THESE DE DOCTORAT DE L'ETABLISSEMENT

UNIVERSITE DE BOURGOGNE FRANCHE-COMTE

Préparée à l'UFR des Sciences médicales et Pharmaceutiques de Besançon

Ecole doctorale N°554

Environnement-Santé

Doctorat en Biochimie et Biologie moléculaire

par

Koceila MEZNAD

Interaction entre l'oncoprotéine E6 d'HPV16 et le métabolisme des ARN messagers

Thèse présentée et soutenue publiquement à Besançon, le 28 novembre 2018

Composition du jury :

Pr. Jean-Luc PRETET	Univ. Bourgogne Franche-Comté	Président
Dr. Murielle MASSON	IREBS, Ilkirch	Rapporteur
Dr. Cyril BOURGEOIS	ENS de Lyon	Rapporteur
Dr. Catherine-Laure TOMASETTO	IGBMC, Ilkirch	Examinatrice
Dr. Alain JUNG	Centre Paul Strauss, Strasbourg	Examinateur
Pr. Michaël BOYER-GUITTAUT	Univ. Bourgogne Franche-Comté	Examinateur
Pr. François AUBIN	Univ. Bourgogne Franche-Comté	Directeur de thèse
Dr. Aurélie BAGUET	Univ. Bourgogne Franche-Comté	Co-directeur de thèse

Remerciements

Je tiens à remercier les membres de mon jury de thèse :

Mme le Docteur Murielle Masson, rapporteur,

Mr le Docteur Cyril Bourgeois, rapporteur,

Mme le Docteur Catherine-Laure Tomasetto, examinatrice

Mr le Docteur Alain Jung, examinateur

Mr le Professeur Michaël Boyer-Guittaut, examinateur

Je vous prie de recevoir toute ma gratitude pour avoir consacré votre temps à l'analyse et la critique de mon travail. Je suis très honoré que mon travail ait été analysé par des experts tels que vous.

Je remercie également mon directeur de thèse, Mr le Professeur François Aubin, Praticien Hospitalier au CHRU de Besançon, pour avoir accepté de diriger mon travail ainsi que pour toute sa confiance.

Je tiens particulièrement à remercier Mme le Docteur Aurélie Baguet de m'avoir permis de mener ce projet, pour sa disponibilité et l'ensemble de ses conseils. Soit assurée de mon profond respect et de ma gratitude.

Mes remerciements vont aussi à Mr le Professeur Jean-Luc Prétet, directeur de l'EA3181, merci pour la confiance et d'avoir su me « re » motiver quand il le fallait, ainsi que pour avoir accepté de faire partie du jury.

Je remercie aussi, Mme le Professeur Christiane Mougin, ancienne directrice du laboratoire, pour m'avoir accueilli dans l'équipe.

A « mes » (nos) techniciennes préférées, Anne et Julie, merci pour votre aide ; plus la thèse avançait plus vous avez été embêtées par mes directives et mes protocoles, parfois, vraiment très INEXPLICITUS.

Aux membres de l'EA3181, d'hier et d'aujourd'hui, Sylvie, Isabelle, Sophie, Adeline, François, Adrien, Philippe et Jérôme, merci pour vos conseils et les bons moments passés aussi bien au laboratoire qu'en dehors.

Un grand merci aussi aux membres du consortium REISO-PAP, c'est un immense honneur que d'avoir pu travailler et apprendre à vos côtés.

Je remercie aussi l'Université de Franche-Comté, la SFR-FED4234, le LabEX LipSTIC, la Ligue contre le Cancer pour le soutien matériel et financier.

Merci à mes amis de Bejaia, de Normandie, de Franche-Comté et de plusieurs endroits du globe, pour les bons moments.

Enfin, je dirai merci à ma famille, mes parents, merci pour être ce que vous êtes, j'espère être à la hauteur de vos espérances.

Table des Matières

Liste des Figures 5
Liste des Tableaux 6
Liste des Abréviations 7
Introduction 11
1. Les papillomavirus humains 13
1.1. Classification des HPV au sein des papillomavirus 13
1.2. Classification des HPV en fonction de leur pathogénicité 14
1.3. Le génome viral 16
1.4. Infection virale 17
1.4.1. Les sites d'infection 17
1.4.2. Le cycle viral productif 18
1.4.2.1. Entrée du virus 18
1.4.2.2. Etablissement et maintenance du génome viral 19
1.4.2.3. Phase productive 20
1.4.2.4. Encapsidation et libération des virions 21
1.5. Clairance ou persistance de l'infection 21
1.5.1. Mécanismes de la clairance de l'infection et de latence 22
1.5.2. Echappement au système immunitaire et persistance de l'infection 22
1.6. Carcinogenèse associée aux HPV-HR 24
1.6.1. Processus de carcinogenèse au niveau du col de l'utérus 24
1.6.1.1. Les lésions précancéreuses 24
1.6.1.2. Progression de la carcinogenèse 25
1.6.2. Dépistage et prévention du cancer du col de l'utérus 26
1.6.2.1. Dépistage 27
1.6.2.2. Vaccination 27
1.6.3. Rôle des protéines E7 et E6 dans la carcinogenèse 28
1.6.3.1. Rôle de la protéine E7 des HPV-HR 28
- Stimulation de la prolifération cellulaire 29
- Instabilité chromosomique 30
- Inhibition de la mort cellulaire 31
- Régulation de l'expression génique et reprogrammation épigénétique 32
- Immunité 32
- Angiogenèse et invasion cellulaire 33

1.6.3.2. Rôle de la protéine E6 des HPV-HR 33
- Inhibition de la voie p53 34
- Inhibition de l'apoptose 35
- Immortalisation cellulaire 36
- Instabilité génétique et chromosomique 36
- Perte de l'adhérence et de la polarité cellulaire 37
- Signalisation, survie et croissance cellulaire 38
- Régulation de l'expression génique 39
- Immunité 40

2. Régulation de l'expression des gènes des HPV-HR 41
2.1. Régulation de la transcription d'HPV 41
2.1.1. Promoteur précoce 41
2.1.1.1. Structure de la LCR 41
2.1.1.2. Régulation par la protéine E2 43
2.1.1.3. Régulation épigénétique 44
2.1.2. Promoteur tardif 45
2.2. Régulation post-transcriptionnelle de l'expression des gènes viraux 46
2.2.1. Epissage de l'ARN et polyadénylation 47
2.2.1.1. L'épissage des ARN 47
2.2.1.2. Régulation de l'épissage alternatif 48
2.2.1.3. La polyadénylation 49
2.2.1.4. Epissage et polyadénylation des transcrits d'HPV16 50
2.2.1.4.1. Régulation de l'épissage et polyadénylation des transcrits d'HPV16 52
2.2.2. Régulation de la stabilité des transcrits viraux 54
2.2.3. Traduction des ARNm viraux 55
2.2.4. Rôle de l'épissage alternatif des ORF E6 et E7 57
3. Le complexe de jonctions des exons 61
3.1. Cycle de l'EJC 61
3.1.1. Composition de l'EJC 61
3.1.2. Assemblage de l'EJC 62
3.1.3. Remodelage de l'EJC 63
3.1.4. Désassemblage et recyclage de l'EJC 65
3.2. Dépôt de l'EJC au sein du transcriptome 66
3.3. Les fonctions de l'EJC 67
3.3.1. Régulation de l'épissage 67

3.3.2. Export nucléo-cytoplasmique des ARNm 68
3.3.3. Stimulation de la traduction 69
3.3.4. Rôle dans la surveillance des ARNm par le NMD 70
3.4. Le Nonsense-mediated mRNA decay 70
3.4.1. Le NMD dépendant de l'EJC 70
3.4.2. Signaux de dégradation par le NMD 72
3.4.3. Cas particuliers : le NMD indépendant de l'EJC 73
3.4.4. Rôle du NMD dans le contrôle de l'expression des gènes 74
3.4.5. Le NMD dans la défense antivirale 76
3.4.6. L'expression des facteurs de l'EJC et du NMD est altérée dans les cancers 77
Objectifs 79
Résultats 83
1. Rôle de l'isoforme E6*I 85
1.1. Résumé de la publication 1 85
2. Rôle de la protéine E6 dans la liaison aux ARN 123
2.1. Introduction 123
2.2. Méthodologie 123
2.3. Résultats 126
2.4. Conclusion 126
3. Implication de l'EJC et du NMD dans la régulation de l'expression des gènes E6 et E7 d'HPV16 127
3.1. Résumé de la publication 2 127
3.2. Résultats supplémentaires : expression des facteurs de l'EJC et du NMD dans les lignées cancéreuses HPV16 positives 168
3.2.1. Introduction 168
3.2.2. Méthodologie 168
3.2.3. Résultats 169
3.2.4. Conclusion 170
4. Régulation de l'épissage d'E6 : Identifier des mutations au niveau des ORF E6 et E7 ayant un impact sur l'épissage 171
4.1. Introduction 171
4.2. Méthodologie 171
4.3. Résultats 172
4.4. Conclusion 175
Discussion et perspectives 177

1. Fonctions spécifiques du gène E6 des HPV-HR.....179
2. La protéine E6, régulateur de l'expression des gènes au niveau post-transcriptionnel ? ...185
3. Régulation de l'expression des transcrits E6 et E7 par l'EJC et le NMD.....187
4. Régulation de l'épissage au sein de l'ORF E6 192
5. Conclusion générale.....193
Références Bibliographiques.....195
Annexe 1 : liste des transcrits dérégulés dans les cellules U-E6All et U-E6*I.....223
Annexe 2 : Liste des communications.....233

