated by confocal microscopy. Insets represent images enlargements and indicate the size of MLN51 and E6 RNA dots.

Figure S4: HPV16 E6 RNAs do not colocalize with P-bodies. HPV16 E6 RNA-FISH were performed with SiHa cells transfected with the peGFP-Dcp1 vector. Dcp1 was recruited to P-bodies. The image corresponds to a Z-section generated by confocal microscopy. Insets represent images enlargements and indicate the size of Dcp1 and E6 RNA dots.

Figure S5: Uncropped Western blotting used in figure 6. Analyses showing endogenous E6, E7 and β -actin protein levels in SiHa cells treated or not with increasing dose of 5azaC or 5azadC.

Table S1. PCR and qPCR primers used in this study.

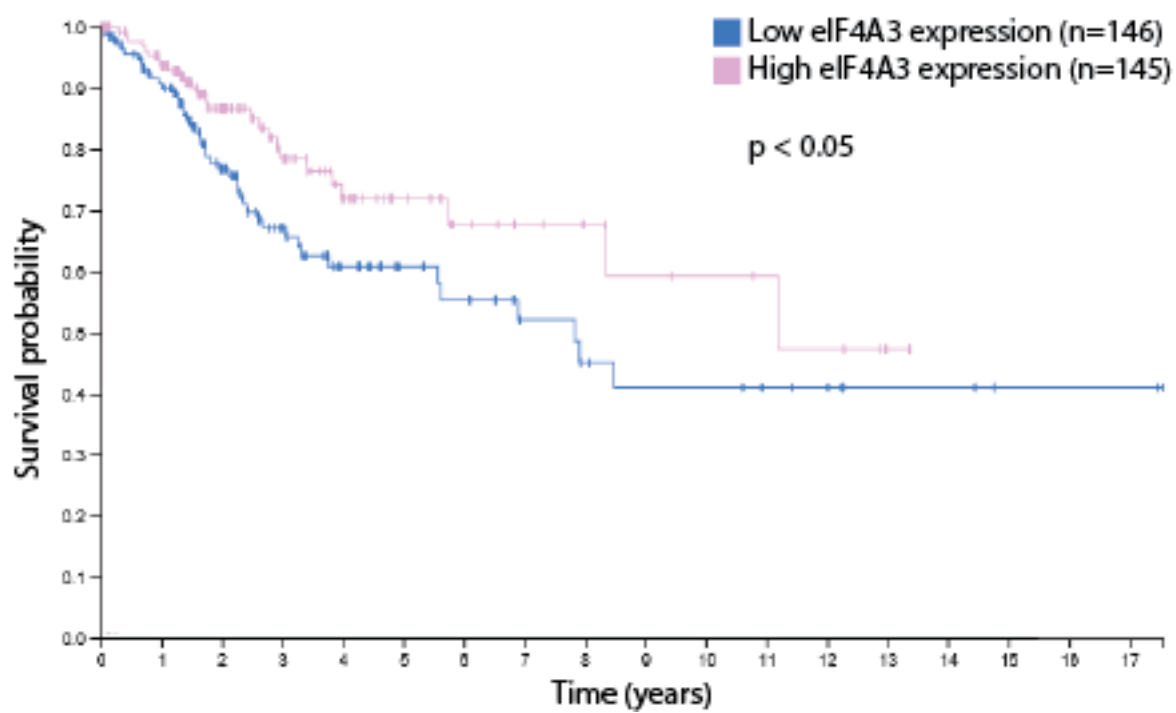


Figure 1

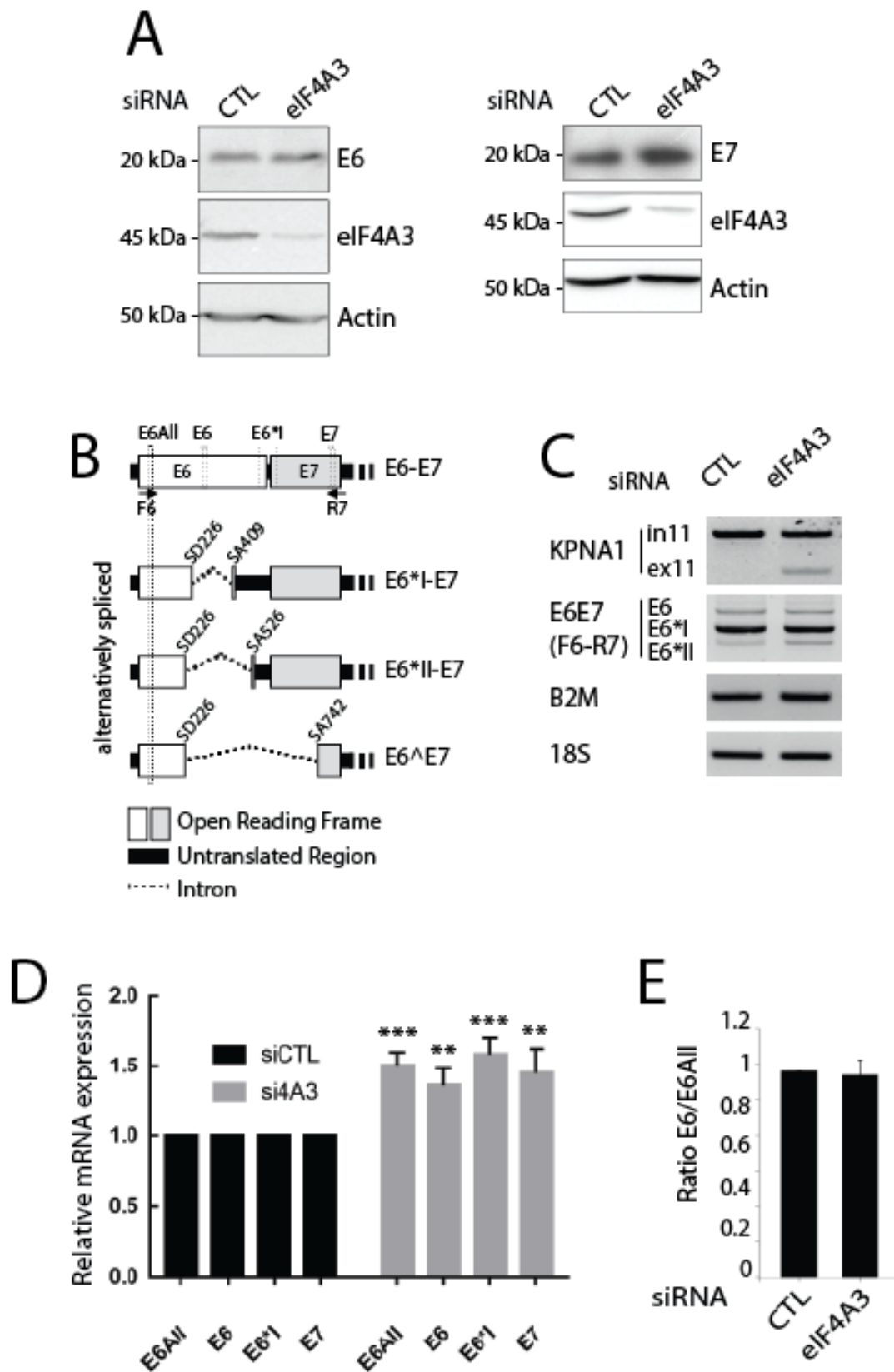


Figure 2

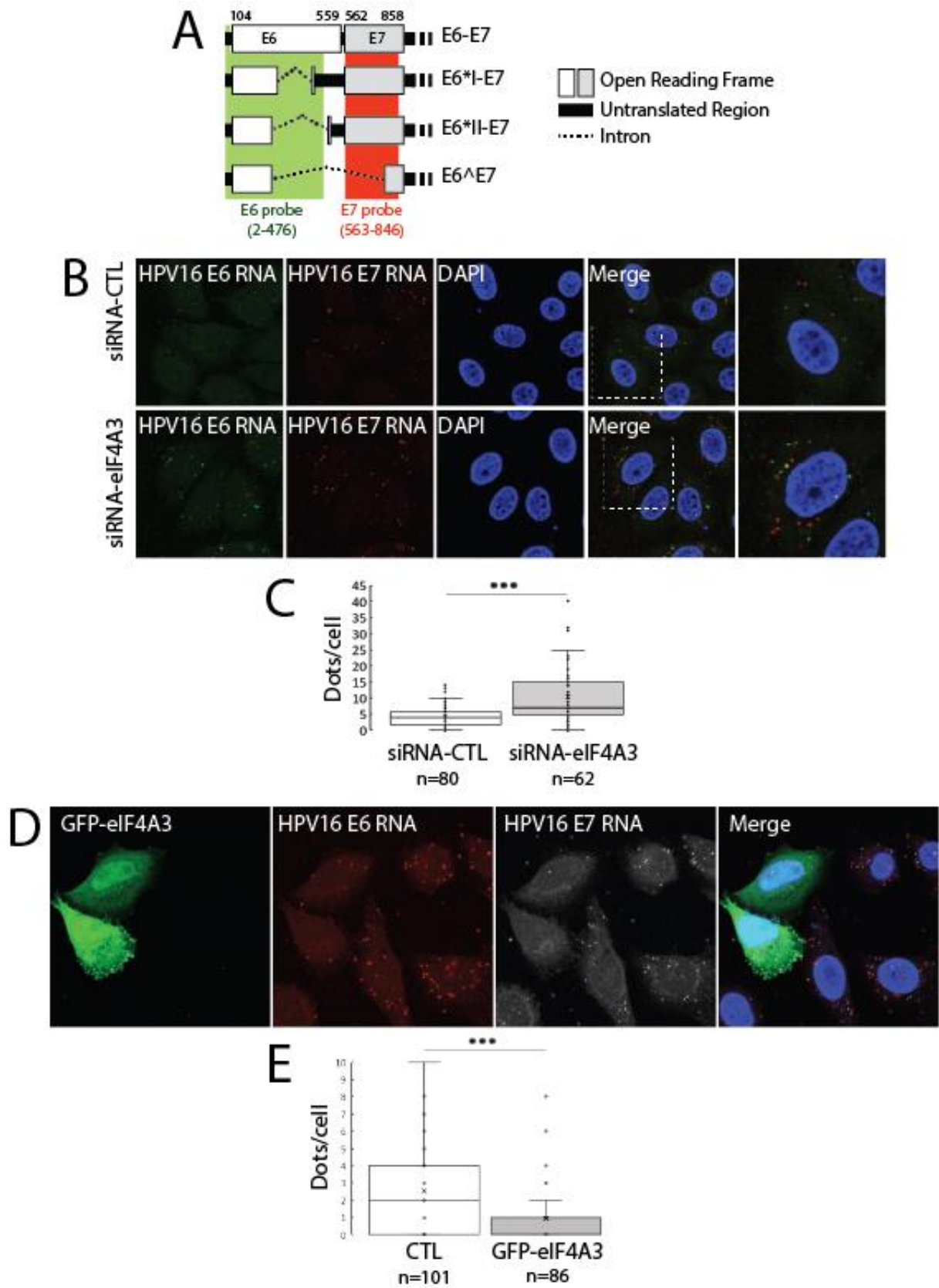


Figure 3

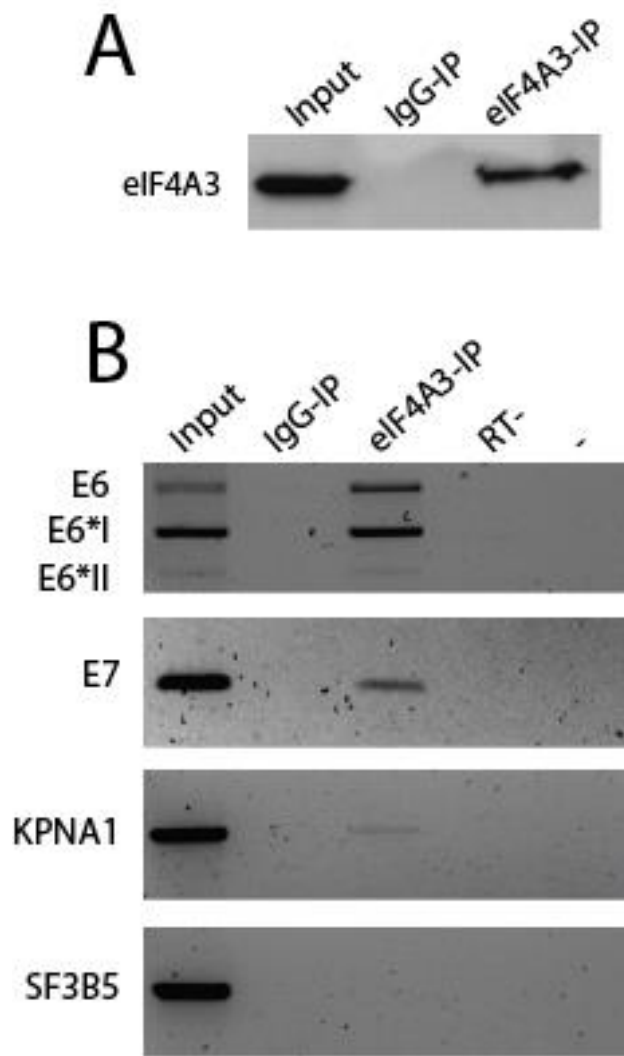


Figure 4

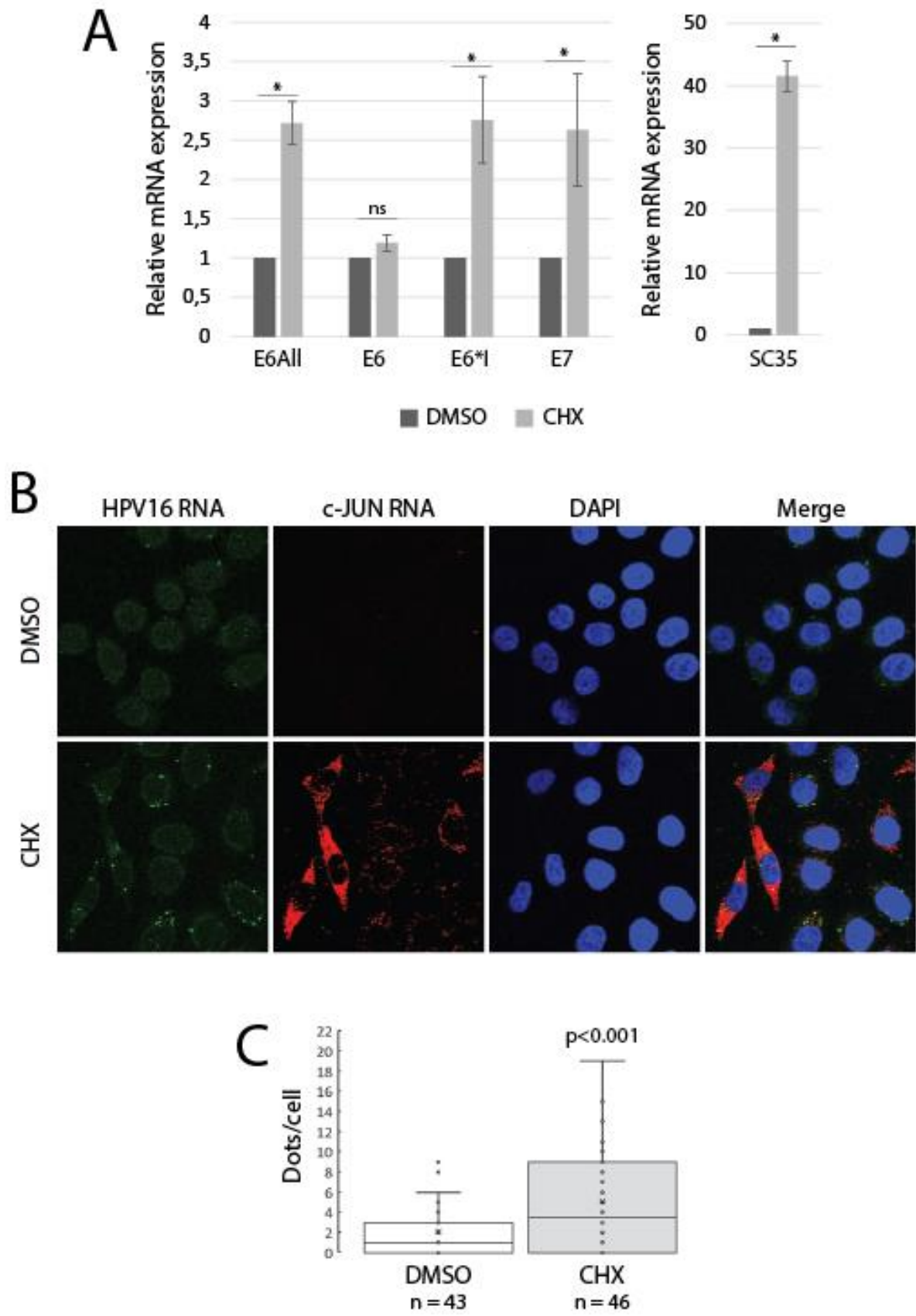


Figure 5

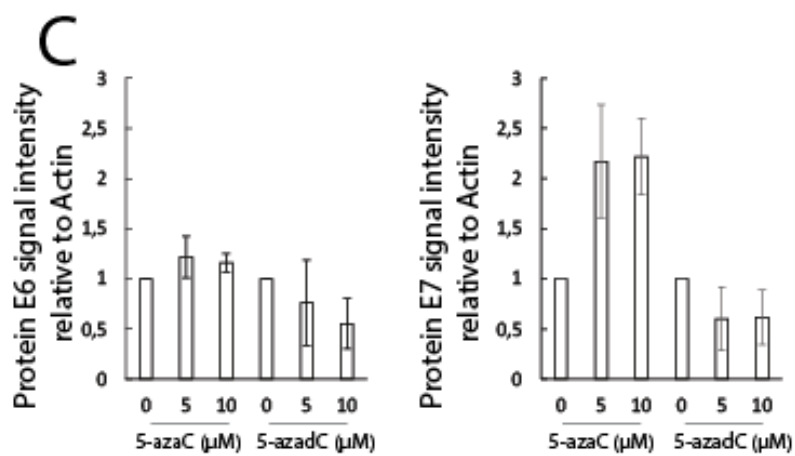
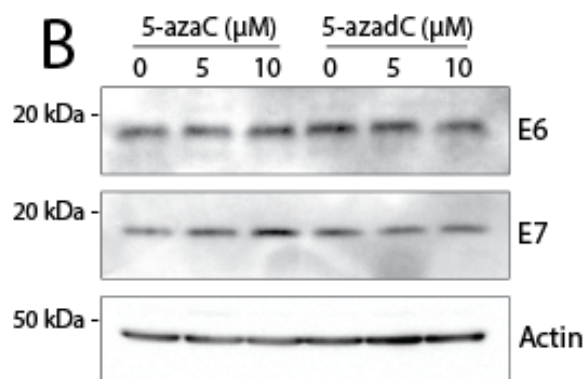
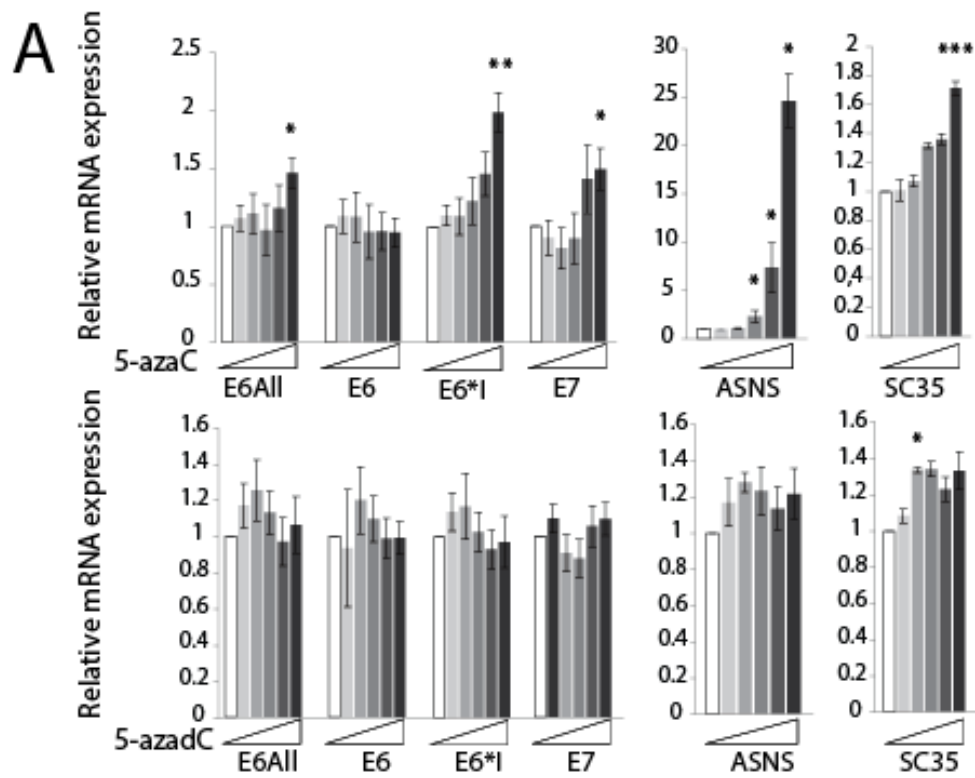