Liste des Figures

Figure 1. Classification des HPV et de quelques PV. 14
Figure 2. Prévalence des infections à HPV dans des cancers associés aux HPV. 15
Figure 3. Structure du génome d'HPV16. 16
Figure 4. L'infection par HPV et la zone de transformation. 18
Figure 5. Expression des gènes viraux durant le cycle productif des alphapapillomavirus. 20
Figure 6. Prévalence des lésions du col de l'utérus induites par les HPV-HR 24
Figure 7. Lésions précancéreuses et cancéreuses induites par les HPV au niveau du col de l'utérus. . 25
Figure 8. Progression de la carcinogenèse du col de l'utérus et principaux évènements moléculaires associés 26
Figure 9. La protéine E7 des HPV-HR 29
Figure 10. Rôle de la protéine E7 des HPV-HR. 29
Figure 11. La protéine E6 d'HPV16. 33
Figure 12. Rôle de l'oncoprotéine E6 HPV-HR. 34
Figure 13. Structure de la LCR d'HPV16. 42
Figure 14. Les sites canoniques d'épissage. 47
Figure 15. Le processus de l'épissage. 48
Figure 16. Le processus de l'épissage altérnatif. 48
Figure 17. La régulation de l'épissage par les protéines SR et hnRNP. 49
Figure 18. Epissage et polyadénylation alternatifs des ARNm d'HPV16. 51
Figure 19. Régulation de l'épissage de l'expression des transcrits précoces et tardifs d'HPV16. 52
Figure 20. Rôles potentiels des ARNm E6*I. 58
Figure 21. Structure et rôle de la protéine E6*I. 59
Figure 22. Structure de l'EJC. 63
Figure 23. Le cycle de l'EJC. 65
Figure 24. La fonction de l'EJC dans le NMD. 71
Figure 25. Les signaux de dégradation par le NMD. 72
Figure 26. Rôle du NMD dans les cellules. 74
Figure 27. Interaction entre la protéine E6 avec ses ARNm 126
Figure 28. Expression des facteurs de l'EJC et du NMD dans les cellules HPV positives. 169
Figure 29. Profils d'épissage des ARNm E6 dans les échantillons de cancers des VADS. 173
Figure 30. Le rôle de l'épissage dans l'ORF E6 des HPV-HR. 184
Figure 31. E6, un régulateur majeur de l'expression des gènes. 186
Figure 32. Interaction entre HPV16 et le NMD. 191

Liste des Tableaux

Tableau I. Classification des différents HPV carcinogènes par l'IARC. 15
Tableau II. Rôle des facteurs de transcription au niveau de la LCR d'HPV16. 43
Tableau III. Les composants de l'EJC. 64
Tableau IV. Liste des amorces utilisées pour quantifier l'expression des facteurs de l'EJC et du NMD. 169
Tableau V. Variations nucléotidiques au niveau de l'ORF E6 d'HPV16 dans les échantillons de cancers des VADS. 174
Tableau VI. Prédictions in silico de l'impact des mutations identifiées au sein de l'ORF E6, sur l'épissage. 175

Liste des Abréviations

5-azaC : 5-azacytidine
5-azadC : 5-deoxy-2-azacytidine
9G8 : Splicing Factor 9G8, appelé aussi SRSF7

A

A : Adénine
aa : acide aminé
Acinus : Apoptotic chromatin condensation inducer in the nucleus
ADN : Acide désoxyribonucléique
AIN : Anal Intraepithelial Neoplasia
Aly/REF : Aly/REF export factor
AP-1 : Activator Protein 1
ARN : Acide ribonucléique, ou RNA
ARNm: ARN messager, ou mRNA
ARNsb : ARN simple brin
ARNt : ARN de transfert
ASAP : Apoptosis and splicing-associated protein
ASC-US : Atypical Squamous Cell of Undetermined Significance
ATM : Ataxia Telangectasia Mutated
ATP : Adenosine TriPhosphate
ATR : Ataxia Telangectasia and Rad3 related

B

Bcl-2 : B-cell lymphoma 2
Bcl-xL : Bcl-2-like protein 1 long isoform
Brd4 : Bromodomain Containing 4
Brm : Brahma homolog

C

C : Cytosine
Ca^{2+} : Calcium
CBC : Cap binding complex
CBP/p300 : CREB Binding protein/Histone acetyltransferase p300
CCL : C-C motif chemokine ligand
CD : Cluster de différenciation
CD36 : CD36 molecule
CDC25A : cell division cycle 25A
Cdk2 : cyclin-dependant kinase 2
cEJC : canonical EJC
CHX : Cycloheximide
CIN : Cervical Intraepithelial Neoplasia
CMH : Complexe Majeur d'Histocompatibilité
MYC : V-Myc Avian Myelocytomatosis Viral Oncogene Homolog
CstF64 : Cleavage Stimulation Factor, 3 Pre-RNA, Subunit 2, 64kDa
C-ter : carboxy-terminale
CWC22 : CWC22 spliceosome associated protein homolog
CXCL : C-X-C motif chemokine ligand
C-x-x-C : C, Cystéine ; x, acide aminé

D

DEAD : Aspartate-Glutamate-Alanine-Aspartate
DECID : Decay-inducing complex
DLG1 : Disc Large protein 1
DNMT1 : DNA-Methyl Transferase 1
DTT : Dithiothreitol

E

E : Early
E2BS : E2-Binding Site
E6AP : E6 Associated Protein
E6Fl : E6 Full length
EDTA : Éthylène diamine tétraacétique
EGF : Epithelial Growth Factor
eIF : eucaryotic Initiation Factor
eIF4A3 :eukaryotic translation initiation factor 4A3
eIF4E :eukaryotic Initiation Factor 4E
EJC : Exon Junction Complex
eRF1 : eucaryotic Release Factor 1
ESE : Exonic Splicing Enhancer
ESS : Exonic splicing Silencer

F-G

FADD : Fas-associated death domain
FasR : Fas cell surface death receptor
FCU : Frottis Cervico-Utérin
G : Guanine
GPX1 : Glutathione Peroxidase 1
GTP : Guanosine TriPhosphate

H

hADA3 : human homologue of yeast alteration/deficiency in activation 3
HAT : Histone Acetyl Transferase
hnRNP : heterogeneous nuclear Ribonucleoprotein
HPV : Human Papillomavirus
HPV- : HPV négatif

HPV+ : HPV négatif
HPV-BR : HPV à bas risque oncogène
HPV-HR : HPV à haut risque oncogène
HRP : horseradish peroxidase
HSF3.1 : Human Splicing Finder 3.1
HSIL : High grade Squamous Intra-epithelial Lesion
HTLV-1 : Human T-cell Leukemia Virus type 1
HuR : Hu-Antigen R, appelée aussi ELAVL1

I-L

IFN : Interféron
IL : Interleukin
IRF : Interferon Responsive Element
ISE : Intronic Splicing Enhancer
ISS : Intronic Splicing Silencer
Kb : Kilo base
L : Late
LCR : Long Control Region
LSIL : Low grade Squamous Intra-epithelial Lesion
LxxLL : L, Leucine ; x, acide aminé

M

m : Masse
MAGI : membrane associated guanylate kinase interacting protein-like
MAGOH : Mago homolog, EJC core component
MAPK : Mitogen-activated protein kinase
MgOAc : magnesium acetate
miRNA : micro-RNA, ou microARN
MLN51 : Metastatic Lymph Node 51
mRNP : messenger Ribonucleoprotein
mTOR : mamalian Target Of Rapamycin

N

Na deoxycholate : sodium deoxycholate
NaCl : Chlorure de sodium
ncEJC : non canonical EJC
NFI : Nuclear Factor I
NFI-X : Nuclear Factor I X-type
NFX1 : Nuclear transcription factor, X-box binding 1
NMD : Nonsense-mediated mRNA Decay
Noxa : PMAIP1, phorbol-12-myristate-13-acetate-induced protein 1
NP-40 : Nonidet P-40
N-ter : amino-terminale
NXF1 : Nuclear RNA export factor 1
NXT1 : Nuclear transport factor 2 like export factor 1

O

ORF : Open Reading Frame
ORI : Origine de réplication du génome viral

P

p600 : 600 kDa retinoblastoma protein-associated factor, aussi appelée UBR4
pA : site poly(A)
PABP : Poly(A) Binding Protein
PABP4 : Poly(A) Binding Protein cytoplasmic 4
pAE : Early pA
pAL : late polyadenylation site
PAP : poly(A) polymérase
PATJ : PALS1-associated TJ protein
PBM : PDZ Binding Motif
P-bodies : Processing bodies ou corps GW
PDZ : PSD-95/DLG/ZO-1 domain
PeIN : Penile Intraepithelial Neoplasia
pH : potential hydrogène
PI3K/AKT : Phosphatidylinositol-4,5-Bisphosphate 3-Kinase/AKT serine/threonine-protein kinase
PIC : pre-initiation complex
Pinin : pinin, desmosome associated protein
PML : Promyelocytic leukemia
polyA+ : poly A positif,
pRb : protéine Retinoblastome
pré-ARNm : précurseur de l'ARNm
PTB : Polypyrimidin Tract Binding Protein
PTC : Premature Termination Codon
PTEN : phosphatase and tensin homolog
PV : Papillomavirus
PVDF : Polyfluorure de vinylidène
PVX : Potato Virus X
PYM1 : PYM homolog 1, exon junction complex associated factor

R

RAD51 : RAD51 recombinase
RBP : RNA binding protein
RIPA : Radioimmunoprecipitation assay
RNA-IP : RNA-Immunoprecipitation
RNA-seq : RNA-sequencing
RNPS1 : RNA Binding Protein With Serine Rich Domain 1
ROS : Reactive-Oxygen Species
RPA : Replication Protein A

RSV : Rous Sarcoma Virus
RT-PCR : Retrotranscription-Polymerase chain reaction
RT-qPCR : RT- quantitative PCR

S

SA : Splice Acceptor site
Sam68 : KH RNA binding domain containing, signal transduction associated 2
SC35 : Splicing Component, 35 KDa
SD : Splice Donnor site
SDS : sodium dodecyl sulfate
SDS-PAGE : sodium dodecyl sulfate–polyacrylamide gel electrophoresis
SF2/ASF : Splicing Factor 2/Alternative-Splicing Factor 1
siRNA : small interfering RNA
SKAR : S6K1 Aly/REF-like target
SMD : Staufen-mediated mRNA Decay
SMG6 : human suppressor with morpho-genetic effect on genitalia 6
snRNA U : U-rich small nuclear RNA
snRNP : small nuclear Ribonucleoprotein
SOD1 : Superoxide Dismutase 1
SP1 : Specificity Protein 1
SR : Serine Argenine Rich
SRp20 : Splicing Factor, Arginine/Serine-Rich, 20-KD
SRSF1 : Serine and arginine rich splicing factor 1

T

T : Thymine
TBE : Tris Borate EDTA
TERT : Telomerase Reverse transcriptase
TFIID : Transcription Initiation Factor TFIID
TGF-β : Transforming Growth Factor bêta
TIP60 : Tat interacting protein, 60kDa
TLR : Toll-Like Receptor
TNF-α : Tumor Necrosis Factor alpha
TRAIL : Tumor necrosis factor (ligand) superfamily, member 10
TREX : TRanscription and Export
TSS : Transcription Start Site

U

U : Uracile
U2AF35 : U2 small nuclear RNA auxiliary factor 1
U2AF65 : U2 small nuclear RNA auxiliary factor 2
UAP56 : U2AF65-associated protein 56
uORF : upstream ORF
UPF3a : Up-Frameshift Suppressor 3 Homolog A
UTR : Untranslated Region

V

v : Volume
VADS : Voies aéro-digestives supérieures
VaIN : Vaginal Intraepithelial Neoplasia
VIN : Vulvar Intraepithelial Neoplasia

X-Z

XRCC1 : X-ray Repair Cross Complementing 1
Y14 : RNA-binding protein Y14, ou RNA binding motif protein 8A
YY1 : Yin Yang 1
ZD : Zinc Domain

Introduction

1. Les papillomavirus humains

1.1. Classification des HPV au sein des papillomavirus

Les papillomavirus (PV) sont des virus non enveloppés, possédant un génome circulaire à ADN double brin qui infectent la peau et les muqueuses. Ils appartiennent à la famille des *Papillomaviridae* qui comprend environ 400 types de PV (Van Doorslaer *et al*., 2017). Au sein des *Papillomaviridae*, la structure des génomes viraux est bien conservée. En effet, le génome viral est composé de trois régions : une région de régulation et deux régions codant des gènes précoces E (Early) et tardifs L (Late) (Van Doorslaer, 2013). La région tardive code les protéines de capside dont L1, qui est très conservée. Les PV sont regroupés au sein de genres et espèces lorsqu'ils présentent des taux d'identité de séquences L1 de 60 à 70% et de 71 à 89% respectivement (Bernard *et al*., 2010; de Villiers, 2013; de Villiers *et al*., 2004). Par ailleurs, au-delà de 90% d'identité de la séquence L1, les PV sont considérés comme des sous-types. Les genres sont identifiés par une lettre grecque suivie du suffixe papillomavirus, tandis que les espèces sont identifiées par un chiffre arabe. Ainsi, HPV16 (Human papillomavirus 16) appartient au genre alphapapillomavirus de l'espèce 9 (Fig. 1), les HPV étant les PV qui infectent spécifiquement l'Homme. De manière intéressante, les genres les plus fournis en types de PV sont : alphapapillomavirus, bêtapapillomavirus et gammapapillomavirus. De la même façon, plus de 90% des HPV font partie de ces trois genres. La majorité des HPV des genres bêta et gamma infectent essentiellement la peau, tandis que le genre alpha comprend des HPV qui infectent la peau ou les muqueuses.

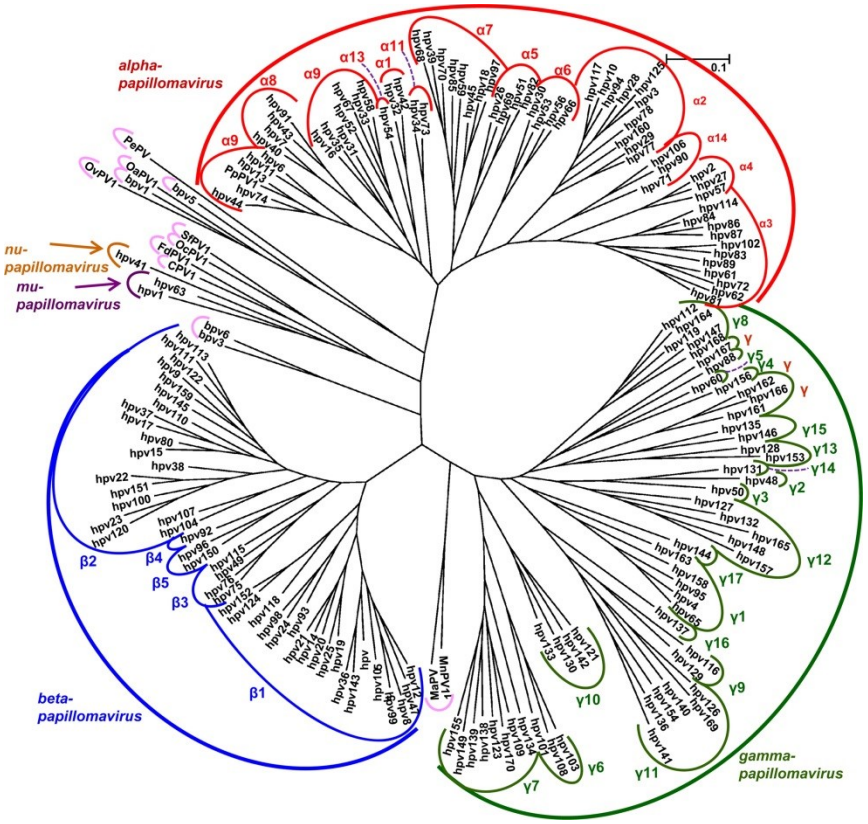


Figure 1. Classification des HPV et de quelques PV. (Adaptée de De Villiers, 2013).

1.2. Classification des HPV en fonction de leur pathogénicité

L'infection par HPV est assez fréquente chez les individus, notamment au niveau de la peau où la transmission se fait par des contacts directs. La majorité de ces infections, par des HPV des genres beta et gamma en particulier, sont présentes sous une forme asymptomatique. Toutefois, les HPV peuvent induire l'apparition de lésions bénignes au niveau de la peau (verrues) ou sur les muqueuses (condylomes ou verrues génitales) pouvant évoluer en lésions cancéreuses (Bravo et Félez-Sánchez, 2015). Ainsi, sur la base de leur pouvoir carcinogène pour l'Homme, les HPV ont été classés par l'IARC (International Agency for Research on Cancer) en plusieurs catégories (Tableau I). Les HPV carcinogènes (types 16, 18, 31, 33 35, 39, 45, 51, 52, 56, 58, 59) appartenant au groupe 1 ont des effets cancérigènes pour l'Homme qui ont été clairement démontrés par des études épidémiologiques et sur la base d'expérimentations animales. Ce groupe de virus est appelé HPV-HR (à haut risque oncogène) en opposition aux HPV-BR (à bas risque oncogène) classés dans le groupe 3. Enfin, les HPV des groupes 2A et 2B sont considérés comme probablement ou possiblement carcinogènes (Tableau 1).

Tableau I. Classification des différents HPV carcinogènes par l'IARC. (Adapté de Bravo et Félez-Sánchez, 2015).

Groupe 1 Carcinogène pour l'Homme	**Groupe 2A** Probablement carcinogène pour l'Homme	**Groupe 2B** Possible carcinogène pour l'Homme	**Groupe 3** Non classifiable	**Groupe 4** Probablement non carcinogène pour l'Homme
Evidence suffisante de la carcinogénicité chez l'Homme et dans des modèles animaux	Certitude/évidence limité quant à leur carcinogénicité pour l'Homme et évidence suffisante pour leur carcinogénicité dans des modèles animaux	Certitude/évidence limité quant à leur carcinogénicité pour l'Homme et preuves insuffisantes pour leur carcinogénicité dans des modèles animaux	Pas d'évidence adéquate pour une carcinogénicité chez l'Homme et dans des modèles animaux	Evidence suggerant l'absence d'effets carcinogène chez l'Homme et dans des modèles animaux
111 agents, dont 8 agents biologiques	**65 agents, dont 3 agents biologiques**	**274 agents, dont 6 agents biologiques**	**504 agents, dont 5 agents biologiques**	**1 agent (pas d'agent biologique)**
- HPV de types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58 et 59	- HPV de type 68	- HPV de types 26, 53, 66, 67, 70, 73 et 82 - HPV de types 5 et 8 (chez les patients atteints d'épidermodysplasie verruciforme) - HPV de types 30, 34, 69, 85, 97 (Classés par analolgie phylo-génétique avec les types d'alphapapillomavirus classés dans le groupe 1)	- HPV du genre beta (sauf les types 5 et 8) et du genre gamma - HPV de types 6 et 11	

Les HPV-HR font tous partie du genre alpha et possèdent un tropisme pour les muqueuses, notamment les muqueuses ano-génitales et orales où ils sont transmis par voie sexuelle. Ainsi, ils sont responsables de l'apparition de 99% et 90% des cas de cancers du col de l'utérus et du canal anal, respectivement (Fig. 2). Ils sont aussi responsables, dans une moindre mesure, des cancers du vagin, de la vulve, du pénis et de l'oropharynx (Crow, 2012; De Vuyst *et al*., 2009; Forman *et al*., 2012). En particulier, les HPV de types 16 et 18 sont les plus retrouvés dans les cancers, et sont responsables de plus de 60% des cancers attribuables aux HPV (Fig. 2).

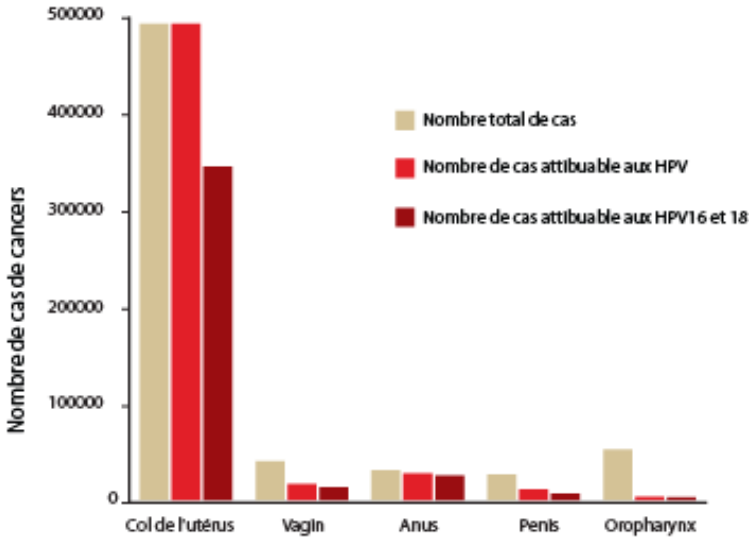


Figure 2. Prévalence des infections à HPV dans des cancers associés aux HPV. Nombre de cancers totaux, attribuable aux HPV ou aux HPV 16 et 18 dans le monde durant l'année 2012 dans les sites indiqués. **(Adaptée de Crow, 2012)**.