Figure 6

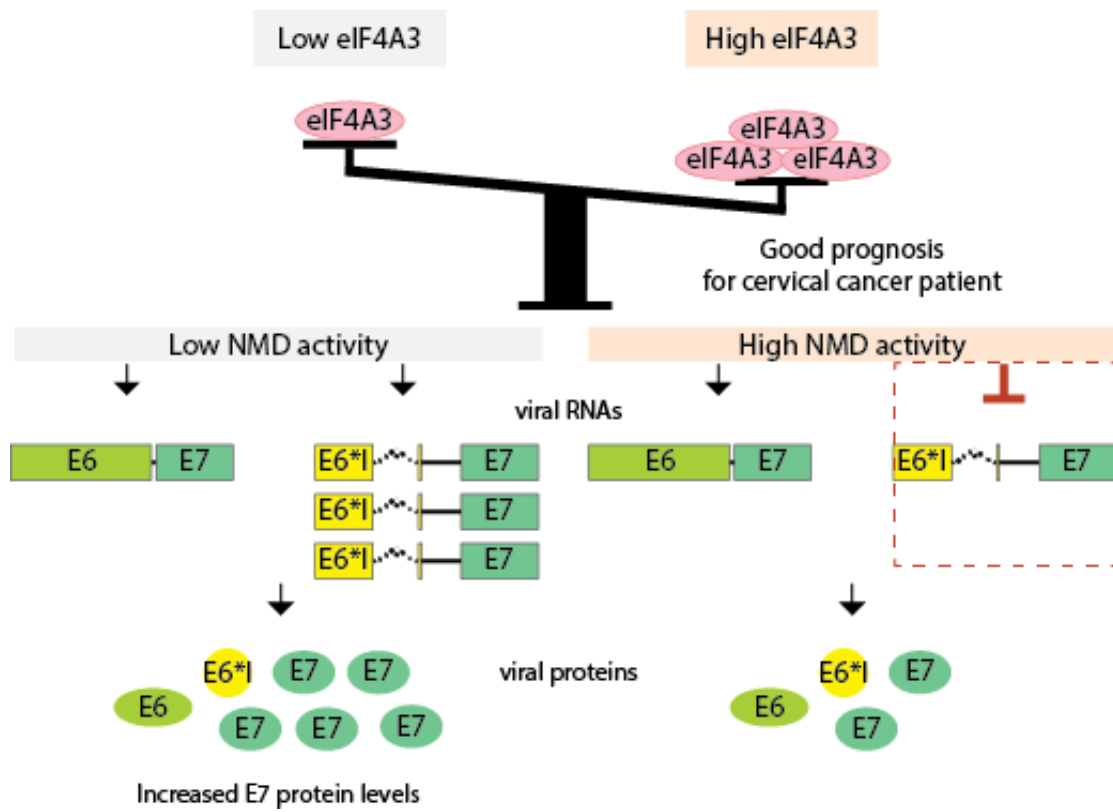


Figure 7

Supplementary Material

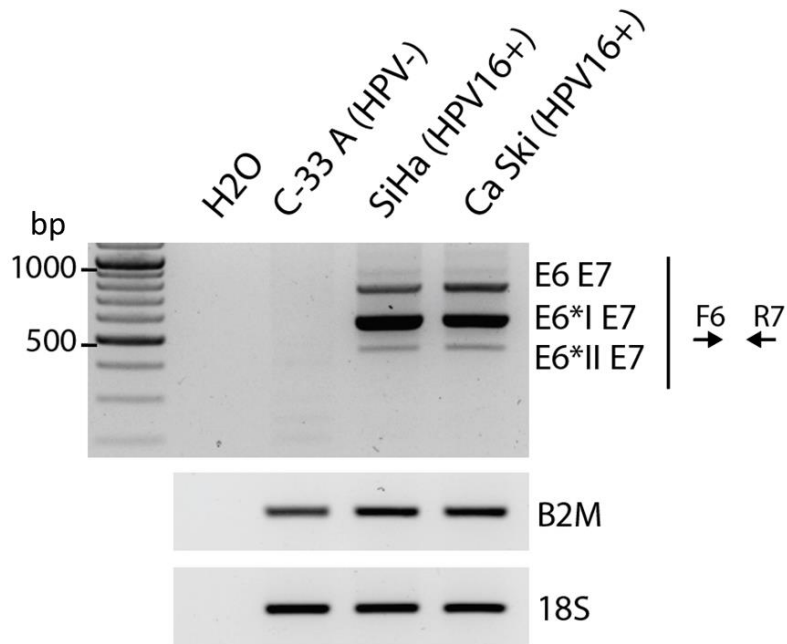


Figure S1: Splicing pattern of HPV16 early transcripts. RT-PCR showing alternative splicing patterns of HPV16 E6/E7 transcripts in SiHa and ca Ski cell lines. C-33A is an HPV-negative cell line derived from cervical carcinoma. B2M and 18S were used as loading controls. Agarose gel images are representative of two independent experiments.

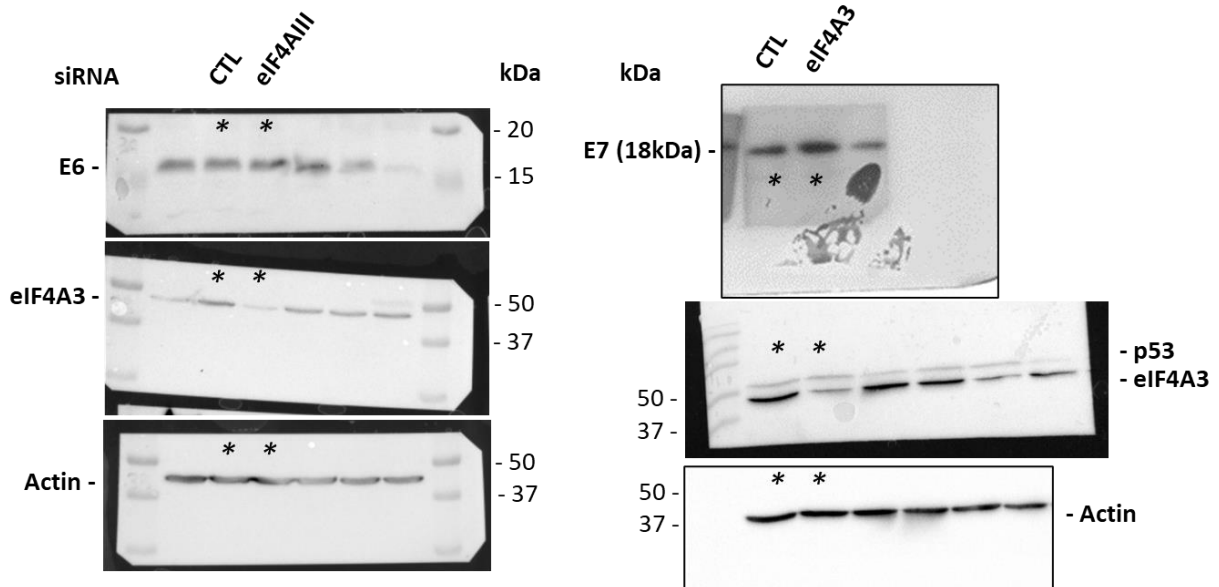


Figure S2: Uncropped Western blotting used in figure 2. Analyses showing endogenous E6, E7, eIF4A3 and β -actin protein levels in CTL or eIF4A3-depleted SiHa cells. * correspond to the lane displayed in figure 2.

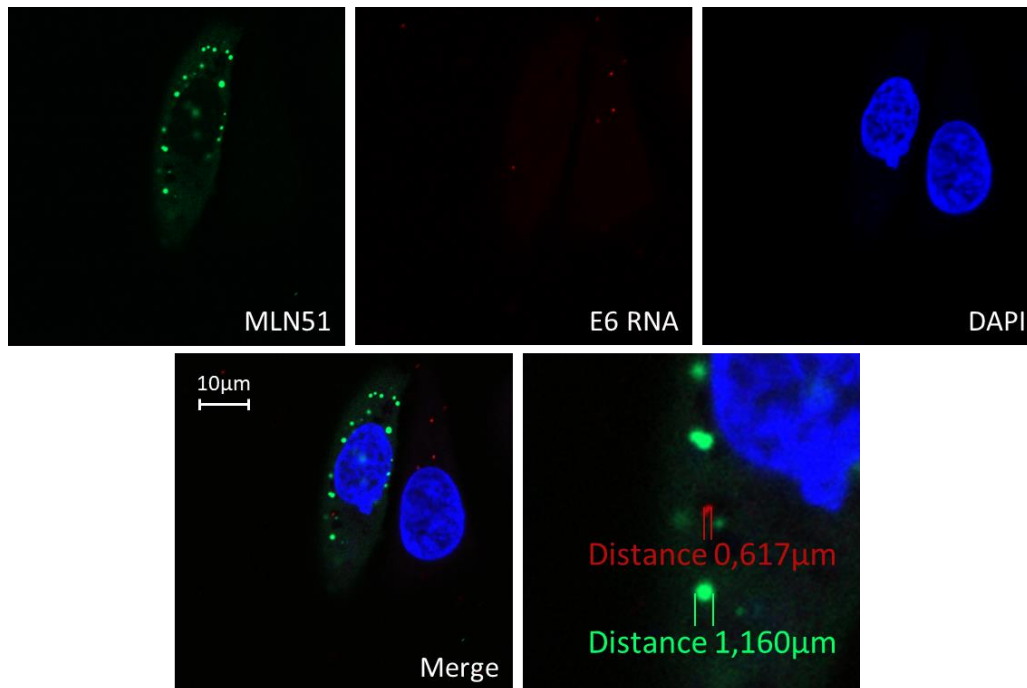


Figure S3: HPV16 E6 RNAs do not colocalize with stress granules. HPV16 E6 RNA-FISH were performed with SiHa cells transfected with the peGFP-MLN51 vector. MLN51 was recruited to stress granules. The image corresponds to a Z-section generated by confocal microscopy. Insets represent images enlargements and indicate the size of MLN51 and E6 RNA dots.