Par ailleurs, des HPV du genre bêta, notamment les types 5 et 8, sont associés à l'apparition de cancers non-mélaniques de la peau chez les patients atteints d'épidermodysplasie verruciforme. Ces patients, présentant une susceptibilité aux infections par les HPV, développent des lésions de la peau telles que des verrues et des cancers au niveau des zones

photo-exposées. De plus, l'infection par des HPV du genre bêta peut être associée à l'apparition de cancers non-mélaniques de la peau chez des patients immunodéprimés (Tommasino, 2017). Ainsi, à l'inverse des HPV-HR (genre alpha) dont les effets carcinogènes sont avérés, le rôle des bêtapapillomavirus dans les cancers cutanés est moins établi, et le processus de carcinogenèse dépend d'autres facteurs, notamment la vulnérabilité des patients aux infections par ces virus. Globalement, les HPV-HR sont responsables de l'apparition d'environ 5% des cancers dans le monde (Forman *et al*., 2012). Ces cancers sont majoritairement des carcinomes épidermoïdes (issus de cellules de l'épithélium épidermoïde ou malpighien), tandis qu'on retrouve aussi des atteintes des épithéliums glandulaires (adénocarcinome), voir des formes mixtes.

1.3. Le génome viral

Les génomes des HPV-HR, tel celui d'HPV16, sont circulaires, comportent une région non codante LCR (Long Control Region) et 8 ORF (Open Reading Frame) ou gènes (Fig. 3). La LCR contient l'origine de réplication, des éléments de régulation transcriptionnelle, dont le promoteur p97 (pour HPV16), permettant l'expression des gènes précoces (E1, E2, E4, E5, E6 et E7). L'expression des gènes L1 et L2 est régulée par la transcription à partir du promoteur tardif p670 (pour HPV16), qui permet aussi l'expression d'E1, E2, E4 et E5.

Les gènes précoces codent les protéines non-structurales qui sont essentielles à la réplication de l'ADN viral (E1 et E2), ainsi que des protéines « accessoires » permettant de contrôler/détourner le cycle de division de la cellule hôte au profit du virus (E4, E5, E6, E7). Les gènes L1 et L2 codent les protéines majeures et mineures de la capside respectivement. De manière intéressante, au cours du cycle viral, le génome des HPV est sous forme d'épisome extrachromosomique alors que lors de la carcinogenèse, on retrouve le plus souvent un génome linéaire et intégré à l'ADN cellulaire (Pett et Coleman, 2007).

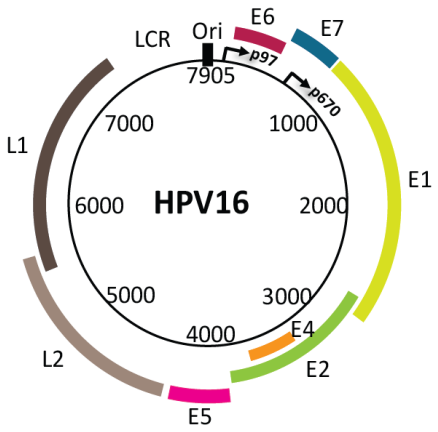


Figure 3. Structure du génome d'HPV16.

1.4. Infection virale

1.4.1. Les sites d'infection

Les HPV infectent principalement les épithéliums épidermoïdes pluristratifiés de la peau et des muqueuses, au sein desquels il peut effectuer un cycle viral productif. Le virus infecte des cellules permissives, qui ont conservé des propriétés de divisions cellulaires. Celles-ci pouvant être des cellules de la couche basale de l'épithélium épidermoïde, des cellules réserves ou souches (Fig. 4). Le cycle viral complet des HPV est dépendant de la différenciation cellulaire ayant lieu au cours de la migration des cellules de la couche basale jusqu'à la couche superficielle la plus différenciée. L'expression des gènes viraux est ainsi coordonnée avec le programme de différenciation cellulaire. On retrouve des épithéliums pavimenteux pluristratifiés dans tous les organes infectés par les HPV, au niveau oral, ano-génital et de la peau. D'ailleurs, la majorité des cancers HPV-induits sont des carcinomes épidermoïdes, issus d'épithéliums malpighiens/pavimenteux. L'accès du virus aux cellules basales de cet épithélium peut se faire grâce à des microlésions, des événements courants au niveau de la peau, mais beaucoup plus rare au niveau des muqueuses, notamment au niveau du col de l'utérus (Fig. 4).

D'autres types de cellules, plus accessibles à une infection par HPV, ont été identifiés au niveau du col de l'utérus. Il s'agit des cellules de l'épithélium glandulaire monostratifié et des cellules cuboïdes situées à la jonction épidermo-glandulaire (Herfs *et al*., 2012) (Fig. 4). Cette jonction est située dans la zone de transformation, dans laquelle se développent la plupart des cancers du col de l'utérus. De façon intéressante, on retrouve ce type de jonction d'épithélium dans le col de l'utérus et dans le canal anal, là où la proportion de cancers induit par HPV est la plus élevée. En outre, l'extrême différence au niveau de la prévalence du cancer du col de l'utérus en comparaison aux autres cancers induits par HPV, suggère l'extrême fragilité de ces jonctions favorisant le développement de cancers (Herfs *et al*., 2017). Les cellules cuboïdes seraient primitives et pourraient se différencier en cellules « réserves », cellules responsables du maintien et du renouvèlement des épithéliums malpighien et glandulaire, et ne supporteraient pas de cycle viral complet (Doorbar, 2016).

Bien que ces cellules cuboïdes soient plus accessibles à une infection par les HPV par rapport aux cellules de l'épithélium malpighien, leur infection par les HPV-HR supporterait davantage une évolution de lésions précancéreuses due aux HPV et aboutissant à des adénocarcinomes, cancers dérivés d'épithéliums glandulaires.

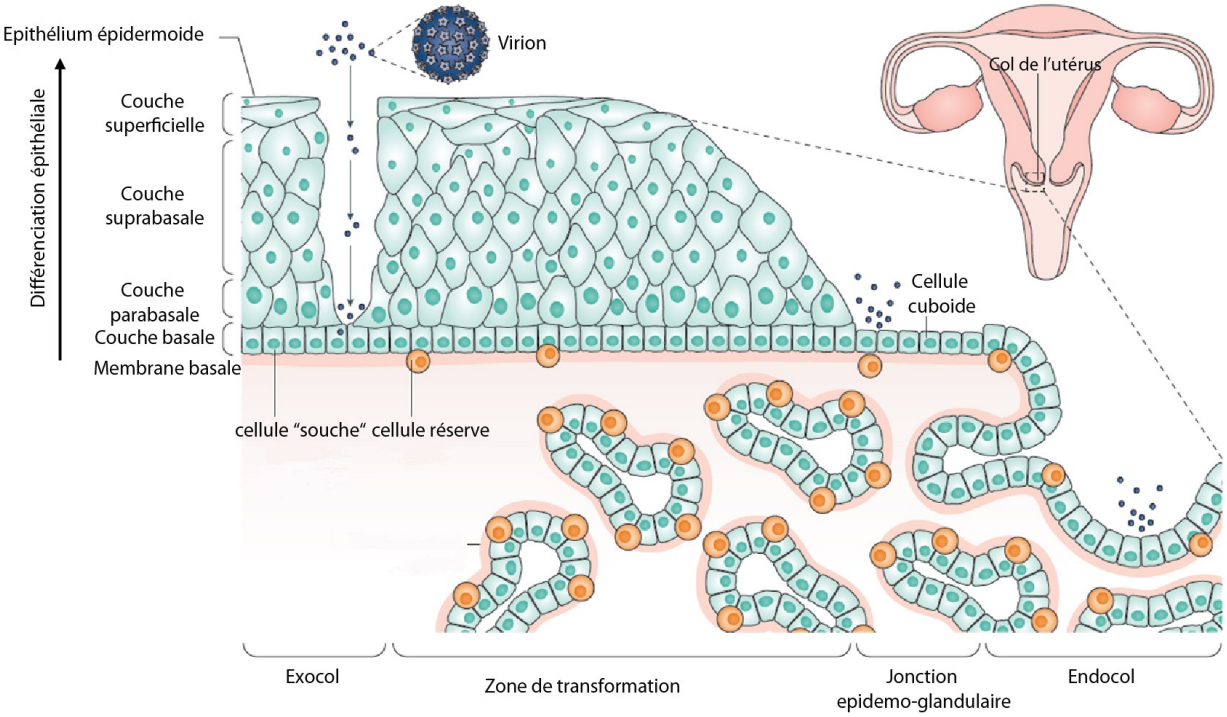


Figure 4. L'infection par HPV et la zone de transformation. (Adaptée de Schiffman *et al.*, 2016).

1.4.2. Le cycle viral productif

Au sein des épithéliums pluristratifiés (exocol, peau), les HPV infectent d'abord les cellules basales (cellules ''souches''), les seules à effectuer la mitose permettant l'entrée dans le noyau et l'amplification du génome viral (Fig. 5). S'en suit l'amplification du génome viral, puis la synthèse des protéines de la capside dans les cellules différenciées aboutissant finalement à la libération des virions lors de la desquamation des cellules apicales. Le modèle le plus étudié est celui d'HPV16, nous allons donc principalement décrire son cycle viral.

1.4.2.1. Entrée du virus

Le modèle le plus illustré et connu de l'infection par les HPV est l'accès direct aux cellules cuboïdes et *via* une microlésion au sein de l'épithélium pavimenteux pluristratifié permettant l'accès aux cellules basales.

Dans un premier temps, HPV16 se lie aux héparanes sulfates au niveau de la matrice extracellulaire ou des cellules basales *via* la protéine L1. Celle-ci subit ensuite des changements de conformation lui permettant de se fixer à son récepteur. Bien qu'il ait été montré que le récepteur de l'EGF (Epithelial Growth Factor Receptor) ou l'intégrine α6 joue un rôle dans l'internalisation du virus, aucun d'entre eux n'est indispensable (Doorbar *et al.*, 2012; Graham, 2017). Le récepteur d'HPV16 reste donc méconnu. En se fixant à ces

différentes protéines et à son récepteur, HPV16 est internalisé par un mécanisme de macropinocytose indépendant des clathrines ou des cavéolines. La fusion entre la vésicule internalisée et le lysosome dont le pH acide dénature la protéine L1 permet la libération de l'ADN viral. Toutefois, l'ADN viral reste lié à la protéine L2 qui permet sa sortie du lysosome *via* leur transfert vers le compartiment trans-golgien. Le complexe L2/ADN viral est ensuite transféré vers le réticulum endoplasmique où il reste protégé de la digestion par les nucléases cytoplasmiques. C'est pendant la mitose que le complexe L2/ADN viral est relargué, à l'intérieur d'une vésicule assurant sa protection, dans le cytoplasme où la protéine L2 interagit avec les microtubules. Enfin, le complexe L2/ADN viral intègre le noyau à la fin de la mitose et avant la formation de la membrane nucléaire.