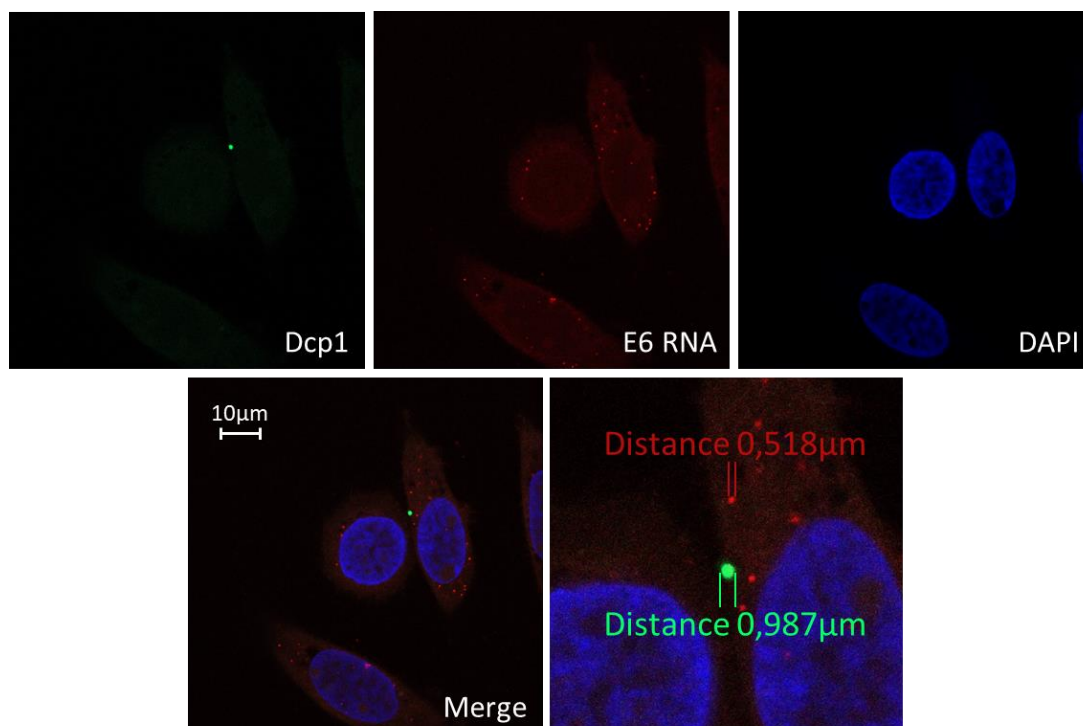


Figure S4: HPV16 E6 RNAs do not colocalize with P-bodies. HPV16 E6 RNA-FISH were performed with SiHa cells transfected with the peGFP-Dcp1 vector. Dcp1 was recruited to P-bodies. The image corresponds to a Z-section generated by confocal microscopy. Insets represent images enlargements and indicate the size of Dcp1 and E6 RNA dots.

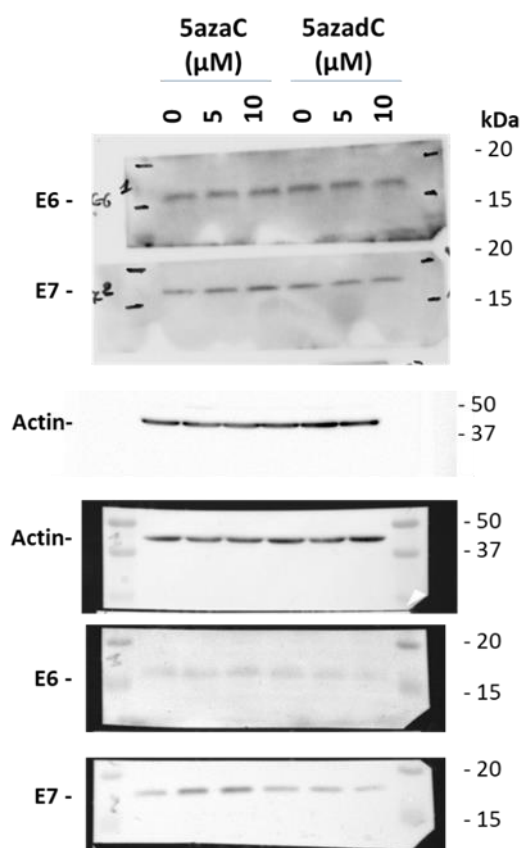


Figure S5: Uncropped Western blotting used in figure 6. Analyses showing endogenous E6, E7 and β -actin protein levels in SiHa cells treated or not with increasing dose of 5azaC or 5azadC. The upper blot is displayed in figure 6B and was revealed on an autoradiographic film. The lower blot is from a second independent experiment revealed with ChemidocTM camera.

Table S1. PCR and qPCR primers used in the study.

Primer	Forward sequence (5'-3')	Reverse sequence (5'-3')	Use
16E6all	GCAATGTTTCAGGACCCACA	TTGTTTGCAGCTCTGTGCAT	qPCR
16E6	TGACTTTGCTTTTCGGGATTT	ACAGCATATGGATTCCCATCTC	qPCR
16E6*I	CTGCGACGTGAGGTGTATTA	TGTCCAGGTGTCTTTGCTTT	qPCR
16E7	ACAAGCAGAACCGACAGAG	GCCCATTAACAGGTCTTCCA	qPCR and PCR
rRNA18S	GCAATTATTCCTCATGAACG	GGCCTCACTAAACCATCCAA	qPCR and PCR
eIF4A3	ACGAGCAATCAAGCAGATCA	AGGTGGCTGTTTTCTGTG	qPCR
SC35	CGGTGTCCTCTTAAGAAAATGATGTA	CTGTACACAACCTGCGCCTTTT	qPCR
ASNS	GGAAGACAGCCCCGATTACT	AGCACGAACTGTTGTAATGTCA	qPCR
SF3B5	ACCGCTACACCATCCATAGC	AGGCTGAAGCATCTTTTCCA	PCR
B2M	GATGAGTATGCCTGCCGTGTG	CAATCCAAATGCGGCATCT	PCR
KPNA1	AGTGGTTTCTCCTGCTTTGC	GCTGATCCTCCAGAAGTTGC	PCR
E6E7ORF	ATGCACCAAAGAGAACTGC (F6)	GCCCATTAACAGGTCTTCCA (R7)	PCR

2.2 Régulation de l'expression de l'oncoprotéine E6 par un agent déméthylant, le 5azadC

ONCOLOGY LETTERS 19: 1074-1081, 2020

DNA demethylation agent 5azadC downregulates HPV16 E6 expression in cervical cancer cell lines independently of TBX2 expression

JEROME PERRARD¹, ADRIEN MOREL², KOCEILA MEZNAD¹, PHILIPPE PAGET-BAILLY¹,
VERONIQUE DALSTEIN^{3,4}, DAVID GUENAT^{1,5}, CELINE MOURAREAU^{3,4}, CHRISTINE CLAVEL^{3,4},
SYLVIE FAUCONNET¹, AURELIE BAGUET¹, CHRISTIANE MOUGIN^{1,5} and JEAN-LUC PRETET^{1,5}

¹Équipe d'Accueil 3181, University of Bourgogne Franche-Comté, Laboratoire d'Excellence Lipoprotéines et Santé, Prévention et Traitement des Maladies Inflammatoires et du Cancer, 25000 Besançon, France; ²Center For Research in Genetics and Genomics-CIGGUR, GENIUIROS Research Group, School of Medicine and Health Sciences, Universidad del Rosario, Bogotá 112041, Colombia; ³Institut National de la Santé et de la Recherche Médicale, Unité Mixte de Recherche-S 1250 Pathologies Pulmonaires et Plasticité Cellulaire, Université de Reims Champagne-Ardenne, Faculté de Médecine; ⁴Centre Hospitalier Universitaire Reims, Laboratoire Biopathologie, 51000 Reims; ⁵Centre National de Référence Papillomavirus, CHU Besançon, Boulevard Alexandre Fleming, 25000 Besançon, France

Received March 7, 2019; Accepted September 9, 2019

DOI: 10.3892/ol.2019.11158

Abstract. HPV16 is the most carcinogenic human papillomavirus and causes >50% of cervical cancers, the majority of anal cancers and 30% of oropharyngeal squamous cell carcinomas. HPV carcinogenesis relies on the continuous expression of the two main viral oncoproteins E6 and E7 that target >150 cellular proteins. Among them, epigenetic modifiers, including DNA Methyl Transferases (DNMT), are dysregulated, promoting an aberrant methylation pattern in HPV-positive cancer cells. It has been previously reported that the treatment of HPV-positive cervical cancer cells with DNMT inhibitor 5-aza-2'-deoxycytidine (5azadC) caused the downregulation of E6 expression due to mRNA destabilization that was mediated by miR-375. Recently, the T-box transcription factor 2 (TBX2) has been demonstrated to repress HPV LCR activity. In the current study, the role of TBX2 in E6 repression was investigated in HPV16 cervical cancer cell lines following 5azadC treatment. A decrease of E6 expression was accompanied by p53 and p21 restoration. While TBX2 mRNA was upregulated in 5azadC-treated SiHa and Ca Ski cells, TBX2 protein was not detectable. Furthermore, the overexpression of TBX2

protein in cervical cancer cells did not allow the repression of E6 expression. The TBX2 transcription factor is therefore unlikely to be associated with the repression of E6 following 5azadC treatment of SiHa and Ca Ski cells.

Introduction

High risk human papillomaviruses (HR HPV) are ~55 nm non-enveloped circular DNA viruses causing anogenital and oropharyngeal cancers (1). Carcinogenic properties of HR HPV are mainly due to the continuous expression of E6 and E7 oncoproteins. Indeed, the most described effects of these two oncoproteins are to bind and abrogate the functions of the tumor suppressor proteins p53 and pRb respectively (2-4). This leads to altered cell cycle regulation with increased cell proliferation rate, immortalization, chromosomal instability driving malignant transformation of infected cells (5). However, E6 and E7 interact with many other proteins and particularly with epigenetic enzymes. Among them, DNMT1 is a DNA methyltransferase maintaining methylation on CpG dinucleotide during cell division. This epigenetic mark generally drives chromatin compaction, thus repressing genes under its dependence. Interestingly, E6 and E7 expression are also regulated by epigenetic mechanisms, and methylation level on viral promoter is higher in HPV-associated cancers than in precancerous lesions or in normal tissues (6). Recently, studies revealed that E6 and E7 expression was downregulated following treatment of HPV positive cancer cell lines by 5aza-2'-deoxycytidine (5azadC), a DNA methyl transferase (DNMT) inhibitor also named decitabine (7-9). The downregulation of E6 was partly due to the upregulation of miR-375 (8,9), known to target the early HPV16 transcripts (10). Indeed, E6 expression was only partially restored by miR-375 inhibitor in 5azadC-treated cells (9).

Correspondence to: Professor Jean-Luc Prétet, Équipe d'Accueil 3181, University of Bourgogne Franche-Comté, Laboratoire d'Excellence Lipoprotéines et Santé, Prévention et Traitement des Maladies Inflammatoires et du Cancer, 19 rue Ambroise Paré, 25000 Besançon, France
E-mail: jean_luc.pretet@univ-fcomte.fr

Key words: epigenetic, DNA methyltransferase inhibitor, 5-aza-2'-deoxycytidine, human papillomavirus-induced cancer, T-Box

This suggests that other mechanisms are involved in E6 downregulation following 5azadC treatment. In fact, although expression of transcription factors known to bind to early viral promoter (SP1, API, YY1 and NF1) is not modified by 5azadC treatment, not all transcription factor expression has been assessed. This is why the role of T-box transcription factor 2 (TBX2), another transcription factor binding viral promoter, has been evaluated.

In 2013, it has been shown that the HPV16 LCR could be repressed, at least *in vitro*, by TBX2, a member of the T-Box protein family (11). This family include transcription factors involved in embryonic development and encoded by highly conserved genes among vertebrates (12). T-Box factors recognize and bind the core sequence GGTGTGA, also known as the T-element (13). Although the HPV16 LCR lacks the canonical T-element, TBX2 is able to repress the p97 activity by binding a sequence located between nt 7564 and nt 7756 (11). Conversely to other T-Box family members, TBX2 has repressive activities thanks to the presence of a strong repression domain located in its C-terminal region.

The role of TBX2 in tumorigenesis remains controversial because of both anti- and pro-tumorigenic activities. On the one hand, TBX2 is able to bypass senescence and activate cell proliferation by repressing p14^{ARF} and p21 in different cancer models (12). TBX2 expression was significantly increased in prostate cancers compared to healthy adjacent tissues (14), suggesting an oncogenic activity. On the other hand, TBX2 inhibits cell cycle progression in lung adenocarcinoma (15) and its expression is decreased in lung cancers compared to normal tissue (16). Furthermore, TBX2 methylation has been associated with a poor prognosis in chronic lymphocytic leukemia (17), endometrial cancer (18) and bladder cancer (19). Thus, TBX2 may also exert an anti-tumor activity depending on the context and cellular origin of cancers.

In the present study, the role of TBX2 was investigated to determine whether, in addition to miR-375, it could induce E6 repression following 5azadC treatment of HPV16 cervical cancer cells.

Materials and methods

Cell lines and cell culture. MCF-7 cells [mammary cancer cells, American Type Culture Collection (ATCC)], SiHa cells (cervical cancer cells, ATCC) and U-2 OS cells (bone osteosarcoma cells, ATCC) were cultured in Dulbecco's modified Eagle's medium (DMEM); Ca Ski cells (cervical cancer cells, ATCC) were grown in Roswell Park Memorial Institute (RPMI) medium while C-33 A and HeLa cells (both cervical cancer cells, ATCC) were cultured in Eagle's minimum essential medium (EMEM). All media were supplemented with 10% (v/v) fetal bovine serum (FBS, Lonza) but not with antibiotics. Cells were cultured at 37°C in a 5% CO₂ humidified incubator.

Drug treatment. SiHa and Ca Ski cells were treated by 5-aza-2'-deoxycytidine (5azadC) (Epigentek). This demethylating agent was dissolved in DMSO at 220 mM and then diluted in appropriate medium at 0.25 or 5.0 μ M to treat SiHa cells (10,000 cells/cm²) and Ca Ski cells (10,000 cells/cm²) for 24, 48, 72 and 96 h. The treatment medium was renewed every

day. Untreated HeLa, C-33 A and MCF-7 cells were harvested as controls for western blotting studies.

U-2 OS, SiHa and Ca Ski cells were transfected as described below.

Plasmid transfection and luciferase assay. U-2 OS cells were plated in 96-well microplates at a density of 15,000 cells per well. They were transfected for 24 h with either pT-REx-DEST30/empty or pT-REx-DEST30/TBX2-3XFlag plus a mixture of pGL3-Luc-I6LCR (obtained by cloning the HPV16 DNA sequence from nt 7135 to nt 105 into the pGL3 plasmid) and pRenilla (Promega) with JetPEI[®] transfection reagent (Polyplus-transfection) according to the manufacturer's recommendations. Cells were lysed by 1X Lysis reagent of Dual-Luciferase Reporter Assay System kit (Promega). The luciferase assay reagent was mixed with lysates and the luminescence was recorded using the TECAN Infinite 200 Pro instrument. Unfortunately, luciferase activity could not have been normalized by Renilla signals because TBX2 strongly repressed the CMV-driven vector pRenilla, as already documented by Schneider and collaborators (11).

SiHa and Ca Ski cells were plated in 6-well plates at a density of 350,000 cells per well and transfected with pT-REx-DEST30/empty or pT-REx-DEST30/TBX2-3XFlag using JetPEI[®] transfection reagent according to the manufacturer's recommendations. At 48 h after plasmid transfection, cells were harvested for checking overexpression efficiency and subsequent analyses *via* RTqPCR and western-blotting.

5azadC and TBX2 combinatory treatment. SiHa and Ca Ski cells were seeded at 10,000 cells/cm² in 6-well plates and treated with 5azadC at 0.25 μ M for 72 h. The treatment medium was renewed every day. Twenty-four hours after the beginning of 5azadC treatment, cells were transfected with pT-REx-DEST30/empty or pT-REx-DEST30/TBX2-3XFlag using JetPEI[®] transfection reagent according to the manufacturer's recommendations. Cells were then cultured for 48 h and harvested for E6 expression analysis by RT-qPCR.

Transfection of siRNA. MCF-7 cells were transfected with 20 nM of siRNA targeting TBX2 (5'-GGA-GCU-GUG-GGACCA-GUU-CTT-3') or control siRNA (SR-CL000-005; Eurogentec) using Lipofectamine 2000 (Thermo Fisher Scientific, Inc.) according to the manufacturer's instructions (ratio siRNA/Lipofectamine of 1:3). Forty-eight hours after transfection, cells were harvested directly in RiboZol[™] solution for RNA extraction or scrapped, centrifuged and lysed in RIPA solution for protein extraction.

RNA extraction and reverse transcription. Total cellular RNAs were isolated by RiboZol[™]-chloroform method (VWR). Briefly, cells were lysed in 500 μ l of RiboZol and 100 μ l of chloroform were added. After a centrifugation at 12,000 g (15 min, 4°C), aqueous phase was harvested and incubated 10 min with 500 μ l of isopropanol. Then, total RNAs were pelleted by centrifugation at 12,000 g, (10 min, 4°C), washed with cold ethanol and dissolved in molecular biology grade water. cDNAs were synthesized using 500 ng of total RNA with the Maxima First Strand cDNA Synthesis kit (Thermo

Fisher Scientific, Inc.) according to the manufacturer's recommendations.

RT-qPCR. Primers were synthesized by Eurogentec. Real-time quantitative PCR was performed using SYBR Green real time PCR master mix (Life Technologies) in the ABI 7500 Real-Time PCR System (Applied Biosystems). The cDNAs were amplified using the following cycling parameters: 95°C for 5 min followed by 40 cycles of 95°C for 30 sec and 60°C for 1 min. Transcript levels for E6 and TBX2 were measured with GAGAACTGCAATGTTTCAGGACC forward and TGTATA GTTGTTCAGCTCTGTGC reverse primers for E6, CTC TGACAAGCACGGCTTCA forward and TGTCGTTGGCTC GCACTATG reverse primers for TBX2 and normalized with β 2M mRNA level measured with GATGAGTATGCCGTGTG forward and CAATCCAATGCGGCATCT reverse primers using the $2^{-\Delta\Delta Ct}$ method (20).

Western blotting. Following different conditions, cells were harvested by scraping in PBS and stored in dry pellets at -80°C until use. Proteins were extracted with radio immunoprecipitation assay (RIPA) lysis buffer [50 mM Tris/HCl, pH 7.4, 150 mM NaCl, 1% (v/v) Nonidet P-40, 0.5% (w/w) Na deoxycholate, 1 mM EDTA, 30 μ g/ml protease inhibitor] (Roche Diagnostics). After sonication, protein concentrations were determined using the Bio-Rad Protein assay (Bio-Rad) according to the manufacturer's recommendations. Each sample was resolved on 12% SDS-PAGE gels and then transferred to Hybond[®] polyvinylidene difluoride (PVDF) membranes (GE Healthcare). The membranes were blocked with 5% nonfat milk overnight at 4°C under constant shaking and then incubated for 2 h at room temperature under constant agitation with primary antibodies: Anti-E6HPV16, 1/1,000 (2E-3F8; Euromedex, Souffelweyersheim, France); DO7 anti-p53, 1/2,000 (554298, BD Biosciences, Le Pont de Claix, France); 6B6 anti-p21, 1/2,000 (554228; BD Biosciences); AC15 anti- β -actin, 1/20,000 (A1978; Sigma-Aldrich); C-17 anti-TBX2, 1/500 (sc-17880). After several washes, membranes were incubated with goat anti-mouse, goat anti-rabbit (BD Pharmingen) or rabbit anti-goat (Agilent) immunoglobulin antibodies conjugated with horseradish-peroxidase. The immune complexes were revealed using an enhanced chemiluminescence detection system with Pierce ECL2 western blotting substrate (Thermo Fisher Scientific, Inc.) using ChemiDoc XRS+. The band densities were normalized against the β -actin internal control and analyzed by Image Lab software (Bio-Rad).

Statistical analysis. A two-tailed unpaired Student's t-test was used to assess differences between two groups. In each case non-treated cells served as reference control. Two-way ANOVA followed by Levene test were used to analyze differences in E6 expression following combinatory treatments. All data were obtained from at least three independent experiments or as specified for each figure and are presented as mean \pm standard deviation (SD). $P < 0.05$ was considered to indicate a statistically significant difference.

Results

5azadC treatment induces HPV16 E6 downregulation and p53 and p21 upregulation. As we previously published, 5azadC

treatment of SiHa cells induced a time-dependent decrease of E6 expression at both mRNA and protein levels (Fig. 1A-C). The effect of 5azadC on E6 repression was observed starting 24 h after treatment, with a maximal decrease in transcripts reaching 60% at 96 h of treatment (Fig. 1C), a result consistent with the decreased expression of E6 observed at the protein level (Fig. 1A and B). In contrast, the effect of 5azadC on E6 RNA expression did not appear dependent on the concentration used (0.25 or 5 μ M) to treat SiHa cells. At the protein levels, 5azadC at 5 μ M seemed to cause sooner and more prominent decrease in E6 expression. As expected, the decreased level of E6 was accompanied by an increased expression of p53 and an up-regulation of the cyclin dependent kinase inhibitor p21.