1.4.2.2. Etablissement et maintenance du génome viral

L'établissement de l'infection se fait grâce à la division de cellules hôtes, permissives, où le génome viral est amplifié à environ 200 copies par cellule (Doorbar *et al*., 2012). C'est notamment à la périphérie des corps nucléaires PML que le génome viral est répliqué par la machinerie cellulaire. Les corps nucléaires PML permettent aussi la transcription des gènes précoces du virus, assurant ainsi la production des protéines E1 et E2, qui sont essentielles à la réplication du génome viral (Bienkowska-Haba *et al*., 2017; Day *et al*., 2004).

E1, qui possède une activité hélicase, se fixe sur l'ORI (origine de réplication du génome viral), ouvre la double hélice et permet la formation de la fourche de réplication. De plus, E1 permet le recrutement de plusieurs protéines cellulaires, notamment les ADN polymérases alpha et delta, la topoisomérase I et la RPA (Replication Protein A), qui assurent la synthèse de l'ADN viral. Par ailleurs, E1 est recruté sur l'ORI grâce à son interaction avec E2, laquelle se fixe au niveau de ses sites de liaison E2BS (E2-Binding Site) situés de part et d'autre de l'ORI au sein de la LCR (Bergvall *et al*., 2013; McBride, 2013). *Via* son interaction avec la protéine Brd4 (Bromodomain Containing 4), E2 recrute les épisomes viraux sur les chromosomes mitotiques, ce qui favorise leur séparation et leur ségrégation au cours des mitoses (Doorbar *et al*., 2012; McBride, 2013). Les génomes viraux sont ainsi répartis et maintenus dans les cellules filles.

Le génome viral est tout d'abord maintenu avec un faible nombre de copies dans les cellules (10-200 copies par cellule), *via* sa réplication en même temps que l'ADN selon un mode semi-conservatif bidirectionnel (Doorbar, 2005). La division des cellules basales permet donc la maintenance du génome d'HPV, alors que l'accomplissement de l'infection par la

production de virions est effectué dans les couches supérieures, au sein des cellules en cours de différenciation.

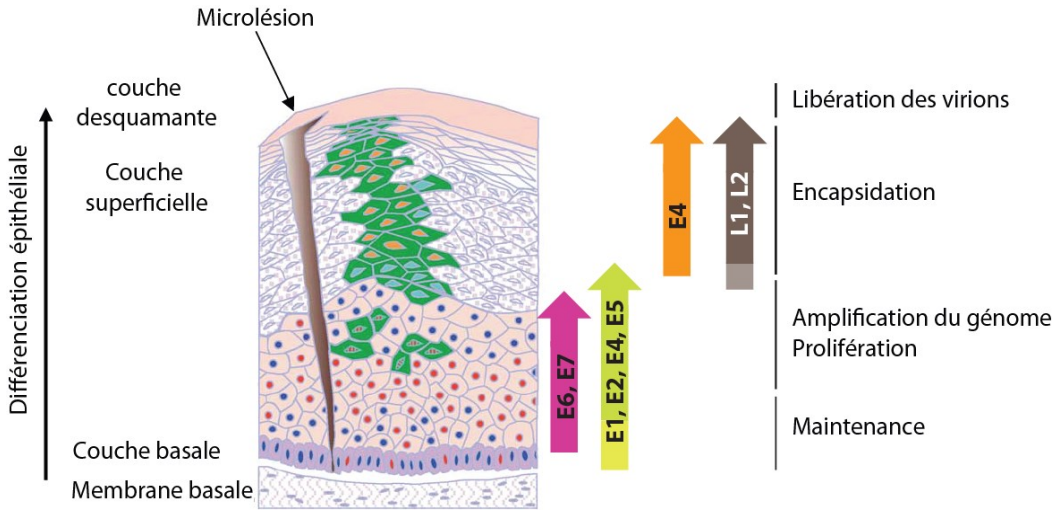


Figure 5. Expression des gènes viraux durant le cycle productif des alphapapillomavirus. (Adaptée de Doorbar, 2005).

1.4.2.3. Phase productive

Quand les cellules épithéliales débutent le processus de différenciation, les HPV enclenchent la phase productive de leur cycle, mettant en place un programme entièrement dédié à la réplication de l'ADN viral d'une manière très efficace. (i) L'activation du promoteur tardif d'HPV16 p670, grâce à la différenciation au niveau des cellules de la couche intermédiaire, permet d'accroître l'expression des protéines virales précoces notamment E1 et E2. (ii) Les cellules sont orientées vers le cycle cellulaire par les protéines virales E7, E6 et E5 (Doorbar, 2005; Doorbar *et al*., 2012).

Une expression élevée des protéines E1 et E2 permet l'amplification du taux de réplication de l'ADN viral. A cette étape, la réplication du génome viral est effectuée selon un mode en « cercle roulant », plus efficace dans l'amplification d'ADN circulaire comparativement au mode bidirectionnel. De plus, la réplication de l'ADN viral a lieu dans la phase G2, une fois le génome cellulaire répliqué, ce qui permet de dédier toute la machinerie cellulaire au profit de la réplication du génome viral (Doorbar *et al*., 2012; Graham, 2017; McBride, 2017). L'entrée des cellules hôtes dans le cycle cellulaire est induite par E7. En effet, E7 séquestre les protéines « suppresseur de tumeur » de la famille pRb (protéine Rétinoblastome) et permet la libération des facteurs E2F stimulant la transition G1-S du cycle cellulaire. Le cycle cellulaire induit par E7 présente une phase G2 anormalement longue, ayant subi vraisemblablement un arrêt (Doorbar *et al*., 2012; Graham, 2017; McBride, 2017). Cet arrêt

est attribué à la protéine E4 et ce, pour favoriser l'amplification du génome viral (Graham, 2017). La protéine E6, altère la différenciation cellulaire et inhibe l'apoptose, collaborant au maintien de la prolifération cellulaire. Pour amplifier son génome, HPV16 recrute aussi des protéines de la voie DDR (DNA Damage Repair pathway) de réparation des dommages à l'ADN, notamment ATM (Aataxia Telangectasia Mutated) et ATR (Ataxia Telangectasia and Rad3 related).

Au terme de cette étape, le nombre de copies de génome viral est amplifié considérablement (jusqu'à 1000 copies par cellule), préparant ainsi à la production du plus grand nombre possible de virions.

1.4.2.4. Encapsidation et libération des virions

Dans les couches épithéliales les plus superficielles et les plus différenciées, l'expression des gènes E6, E7 et E5 tend à diminuer au profit de l'expression de L1 et L2, ce qui arrête alors la prolifération cellulaire. Les protéines L1 et L2 exprimées sont adressées au noyau où aura lieu l'encapsidation des génomes viraux. La protéine E4 est aussi exprimée, assurant le réarrangement du cytosquelette, pour fragiliser les cellules et favoriser la libération des virions (Doorbar *et al*., 2012; Graham, 2017). Les virions sont libérés par le processus de desquamation des cellules, sans lyser les cellules, évitant ainsi d'activer le système immunitaire. Les virions, protégés dans les cellules desquamées sont donc libérés dans le canal vaginal où ils peuvent se transmettre par contacts sexuels ou réinfecter les tissus environnants.

1.5. Clairance ou persistance de l'infection

Les infections par les HPV, au niveau du col de l'utérus, peuvent être éliminées par le système immunitaire ou persister et induire un risque de développer des lésions précancéreuses et cancéreuses. La réaction immunitaire dirigée contre les HPV permet la clairance de la plupart des infections (~90%), au bout de deux ans (Shulzhenko *et al*., 2014). Cette clairance correspondant à l'absence de détection d'un HPV après l'infection, peut être liée à l'élimination complète du virus ou sa diminution à des taux indécelables par les méthodes utilisées actuellement en diagnostic. En effet, certaines infections persisteraient sous une forme latente, indétectable par le système immunitaire et pourraient se réactiver plus tard suite à une immunosuppression. Enfin, environ 10% des infections échappent au système immunitaire et persistent. La persistance étant définie comme la détection d'une infection (d'ADN viral) à au moins 12 mois d'intervalle.

L'élimination ou la persistance de l'infection par un HPV serait dépendante entre autres : (i) de l'efficacité du système immunitaire des patients à mettre en place une réaction efficace contre le virus et (ii) du génotype viral et notamment de la capacité du virus à mettre en place des mécanismes d'échappement.

1.5.1. Mécanismes de la clairance de l'infection et de latence

De façon intéressante, les patientes éliminant les infections virales ont tendance à développer des réponses immunitaires dirigées contre les protéines E2, E6 et E7 d'HPV16 (Dillon *et al*., 2007; Nakagawa *et al*., 2000). De plus, l'infiltration de lymphocytes T CD4+ (Monnier-Benoit *et al*., 2006) et cytotoxiques actifs T CD8+ Granzyme B+ au niveau de l'épithélium est plus importante chez les patientes réussissant à éliminer l'infection par rapport à celles où l'infection persiste (Woo *et al*., 2010). La clairance semble aussi être plus efficace chez les patientes ayant démontré une séroconversion et produisant des anticorps qui empêcheraient la réinfection (Foresta *et al*., 2015). Par ailleurs, Ryser et collaborateurs suggèrent que la stochasticité de la dynamique des cellules basales pourrait aussi participer à la clairance des infections, particulièrement chez les patients (es) immunocompétents (es) (Ryser *et al*., 2015).

L'existence d'une phase latente des infections HPV étant très discutée, quelques observations suggèrent cette possibilité. Chez le lapin, la détection, en l'absence d'infection apparente, de l'ADN du *Rabbit oral PapillomaVirus* en faibles copies et sans activité transcriptionnelle suggère sa présence sous un état latent (Maglennon *et al*., 2011). Dans plusieurs autres modèles de lapin, une immunosuppression induit l'apparition d'infections par les PV sur la localisation d'une infection précédente (Doorbar, 2013). Enfin, la détection d'HPV chez des patientes immunodéprimées et abstinentes à des taux similaires à ceux observés chez les patientes sexuellement actives, pointent encore vers l'existence potentielle d'infections latentes.