E6 expression was also investigated in Ca Ski cells treated with 0.25 and 5 μ M of 5azadC up to 96 h (Fig. 1D-F). As in SiHa cells, E6 expression was downregulated at both protein and mRNA levels. The downregulation of E6 in Ca Ski cells was also accompanied by the restoration of p53 and p21 expression.

5azadC treatment induces TBX2 mRNA expression. A recent report has shown that TBX2 was able to bind and repress the HPV16 LCR (11). To confirm this, the pT-REx-DEST30/TBX2-3XFlag was co-transfected with pGL3-Luc-16LCR plasmid in U-2 OS cells and the p97 promoter activity was measured. As shown in Fig. 2, the overexpression of TBX2 induced a reduction of the relative luciferase activity of 40% compared to cells transfected with the empty vector. TBX2 transfection efficiency in U-2 OS cells was assessed by western blotting (Fig. S1).

Then, the expression of TBX2 was studied in 5azadC-treated SiHa and Ca Ski cells. In a first set of experiments, the specificity of TBX2 primers and anti-TBX2 antibody was confirmed in MCF-7 cells, known to express TBX2, treated by TBX2 siRNA (Fig. S1). As shown in Fig. 3A, a slight increase of TBX2 RNA level was observed in SiHa cells treated with 0.25 μ M of 5azadC at 96 h. In Ca Ski cells, no variation of TBX2 expression was observed with the same treatment (Fig. 3B). In contrast, a clear increase of TBX2 RNA was observed in both cell lines for the highest concentration (5 μ M) of 5azadC, especially at 96 h, reaching 3.5 and 3.8-fold in SiHa and Ca Ski cells respectively (Fig. 3A and B). It is interesting to note that the pattern of variation of E6 mRNA (Fig. 1C and F) was not inversely related to that of TBX2.

The endogenous expression of TBX2 was then analyzed by western blotting. The TBX2 protein was not detected in SiHa and Ca Ski cells either treated or not with the highest concentration of 5azadC during 96 h (Fig. 3C). Furthermore, TBX2 was neither detected in HeLa (HPV18 positive) cells and C-33 A (HPV negative) cervical cancer cell lines (Fig. 3D). In contrast, a strong signal was observed in MCF-7 cells (a positive control for TBX2 expression) or in SiHa cells transfected with pT-REx-DEST30/TBX2-3XFlag (Fig. 3C and D). It is noteworthy that TBX2 mRNA expression is 700 and 90-fold more important in MCF-7 compared to Ca Ski and SiHa cells, respectively (not shown). Thus, these data indicate that TBX2 is not expressed or at a very low level in cervical cancer cell lines compared to MCF-7 cells.

Ectopic TBX2 expression does not repress endogenous E6 expression. Then, we investigated whether ectopic TBX2

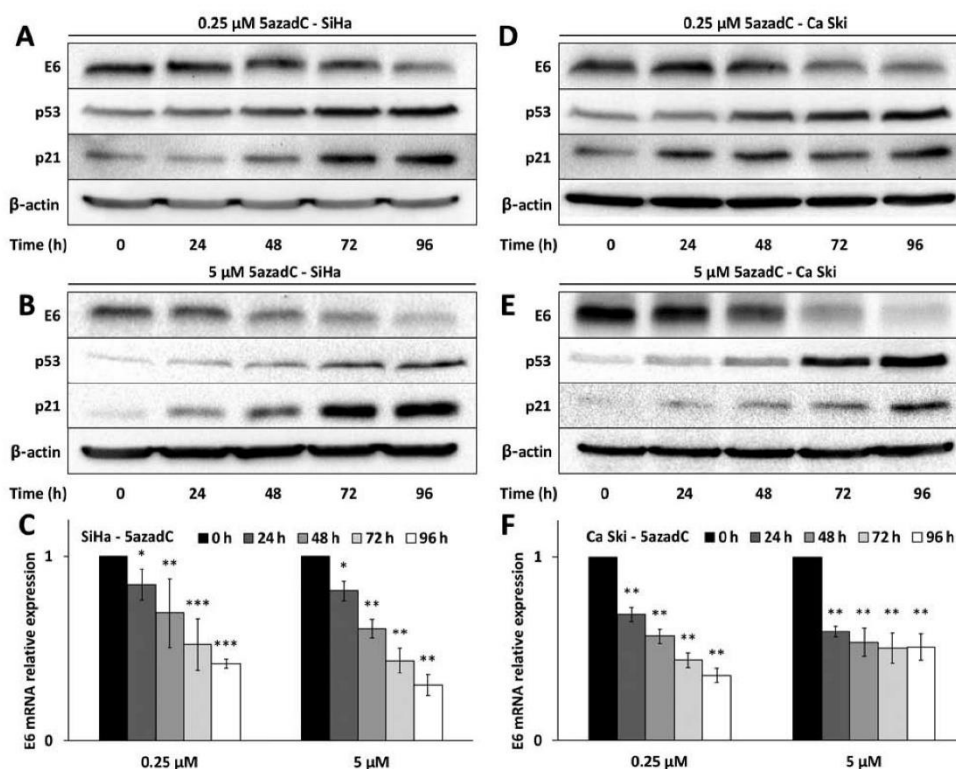


Figure 1. 5azadC treatment downregulates E6 expression. E6, p53, p21 and β -actin protein expression was analyzed by western blotting in the (A) 0.25 μ M and (B) 0.5 μ M treatment in SiHa cells. (C) E6 mRNA relative expression was quantified by RT-qPCR in SiHa cells treated with 0.25 or 5 μ M of 5azadC. E6, p53, p21 and β -actin protein expression was analyzed by western blotting in the (D) 0.25 μ M and (E) 0.5 μ M treatment in Ca Ski cells. (F) E6 mRNA relative expression was quantified by RT-qPCR in Ca Ski cells treated with 0.25 or 5 μ M of 5azadC. The data are presented as mean values from at least three independent experiments. Error bars represent the standard deviation and P-values were calculated using a Student's t-test using non-treated cells as a reference control. *P<0.05, **P<0.01 and ***P<0.001 vs. control. 5azadC, 5-aza-2'-deoxycytidine; RT-q, reverse transcription-quantitative.

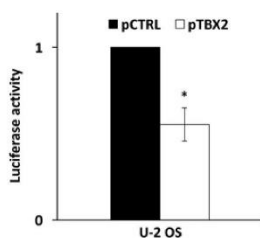


Figure 2. TBX2 ectopic expression induced relative Luciferase activity repression. Luciferase activity of the reporter plasmid was measured in U-2 OS cells transfected with the pT-REx-DEST30/TBX2-3XFlag (pTBX2) or with the control for 48 h. The data are presented as mean values from two independent experiments. Error bars represent the standard deviation and P-values were calculated using a Student's t-test. Non-treated cells were used as a reference control: *P<0.05. TBX2, T-box transcription factor 2.

expression could downregulate endogenous E6 expression in SiHa and Ca Ski cells transfected with pT-REx-DEST30/TBX2-3XFlag. As shown in Fig. 4A, TBX2 protein was readily

detected in transfected SiHa and Ca Ski cells. But any changes were observed in E6 protein (Fig. 4A) and mRNA (Fig. 4B) expression following ectopic TBX2 expression. In keeping with this observation, TBX2 unlikely downregulates endogenous E6 expression in cervical cancer cell lines.

TBX2 overexpression and 5azadC combinatory treatment does not enhance E6 repression. In order to rule out the possibility that TBX2 was involved in E6 repression following 5azadC treatment, TBX2 overexpression and 5azadC combinatory treatment were conducted. TBX2 overexpression has been confirmed by RT-qPCR in cells treated or not with 5azadC (Fig. S1).

Fig. 4C shows that E6 RNA relative expression was significantly decreased following 5azadC treatment in SiHa (P=0.03) and Ca Ski (P=0.001) cells. In contrast, no difference in E6 RNA relative expression was observed between the pT-REx-DEST30/TBX2-3XFlag-(pTBX2) and the pT-REx-DEST30/empty-(pCTRL) transfected cells, whether they were treated by 5azadC or not. Thus, overexpression of TBX2 did not enhance 5azadC-induced E6 repression.

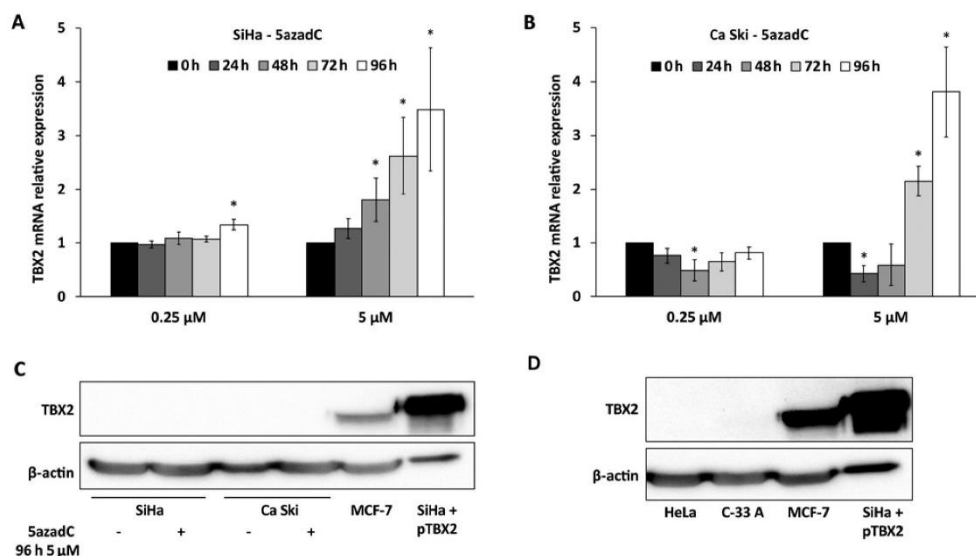


Figure 3. TBX2 mRNA detected in cervical cancer cell lines. TBX2 mRNA relative expression was analyzed using reverse-transcription-quantitative PCR in (A) SiHa cells and in (B) Ca Ski cells treated with 0.25 or 5 μM of 5azadC for a period of 24, 48, 72 or 96 h. TBX2 protein expression was analyzed using western blotting in (C) SiHa and Ca Ski cells treated with 5 μM of 5azadC during 96 h or the control and in (D) HeLa and C-33 A cervical cancer cell lines and in MCF-7 breast cancer cells or SiHa cells transfected with pT-REx-DEST30/TBX2-3XFlag (SiHa + pTBX2) that served as positive controls. The data are presented as mean values from at least three independent experiments. Error bars represent the standard deviation and P-values were calculated using a Student's t-test using non-treated cells as a reference control. *P<0.05. TBX2, T-box transcription factor 2; 5azadC, 5-aza-2'-deoxycytidine.