1.5.2. Echappement au système immunitaire et persistance de l'infection

A l'instar d'autres virus, les HPV ont mis en place des stratégies d'échappement au système immunitaire très efficaces. Le cycle viral n'engendre, ni apoptose, ni lyse des cellules infectées, ayant lieu exclusivement dans l'épithélium, limitant ainsi l'induction d'inflammation et l'activation du système immunitaire (Grabowska et Riemer, 2012). De plus, les HPV empêchent le déclenchement d'une réaction immunitaire innée et restreignent le

développement d'une immunité adaptative. En effet, les protéines E7 et E6 inhibent respectivement l'expression du récepteur TLR9 (Toll-Like Receptor), senseur de la présence d'ADN viral dans les cellules, ainsi que l'activité de ses effecteurs IRF-1, IRF-3 (Interferon Responsive Element) et NF-*κ*B (Nuclear Factor-kappa B) (Grabowska et Riemer, 2012; Westrich *et al*., 2017). Par conséquent, l'expression des gènes cibles de cette voie, tels les interférons (*IFN-α* et *IFN-β*), les cytokines proinflammatoires IL-1 (Interleukin), IL-12, TNF-α (Tumor Necrosis Factor alpha) ou encore les chémokines (IL-8 ; CXCL14, C-X-C motif chemokine ligand 14 ; et CCL2, C-C motif chemokine ligand 2), est inhibée. Cela empêche non seulement la mise en place d'une réponse médiée par les IFN, mais aussi le recrutement d'autres cellules immunitaires sur le site de l'infection (Grabowska et Riemer, 2012; Westrich *et al*., 2017). De plus, le recrutement de cellules de Langerhans, pouvant présenter les antigènes aux lymphocytes, est perturbé par l'inhibition de l'expression de la cadhérine E et de CCL20 dans les cellules infectées (Smola *et al*., 2017; Westrich *et al*., 2017). Ainsi, une diminution du recrutement des cellules CD8+ a été corrélée à une plus longue persistance de l'infection virale.

Par ailleurs, les protéines virales sont exprimées faiblement dans les cellules hôtes, limitant leur présentation à la surface cellulaire par le CMH de classe I (Complexe majeur d'histocompatibilité), et par conséquent leur reconnaissance par les lymphocytes cytotoxiques. L'expression du CMH de classe I et son adressage à la surface cellulaire sont en plus inhibés par les protéines E7 et E5 des HPV (Grabowska et Riemer, 2012; Westrich *et al*., 2017).

Du reste, les séroconversions qui permettent la production d'anticorps neutralisants, n'ont été observées que chez moins de 20% des patientes infectées par des HPV-HR (Wang *et al*., 2004). De plus, les anticorps produits ne fournissent pas une bonne protection. L'absence d'anticorps neutralisants pourrait ainsi contribuer à la persistance de l'infection par réinfection de la muqueuse. Par ailleurs, l'environnement immunosuppressif observé au niveau de la muqueuse avec la présence de cellules immunitaires tolérogènes sécrétant de l'IL10 et du TGF-*β* (Transforming Growth Factor bêta), favorise le maintien de l'infection (Patel et Chiplunkar, 2009).

De façon intéressante, la clairance des infections par des HPV-BR non carcinogènes est rapide et efficace, tandis que celle des HPV-HR, oncogènes, est plus longue, notamment pour HPV16 (Franco *et al*., 1999; Jaisamrarn *et al*., 2013). La persistance de l'infection est un des facteurs menant à la carcinogenèse induite par les HPV (Dalstein *et al*., 2003; Wallin *et al*., 1999).

1.6. Carcinogenèse associée aux HPV-HR

1.6.1. Processus de carcinogenèse au niveau du col de l'utérus

Le processus de carcinogenèse le plus étudié est celui du cancer du col de l'utérus, dont nous allons donc décrire le déroulement. Le développement du cancer du col de l'utérus suite à l'infection par les HPV-HR est un processus pouvant durer de 10 à 20 ans. C'est la persistance de l'infection par des HPV-HR qui induit l'apparition de lésions précancéreuses, lesquelles peuvent être éliminées ou évoluer en lésions cancéreuses (Fig. 6). Le pic de prévalence de l'infection par les HPV au niveau du col utérin, se situe vers l'âge de 20 ans et diminue au cours des années suivantes, suite à leur élimination par le système immunitaire. Quant aux infections ayant persisté, elles peuvent conduire au développement de lésions précancéreuses, apparaissant quelques années après l'infection (Fig. 6). Ces lésions précancéreuses peuvent soit régresser spontanément, rétablissant l'homéostasie de l'épithélium du col-utérin, soit évoluer en cancer invasif comme décrit dans le paragraphe 1.6.1.2.

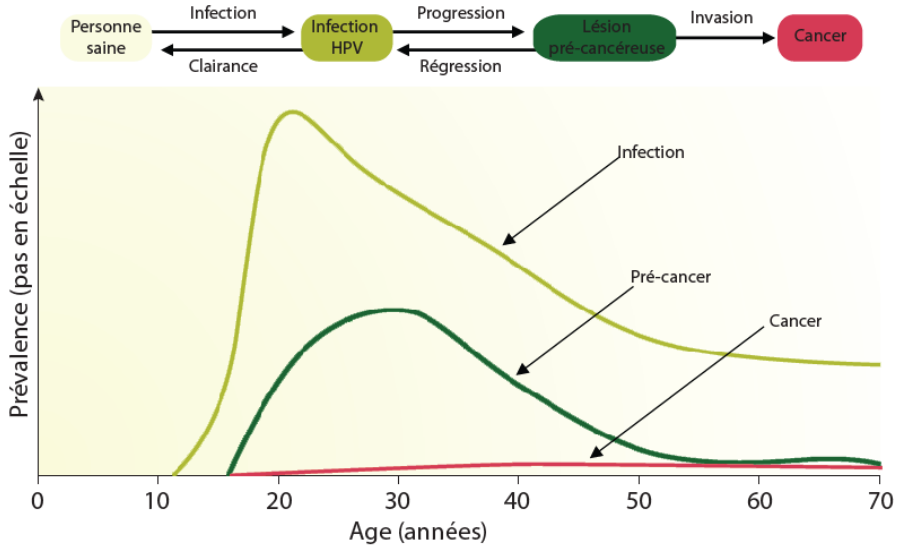


Figure 6. Prévalence des lésions du col de l'utérus induites par les HPV-HR. (Adaptée de Schiffman *et al.*, 2016).

1.6.1.1. Les lésions précancéreuses

Les lésions précancéreuses induites par les HPV présentent différentes gravités et ont des risques distincts d'évolution en cancer. Au niveau histologique, on distingue les lésions néoplasiques intra-épithéliales de grade 1 (CIN1, Cervical Intraepithelial Neoplasia I), CIN2 et CIN3 correspondant à l'atteinte d'un, deux et trois tiers de la hauteur de l'épithélium malpighien du col de l'utérus respectivement (Fig. 7). Une nomenclature équivalente est aussi

utilisée pour les lésions vaginales (VaIN), vulvaires (VIN), péniennes (PeIN) et anales (AIN). Le stade CIN1 correspondant au grade cytologique de lésion intraépithéliale de bas grade LSIL (Low grade Squamous Intra-epithelial Lesion), traduit la présence d'une infection productive dont les cellules superficielles subissent une différenciation. Ces lésions pouvant être causées par l'infection par un HPV-HR ou -BR, ont un faible risque d'évoluer en cancer, mais peuvent évoluer vers les stades CIN2 et CIN3/CIS (Cancer *in situ*). Correspondant à des lésions de haut grade HSIL (High grade Squamous Intra-epithelial Lesion) au niveau cytologique, les lésions CIN2/3 sont caractérisées par des altérations de plus en plus importantes de la morphologie cellulaire et de l'organisation tissulaire sur les 2/3 (CIN2) ou toute la hauteur (CIN3/CIS) de l'épithélium. Ces lésions sont principalement liées à des infections par les HPV-HR. En outre, les lésions CIN2/3 se distinguent par un cycle viral abortif, sans production de virion et sont appelées infections transformantes.

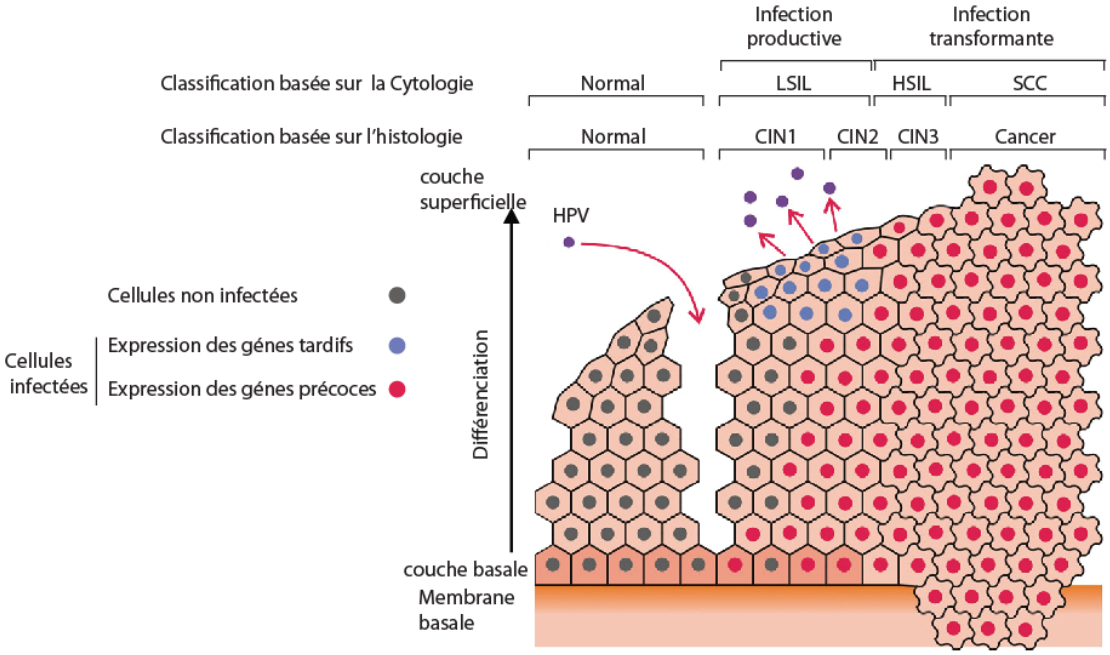


Figure 7. Lésions précancéreuses et cancéreuses induites par les HPV au niveau du col de l'utérus. (Adaptée de Durzynska *et al.*, 2017).

1.6.1.2. Progression de la carcinogenèse

Les lésions précancéreuses CIN2/3 ont environ 30-40% de probabilité d'évoluer en cancer (Fig. 8). La persistance et la progression des lésions précancéreuses sont principalement entretenues et dirigées par une expression accrue des oncoprotéines virales E6 et E7, qui stimulent la prolifération cellulaire, inhibent la différenciation et engendrent une instabilité favorisant l'apparition d'anomalies génomiques promouvant la transformation tumorale. D'ailleurs, l'apparition de ces anomalies corrèle avec le développement du cancer du col de

l'utérus. Cela favorise aussi l'intégration du génome viral dans celui des cellules hôtes, qui est progressive tout le long de la carcinogenèse. Celle-ci est de fait accompagnée par une perte des épisomes (Fig. 8).

L'intégration du génome viral conduit dans la majorité des cas à une perte de l'ORF codant E2, un répresseur de l'expression des gènes E6 et E7. L'augmentation de l'expression d'E6 et E7 qui en résulte, favorise l'immortalisation et la transformation cellulaire (Fig. 8).

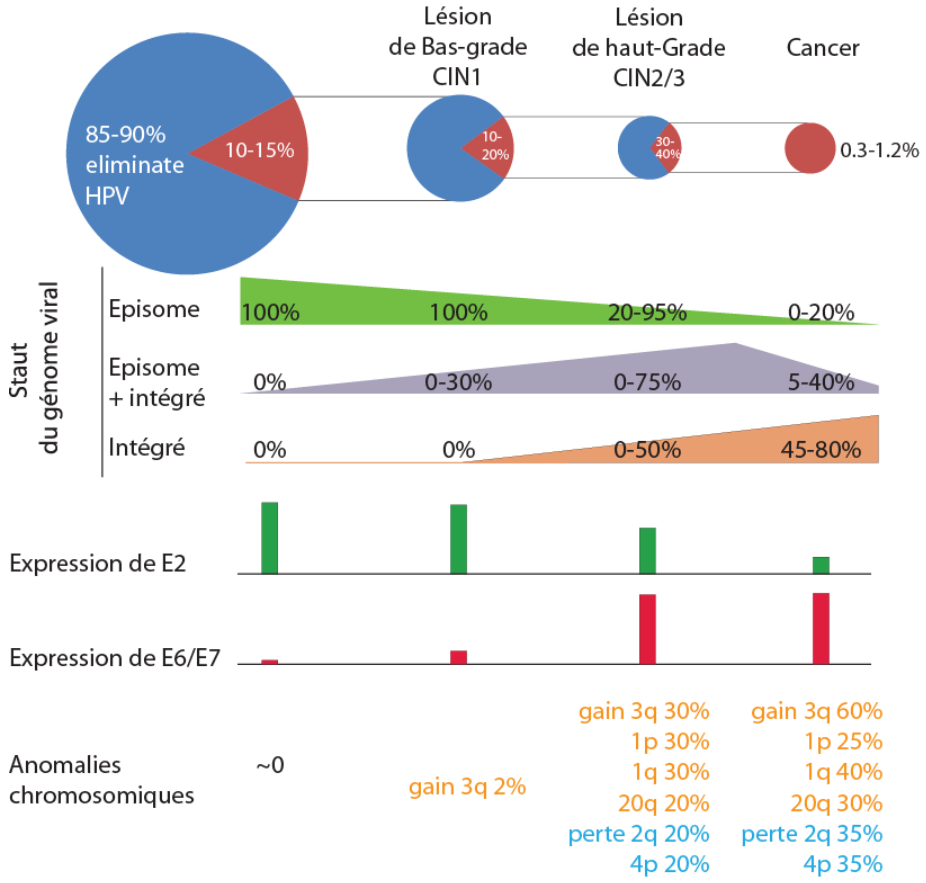


Figure 8. Progression de la carcinogenèse du col de l'utérus et principaux évènements moléculaires associés. (Adaptée de Shulzhenko *et al.*, 2014).

1.6.2. Dépistage et prévention du cancer du col de l'utérus

Les cancers du col de l'utérus (et les autres cancers liés aux HPV) sont traités par chirurgie et/ou par radio- et chimiothérapies et ce, selon la taille de la tumeur et sa localisation. Leur lien de causalité avec les HPV-HR et le temps d'apparition du cancer (précédé de lésions précancéreuses) implique que le cancer du col de l'utérus fasse l'objet de stratégies de prévention primaire basée sur la vaccination et de prévention secondaire par FCU (Frottis Cervico-Utérin) pour dépister d'éventuelles lésions précancéreuses.

1.6.2.1. Dépistage

Le dépistage est réalisé chez les femmes âgées de 25 à 65 ans par des FCU. Les frottis sont effectués tous les 3 ans à la suite de 2 FCU aux profils cytologiques normaux, effectués dans un intervalle d'un an. Le FCU consiste à déceler des anomalies cytologiques précancéreuses et à les prendre en charge selon les recommandations de la haute autorité de santé. Si la cytologie montre la présence d'un ASC-US (Atypical Squamous Cells of Undetermined Significance), un triage par test HPV est recommandé en France. Le test HPV est aussi préconisé par le CNOGF pour le suivi des femmes ayant été traitées pour une lésion de haut grade (HSIL). Dans certains pays, le test HPV est utilisé en 1ère intention ou associé à un FCU, afin de compenser le manque de sensibilité de celui-ci (Crow, 2012; Goodman, 2015; Kessler, 2017; Schiffman et al., 2016). La persistance de l'infection et le développement des lésions HSIL étant influencés par le type d'HPV infectant le col de l'utérus, un génotypage permet de mieux détecter les lésions à risque de progression et d'optimiser la prise en charge des patientes.

1.6.2.2. Vaccination

Les cancers du col de l'utérus étant dus à un virus, en l'occurrence les HPV-HR, prévenir l'infection par la vaccination limite le développement de cancers. En utilisant des VLP (Virus like-particle) formées par les protéines L1 de plusieurs HPV, 3 vaccins ont été mis au point et sont disponibles en France :

- Cervarix ®, dirigé contre HPV16 et HPV18, les génotypes responsables de la majorité des cancers.
- Gardasil ®, dirigé contre HPV16 et HPV18 ainsi que HPV6 et HPV11 responsables de l'apparition des condylomes.
- Gardasil 9®, dirigé contre HPV16, HPV18, HPV31, HPV33, HPV45, HPV52, HPV58, HPV6 et HPV11.

Les recommandations vaccinales, similaires pour le Cervarix ®, Gardasil ® et Gardasil 9®, préconisent de vacciner les jeunes filles de 11 à 14 ans (jusqu'à 13 ans pour Gardasil), avec 2 doses de vaccin dans un intervalle de 6 mois. Un schéma de vaccination pour les filles plus âgées (15-19 ans ; 14-19 ans pour Gardasil ®) préconise 3 injections espacées de 2 et 4 mois (1 et 5 mois pour Cervarix ®). Par ailleurs, le vaccin est conseillé pour les hommes (jusqu'à 26 ans) ayant des relations homosexuelles, étant donné le risque élevé de développer un cancer du canal anal. Un schéma de vaccinations avec 3 doses y est préconisé.

L'efficacité et la sûreté des vaccins Cervarix ®et Gardasil ® a été montré par plusieurs études. Le nombre de lésions du col de l'utérus (CIN1 et CIN2/3) et du canal anal (AIN), associées à HPV16 et 18 a ainsi baissé de 100% (85,5-100), suite à la vaccination par le Cervarix ® ou Gardasil ® (Lehtinen et al., 2012; Mejilla et al., 2017; Smola et al., 2017). De façon intéressante, les vaccins bivalents font aussi baisser, d'environ 60%, le nombre de lésions CIN3 associées à d'autres HPV-HR, notamment HPV31 et HPV45, *via* une réaction croisée (Einstein et al., 2011; Wheeler et al., 2012). Il faudra tout de même attendre quelques années afin d'étudier l'effet de la vaccination sur le développement de cancers. De plus, les études menées sur la sûreté des vaccins, n'ont pas trouvé de rapport de causalité entre les vaccins et le développement d'une quinzaine de maladies auto-immunes (dont la sclérose en plaque, le lupus érythémateux disséminé) (Arnheim-Dahlström et al., 2013; Miranda et al., 2017; Andrews et al., 2017).

Toutefois, malgré l'efficacité, l'immunogénicité et la tolérance des vaccins HPV, la couverture vaccinale reste très limitée dans plusieurs pays, en particulier en France où elle n'atteint pas 20% chez les filles. Cette faible couverture vaccinale peut induire une persistance du risque de développement de cancers associés aux HPV, nécessitant la poursuite des investigations afin d'améliorer le dépistage de ces cancers. De plus, le vaccin étant préconisé pour les filles âgées de 11 à 14 ans, une grande partie de la population ayant dépassé cet âge lors de l'introduction des campagnes de vaccination n'a pas été vaccinée. Ainsi, on peut parler de *shift* de protection du développement de cancers invasifs d'environ 10-20 années après la vaccination. Enfin, la non vaccination des garçons réduit l'immunité de troupeau (*Herd Immunity*) et permet au virus de persister dans la population.

1.6.3. Rôle des protéines E7 et E6 dans la carcinogenèse

1.6.3.1. Rôle de la protéine E7 des HPV-HR

La protéine E7 est constituée de 98 aa (acides aminés) et présente une localisation nucléaire, nucléolaire et cytoplasmique (Roman et Munger, 2013). Elle comporte trois domaines conservés lui conférant ses caractéristiques fonctionnelles (Fig. 9). Deux régions CR1 et CR2 (Conserved Region), situées sur la partie N-ter (amino-terminale) sont nommées ainsi par rapport à leur similarité aux domaines correspondants dans la protéine E1A des adénovirus. Sur la partie C-ter (carboxy-terminale) se trouve un domaine de liaison au zinc, constitué de 2 motifs C-x-x-C (C, Cystéine ; x, acide aminé) et la séquence d'export nucléaire. La séquence de localisation nucléaire n'a pas encore été identifiée.

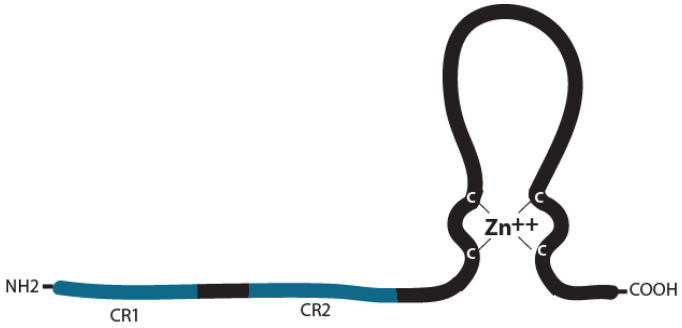


Figure 9. La protéine E7 des HPV-HR.