Discussion

In this study, we confirmed that the treatment of HPV16 positive cancer cell lines with 5azadC leads to the downregulation of E6 at both mRNA and protein levels (7-9). As expected, a p53 and p21 restoration was observed in treated cells confirming the functional loss of E6. Interestingly, the treatment of the tongue HPV16-positive cancer cell line UPCI:SCC090 with 5azadC also leads to the decrease of E6 and to the increase of p53 and p21 expression (Fig. S2). These results, in line with those obtained by Stich *et al* with two other head and neck cancer-derived cell lines (UM-SCC-47 and UM-SCC-104) (8), suggest that the effect of 5azadC treatment on E6 repression is independent of the tumor origin. Since 5azadC (decitabine) is already used to treat myelodysplastic syndromes (21) and acute myeloid leukemia (22) its usefulness in the treatment of HPV-associated cancers probably deserves to be addressed. In this line, Biktasova *et al* recently reported that the treatment of patients presenting HPV-positive head and neck cancers with 5-azacytidine, a 5azadC analogue, induced E6 and E7 RNA repression in their tumors accompanied by reactivation of the p53 pathway and apoptosis induction (23). This is the first study presenting the potential effects of a demethylating treatment against human HPV+ tumors in a clinical trial.

In SiHa cells, the effect of 5azadC on E6 RNA repression was not dependent on the concentration used since E6 downregulation was similar whatever the concentration was (0.25 μM or 5 μM) at each time point over 96 h. Similarly, Stich *et al* showed no dose effect of 5azadC treatment on

E6*/E7 mRNA downregulation even with a concentration as low as 100 nM (8). In contrast, a clear time-dependent effect of 5azadC treatment was observed on E6 downregulation with a maximum achieved at 96 h. This may reflect the 5azadC mechanism of action that requires successive cell divisions to passively demethylate DNA leading to the progressive upregulation of factor(s) involved in direct or indirect E6 repression. However, such a time dependent effect was not reproduced in Ca Ski cells treated with 5 μM of 5azadC. At this concentration, a cut-off effect was observed starting at 24 h of treatment, an observation consistent with a more sensitive phenotype of Ca Ski cells regarding 5azadC effects on E6 downregulation.

In order to highlight the mechanism involved in E6 downregulation following 5azadC treatment, we explored the role of the transcription factor TBX2. Indeed, Schneider and collaborators have recently proposed that TBX2 might decrease the HPV gene expression through LCR inhibition (11). Furthermore, they mapped the minimal sequence required for TBX2 inhibition from nt 7564 to nt 7756 on the LCR. A first series of experiments permitted to confirm in an *in vitro* assay that the overexpression of TBX2 downregulated HPV16 LCR activity.

In both SiHa and Ca Ski cells, TBX2 mRNA expression was up-regulated following 5azadC exposure in a time and dose dependent manner. However, this regulation pattern differed substantially from the one observed for E6. This was especially true for the lowest concentration of 5azadC that had weak effects on TBX2 mRNA upregulation while E6 mRNA was clearly downregulated. Surprisingly, the TBX2

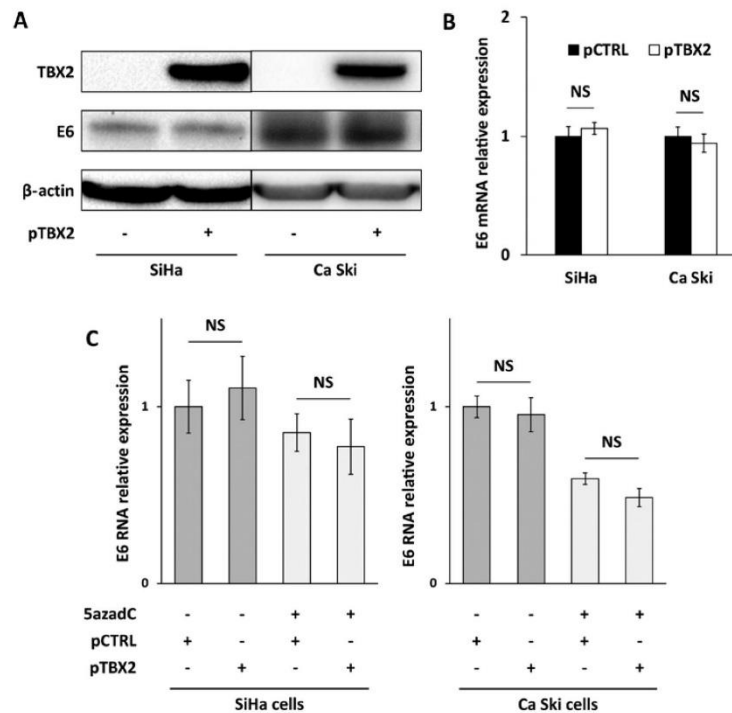


Figure 4. TBX2 is not involved in E6 repression *in vivo*. (A) E6 protein and (B) mRNA expression in SiHa and Ca Ski cells transfected for 48 h with pT-REx-DEST30/TBX2-3XFlag (pTBX2). The detection of TBX2 using western blotting confirmed the efficiency of transfection. E6 RNA expression in (C) SiHa and Ca Ski cells treated with 0.25 μ M of 5azadC and transfected with the pT-REx-DEST30/TBX2-3XFlag (pTBX2) or the pT-REx-DEST30/empty (pCTRL). The data are presented as mean values from three independent experiments. Error bars represent the standard deviation and P-values were calculated using a Student's t-test using non-treated cells as a reference control and a two-way ANOVA. TBX2, T-box transcription factor 2; 5azadC, 5-aza-2'-deoxycytidine; NS, not significant.

protein was undetectable in SiHa and Ca Ski cells as well as in two other cervical cancer cell lines infected by HPV18 (HeLa cells) or not infected by HPV (C-33 A cells). In contrast TBX2 protein was clearly detected in MCF-7 cells and in SiHa cells transfected with pT-REx-DEST30/TBX2-3XFlag that served as positive controls. While we can hypothesize that TBX2 protein is not expressed in cervical cancer we cannot rule out that its expression is very low in cervical cancer cells (under the limit of detection of the western blotting assay). Indeed, Schneider *et al* reported only a very faint band for TBX2 in their western blotting experiments performed with SiHa cell protein extracts. In the present study, TBX2 protein was not detected even after the treatment of cervical cancer cell lines with 5 μ M of 5azadC during 96 h, whereas the relative expression of the corresponding mRNA was increased more than 3-fold. Thus, we provide no evidence that the downregulation of E6 expression following 5azadC treatment is mediated through an increased expression of endogenous TBX2.

The increased expression of TBX2 mRNA transcripts without effective translation could be explained by the fact that Ca Ski and SiHa cells expressed at extremely low level TBX2 mRNA (700-fold less in Ca Ski and 90-fold less in SiHa cells), compared to TBX2 mRNA levels in MCF-7 cells as determined

by RT-qPCR (data not shown). These results are in line with the Human Protein Atlas data, that report no TBX2 mRNA expression in SiHa cells (0.0 TPM) while TBX2 is expressed at 62.0 TPM in MCF-7 cells). Another explanation could be the inhibition of translation due to mRNA sequestration in the nucleus, or due to miRNA interference.

Because TBX2 protein was shown to inhibit the activity of a cloned HPV16 LCR in overexpression experiments [(11) and present study], we wondered to assess whether it could also repress the endogenous HPV16 LCR in Ca Ski and SiHa cells. While TBX2 was readily detected in both transfected cell lines, no variation in E6 expression at both protein and mRNA expression was evidenced. Similarly, TBX2 overexpression in the presence of 5azadC did not enhance E6 repression that is solely due to 5azadC. These observations again reinforce the idea that TBX2 is unlikely involved in the regulation of E6 expression in both cell lines.

A limitation of the present study is the lack of TBX2 target assessment. No specific TBX2 target is described in cervical cancer cell models, in the literature. This is why TBX2 targets described in cell lines derived from other tissues [p21 (24), p14/p16 (25-27), p27 (28), PTEN (29) and NDRG1 (30)] were tested. However, none of these targets was significantly

repressed following TBX2 transfection. There is no clear explanation to this observation, but it has to be noted that p21, p27, p14/p16 are already deregulated by HPV oncogenes in cervical cancer cells.

Epigenetic regulation of HPV16 oncogene expression is very likely since the use of a DNA demethylating agent leads to E6 mRNA and protein repression. Indirect effects linked to the re-expression of miR-375, that targets HPV16 early transcripts has been well documented (8,9). By contrast, the involvement of TBX2 in E6 repression is very unlikely in SiHa and Ca Ski cells treated by 5azadC even if this transcription factor may exert inhibitory effect on HPV16 LCR cloned upstream a reporter gene. In particular, combinatory experiments conducted in the present study permitted to confirmed that TBX2 was not involved in E6 repression even in a context of demethylated DNA. Whether the structure of chromatin affect the accessibility of TBX2 to occupy its binding site in the native HPV16 LCR remains a challenging question.

Acknowledgements

The authors would like to thank Dr Schneider (University Medical Center of Johannes Gutenberg University Mainz, Mainz, Germany) for providing us with the pT-REx-DEST30/TBX2-3XFlag, Dr Caroline Demeret (Pasteur Institute, Paris, France) for providing us with pGL3-Luc-16LCR and Dr F Monnier (Centre Hospitalier Universitaire de Besançon, Besançon, France) for aiding in the statistical analysis of the present study. The authors would also like to thank Julie Durel and Anne Peigney (both, Université Bourgogne Franche-Comté, Université de Franche Comté, Besançon, France) for excellent technical assistance.

Funding

The present study was supported by research grants from La Ligue Contre le Cancer (grant no. CCIR-GE) and the Conseil Régional de Franche-Comté (grant no. 2016Y7570-2016Y7571). J. Perrard and A. Morel were recipients of a predoctoral scholarship from the Conseil Régional de Franche-Comté and K. Meznad was the recipient of a predoctoral scholarship from Ministère de l'Enseignement Supérieur et de la Recherche Scientifique.

Availability of data and materials

All data generated or analyzed during the present study are included in this published article.

Authors' contributions

JP, AM, ChM, VD, CC and JLP conceived and designed the experiments. JP, KM, CeM and PPB performed the experiments. JP, AB, SF, VD, DG and CC analyzed the data. JP, AB, DG, ChM and JLP wrote the manuscript.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Doorbar J, Egawa N, Griffin H, Kranjec C and Murakami I: Human papillomavirus molecular biology and disease association. *Rev Med Virol* 25 (Suppl 1): S2-S23, 2015.
2. Scheffner M, Werness BA, Huibregtse JM, Levine AJ and Howley PM: The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell* 63: 1129-1136, 1990.
3. Münger K, Scheffner M, Huibregtse JM and Howley PM: Interactions of HPV E6 and E7 oncoproteins with tumour suppressor gene products. *Cancer Surv* 12: 197-217, 1992.
4. Huh K, Zhou X, Hayakawa H, Cho JY, Libermann TA, Jin J, Harper JW and Mungler K: Human papillomavirus type 16 E7 oncoprotein associates with the cullin 2 ubiquitin ligase complex, which contributes to degradation of the retinoblastoma tumor suppressor. *J Virol* 81: 9737-9747, 2007.
5. zur Hausen H: Human papillomaviruses in the pathogenesis of anogenital cancer. *Virology* 184: 9-13, 1991.
6. Jacquin E, Baraquin A, Ramanah R, Carcopino X, Morel A, Valmary-Degano S, Bravo IG, de Sanjosé S, Riethmuller D, Mougins C and Prétet JL: Methylation of human papillomavirus Type 16 CpG sites at E2-binding site 1 (E2BS1), E2BS2, and the Sp1-binding site in cervical cancer samples as determined by high-resolution melting analysis-PCR. *J Clin Microbiol* 51: 3207-3215, 2013.
7. Zhang C, Deng Z, Pan X, Uehara T, Suzuki M and Xie M: Effects of methylation status of CpG sites within the HPV16 long control region on HPV16-positive head and neck cancer cells. *PLoS One* 10: e0141245, 2015.
8. Stich M, Ganss L, Puschhof J, Prigge ES, Reuschenbach M, Guterrez A, Vinokurova S and von Knebel Doeberitz M: 5-aza-2'-deoxycytidine (DAC) treatment downregulates the HPV E6 and E7 oncogene expression and blocks neoplastic growth of HPV-associated cancer cells. *Oncotarget* 8: 52104-52117, 2016.
9. Morel A, Bague A, Perrard J, Demeret C, Jacquin E, Guenat D, Mougins C and Prétet JL: 5azadC treatment upregulates miR-375 level and represses HPV16 E6 expression. *Oncotarget* 8: 46163-46176, 2017.
10. Jung HM, Phillips BL and Chan EK: miR-375 activates p21 and suppresses telomerase activity by coordinately regulating HPV E6/E7, E6AP, CIP2A, and 14-3-3 ζ . *Mol Cancer* 13: 80, 2014.
11. Schneider MA, Scheffer KD, Bund T, Boukhalouk F, Lambert C, Cotarelo C, Pflugfelder GO, Florin L and Spoden GA: The transcription factors TBX2 and TBX3 interact with human papillomavirus 16 (HPV16) L2 and repress the long control region of HPVs. *J Virol* 87: 4461-4474, 2013.
12. Wansleben S, Peres J, Hare S, Goding CR and Prince S: T-box transcription factors in cancer biology. *Biochim Biophys Acta* 1846: 380-391, 2014.
13. Abrahams A, Parker MI and Prince S: The T-box transcription factor Tbx2: Its role in development and possible implication in cancer. *IUBMB Life* 62: 92-102, 2010.
14. Du WL, Fang Q, Chen Y, Teng JW, Xiao YS, Xie P, Jin B and Wang JQ: Effect of silencing the T-Box transcription factor TBX2 in prostate cancer PC3 and LNCaP cells. *Mol Med Rep* 16: 6050-6058, 2017.
15. Khalil A, Dekmak B, Boulos F, Kantrowitz J, Spira A, Fujimoto J, Kadara H, El-Hachem N and Nemer G: Transcriptomic alterations in lung adenocarcinoma unveil new mechanisms targeted by the TBX2 subfamily of tumor suppressor genes. *Front Oncol* 8: 482, 2018.
16. Khalil AA, Sivakumar S, Lucas FAS, McDowell T, Lang W, Tabata K, Fujimoto J, Yatabe Y, Spira A, Scheet P, *et al*: TBX2 subfamily suppression in lung cancer pathogenesis: A high-potential marker for early detection. *Oncotarget* 8: 68230-68241, 2017.

17. Rani L, Mathur N, Gupta R, Gogia A, Kaur G, Dhanjal JK, Sundar D, Kumar L and Sharma A: Genome-wide DNA methylation profiling integrated with gene expression profiling identifies PAX9 as a novel prognostic marker in chronic lymphocytic leukemia. *Clin Epigenetics* 9: 57, 2017.
18. Farkas SA, Sorbe BG and Nilsson TK: Epigenetic changes as prognostic predictors in endometrial carcinomas. *Epigenetics* 12: 19-26, 2017.
19. Kandimalla R, van Tilborg AA, Kompier LC, Stumpel DJ, Stam RW, Bangma CH and Zwarthoff EC: Genome-wide analysis of CpG island methylation in bladder cancer identified TBX2, TBX3, GATA2, and ZIC4 as pTa-specific prognostic markers. *Eur Urol* 61: 1245-1256, 2012.
20. Livak KJ and Schmittgen TD: Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-(Delta Delta C(T))} method. *Methods* 25: 402-408, 2001.
21. Gangat N, Patnaik MM and Tefferi A: Myelodysplastic syndromes: Contemporary review and how we treat. *Am J Hematol* 91: 76-89, 2016.
22. Nieto M, Demolis P, Béhanzin E, Moreau A, Hudson I, Flores B, Stemplewski H, Salmonson T, Gisselbrecht C, Bowen D and Pignatti F: The European Medicines Agency review of decitabine (Dacogen) for the treatment of adult patients with acute myeloid leukemia: Summary of the scientific assessment of the committee for medicinal products for human use. *Oncologist* 21: 692-700, 2016.
23. Biktasova A, Hajek M, Sewell A, Gary C, Bellinger G, Deshpande HA, Bhatia A, Burtneß B, Judson B, Mehra S, *et al*: Demethylation therapy as a targeted treatment for human papillomavirus-associated head and neck cancer. *Clin Cancer Res* 23: 7276-7287, 2017.
24. Prince S, Carreira S, Vance KW, Abrahams A and Goding CR: Tbx2 directly represses the expression of the p21(WAF1) cyclin-dependent kinase inhibitor. *Cancer Res* 64: 1669-1674, 2004.
25. Jacobs JJ, Keblusek P, Robanus-Maandag E, Kristel P, Lingbeek M, Nederlof PM, van Welsem T, van de Vijver MJ, Koh EY, Daley GQ and van Lohuizen M: Senescence bypass screen identifies TBX2, which represses Cdkn2a (p19(ARF)) and is amplified in a subset of human breast cancers. *Nat Genet* 26: 291-299, 2000.
26. Vance KW, Carreira S, Brosch G and Goding CR: Tbx2 is overexpressed and plays an important role in maintaining proliferation and suppression of senescence in melanomas. *Cancer Res* 65: 2260-2268, 2005.
27. Harrelson Z, Kelly RG, Goldin SN, Gibson-Brown JJ, Bollag RJ, Silver LM and Papaioannou VE: Tbx2 is essential for patterning the atrioventricular canal and for morphogenesis of the outflow tract during heart development. *Development* 131: 5041-5052, 2004.
28. Lütke TH, Rudat C, Wojahn I, Weiss AC, Kleppa MJ, Kurz J, Farin HF, Moon A, Christoffels VM and Kispert A: Tbx2 and Tbx3 act downstream of shh to maintain canonical wnt signaling during branching morphogenesis of the murine lung. *Dev Cell* 39: 239-253, 2016.
29. Zhu B, Zhang M, Williams EM, Keller C, Mansoor A and Davie JK: TBX2 represses PTEN in rhabdomyosarcoma and skeletal muscle. *Oncogene* 35: 4212-4224, 2016.
30. Redmond KL, Crawford NT, Farmer H, D'Costa ZC, O'Brien GJ, Buckley NE, Kennedy RD, Johnston PG, Harkin DP and Mullan PB: T-box 2 represses NDRG1 through an EGFR1-dependent mechanism to drive the proliferation of breast cancer cells. *Oncogene* 29: 3252-3262, 2010.

Figure S1. TBX2 primers and anti-TBX2 antibody specificity and TBX2 transfection controls. TBX2 silencing by transfection with TBX2 siRNA (siTBX2) in MCF-7 confirmed using (A) western blotting and (B) reverse-transcription-quantitative PCR. (C) TBX2 protein expression in U-2 OS cells transfected for 48 h with the pT-REx-DEST30/TBX2-3XFlag (pTBX2). (D) TBX2 RNA expression in SiHa and Ca Ski cells treated with 0.25 μ M of 5azadC and transfected with pT-REx-DEST30/TBX2-3XFlag (pTBX2) or the pT-REx-DEST30/empty (pCTRL). The data are presented as mean values from three independent experiments. Error bars represent the standard deviation. TBX2, T-box transcription factor 2; si, small interfering; 5azadC, 5-aza-2'-deoxycytidine.

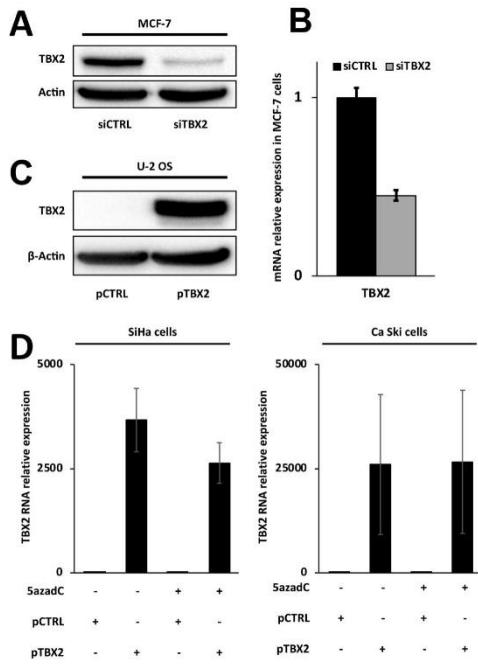


Figure S2. 5azadC treatment downregulates E6 expression in the UPCI:SCC090 HPV-positive tongue cancer cells. (A) E6 mRNA relative expression in UPCI:SCC090 cells treated with 250 nM of 5azadC for 96 h. (B) E6, p53, p21 and actin protein expression in the UPCI:SCC090 cells treated with 250 nM of 5azadC during 96 h. 5azadC, 5-aza-2'-deoxycytidine.