La protéine E7 ne possédant pas d'activité enzymatique, ni de capacité de liaison à l'ADN, elle agit *via* son interaction avec des partenaires cellulaires dont elle module les activités (Fig. 10). Ses domaines CR1 et CR2, désordonnés sur le plan structurel, lui confèrent la capacité de se lier avec de très nombreux facteurs cellulaires.

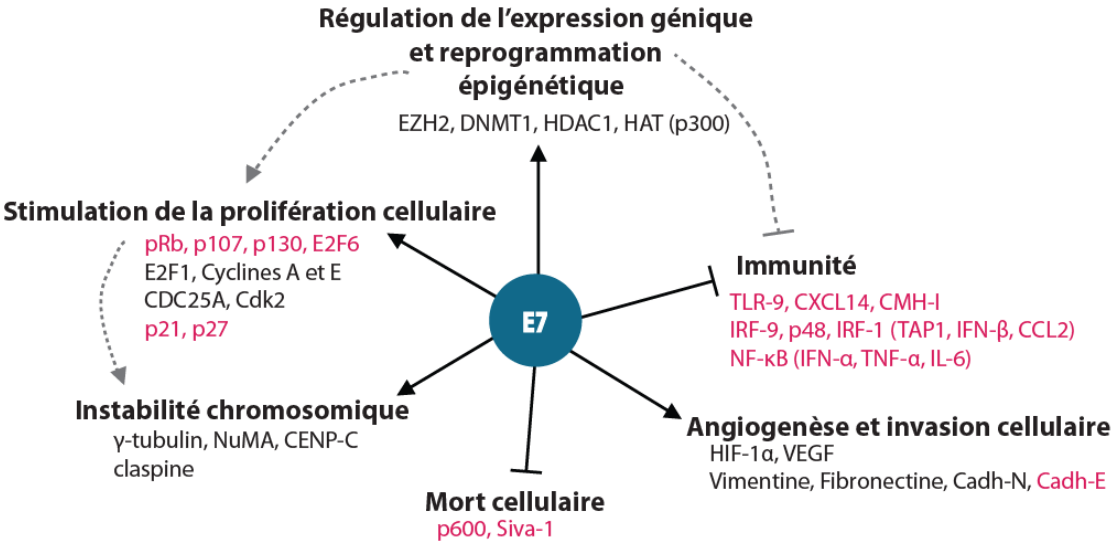


Figure 10. Rôle de la protéine E7 des HPV-HR. (D'après Mittal et Banks, 2017; Westrich *et al.*, 2017).

- Stimulation de la prolifération cellulaire

E7 stimule la prolifération cellulaire en ciblant particulièrement l'entrée en phase G1 et la transition G1/S du cycle cellulaire régulées par les protéines de la famille rétinoblastome pRb, p107 et p130 (Fig. 10). Ces protéines à poches séquestrent les facteurs de transcription de la famille E2F, permettant l'entrée et l'avancé dans le cycle cellulaire. En se fixant aux protéines pRb, p107 et p130 et en les empêchant d'interagir avec E2F, E7 permet de lever la répression (Mittal et Banks, 2017a). E7 se lie aussi à E2F1 et augmente son activité transactivatrice (Hwang et al., 2002). E2F1 peut donc induire la transcription de gènes tels les cyclines E et A, CDC25A (cell division cycle 25A) impliqués dans l'entrée en phase S.

L'impact d'E7 des HPV-HR sur l'entrée en phase S est remarquablement plus important que celles produites par les HPV-BR. E7 des HPV-HR possédant des interactions plus fortes avec les protéines E2F1, pRb, p107 et p130, avec en sus la capacité d'orienter les trois dernières vers la dégradation (Doorbar, 2013; McLaughlin-Drubin et Münger, 2009). E7 permet ainsi l'expression des gènes contenant une boite E dans leurs promoteurs, et dépendant des facteurs activateurs de la famille E2F (E2F1, E2F2, E2F5). Par ailleurs, E7 se lie à E2F6, dont l'expression est médiée par E2F1 et l'empêche de réprimer la transcription des gènes à boite E, par reprogrammation épigénétique *via* PRC (polycomb repressor complex).
Ainsi, les cyclines E, A et la phosphatase CDC25A, produites suite à l'activation d'E2F1, permettent d'activer la Cdk2 (cyclin-dependant kinase 2), qui va à son tour consolider l'entrée et l'avancement en phase S, en renforçant l'expression d'E2F et des gènes impliqués dans la réplication de l'ADN, tels l'ADN polymérases alpha, RPA, PCNA (Proliferating Cell Nuclear Antigen). E7 permet aussi de stimuler l'activité de la Cdk2 en : (i) s'y liant directement et (ii) en empêchant son inhibition par les CKI (Cyclin Kinase Inhibitor), p21 et p27 (McLaughlin-Drubin et Münger, 2009). Par l'inhibition de p21 et p27, E7 permet aussi d'inhiber les voies de signalisation du TGF-β et d'éluder les effets antiprolifératifs de ce dernier. E7 permet d'outrepasser les mécanismes de contrôle de la mitose, notamment ceux impliquant les fuseaux mitotiques.

Par ailleurs, la protéine E7 permet de stimuler la croissance et la résistance aux signaux antiprolifératifs en dérégulant l'expression de nombreux gènes, *via* l'induction d'une reprogrammation épigénétique.

- Instabilité chromosomique

L'activité accrue des complexes Cdk2/cycline E induite par E7, conduit à une surduplication des centrosomes *via* le recrutement exagéré de la protéine ALK4 (Activin Receptor-Like Kinase 4) (Korzeniewski *et al*., 2011). En se liant avec la γ-tubuline et en altérant sa localisation au niveau des centrosomes, E7 favorise l'apparition de centrosomes surnuméraires (Nguyen *et al*., 2007). Les mitoses multipolaires résultantes de cette surduplication induisent une ségrégation chromosomale aberrante aboutissant *in fine* à l'apparition d'aneuploïdies. E7 induit aussi la délocalisation des protéines NuMA1 (Nuclear Mitotic Apparatus 1) et Dynéine des microtubules, engendrant des défauts d'organisation des fuseaux mitotiques et d'alignement des chromosomes (Nguyen et Münger, 2009; Nguyen *et al*., 2008).

Par l'intervention d'E7 et E6, les cellules comportant des défauts d'alignement des chromosomes durant la métaphase peuvent continuer leur division cellulaire ou sortir de la mitose sans cytodiérèse, produisant des cellules tétraploïdes. En temps normal, ces cellules sont bloquées en G1, mais l'expression d'E7 et E6 permettent de les rediriger vers un nouveau cycle cellulaire. De plus, E7 interagit avec la protéine CENP-C (centromère C), empêche sa liaison avec l'ADN *α*-satellite pouvant provoquer des défauts d'attachements des chromosomes sur le fuseau mitotique (Yaginuma *et al*., 2015).

Par ailleurs, l'expression d'E7 et E6 induit l'apparition de cassures double brin de l'ADN, entraînant la formation de ponts entre chromosomes pendant l'anaphase et favorisant les réarrangements chromosomiques (Korzeniewski *et al*., 2011). L'apparition de cassures double brin de l'ADN pourrait résulter du stress réplicatif, engendré par la réplication de l'ADN induite par E7.

En outre, E7 ainsi qu'E6 permettent de passer le point de contrôle G2/M malgré la présence de lésions de l'ADN. Alors qu'E6 inhibe p53, principal régulateur du point G2/M, E7 induit la dégradation de la claspine, acteur de l'arrêt du cycle cellulaire médié par ATR et Chk1 (Checkpoint kinase 1) (Korzeniewski *et al*., 2011).

- Inhibition de la mort cellulaire

Lorsqu'une cellule épithéliale qui n'adhère plus à la lame basale tente de proliférer, elle est dirigée vers une mort cellulaire par perte d'encrage, l'anoïkis. L'expression d'E7, qui se lie et inhibe p600, une protéine clé pour l'anoïkis, permet aux cellules de résister et d'outrepasser ce mécanisme de contrôle (Huh et al., 2005).

Par ailleurs, E7 des HPV-HR peut inhiber l'apoptose en empêchant l'interaction entre la protéine pro-apoptotique Siva-1 (Apoptosis Regulatory Protein Siva) et la protéine anti-apoptotique Bcl-x_L (Bcl-2-like protein 1 long isoform), ainsi qu'en augmentant l'expression de Bcl-2 (B-cell lymphoma 2) (Liu et al., 2008; Severino et al., 2007). E7 permet aussi d'inhiber la mort cellulaire induite par le TNF-*α* (Thompson et al., 2001). Toutefois, c'est principalement E6 qui permet aux HPV d'inhiber l'apoptose (Voir ci-dessous).

- Régulation de l'expression génique et reprogrammation épigénétique

La régulation épigénétique de l'expression des gènes met en jeu notamment la méthylation de l'ADN et les modifications post-traductionnelles des histones permettant le remodelage de la chromatine. Ces modifications influencent le recrutement des facteurs de transcription au niveau des promoteurs et par conséquent leurs niveaux de transcription.

E7 possède la capacité d'interagir par les protéines impliquées dans la régulation épigénétique et d'altérer la régulation de l'expression des gènes. E7 augmente l'expression des histones déméthylases KDM6A (Lysine demethylase 6A) et KDM6B, et diminue la méthyl transférase EZH2 (Enhancer of zeste 2 polycomb repressive complex 2 subunit), provoquant ainsi la déméthylation de la Lysine 27 de l'Histone 3, favorisant ainsi l'expression de gènes à Homeoboîte tel que celui codant p16 (Durzynska et al., 2017; McLaughlin-Drubin et al., 2011). E7 peut aussi influencer le recrutement d'EZH2 au sein du promoteur de TLR-9, inhibant ainsi son expression. L'inhibition de l'expression de TLR-9 est aussi confortée par la désacétylation de l'histone 4, suite au recrutement de HDAC1 (Histone deacetylase 1) par E7. HDAC1 est aussi recrutée par E7 vers les promoteurs des gènes sous la dépendance d'IRF-1 (*CCL2*, *IFN-β*, *TAP1*, *CMH-I*), inhibant leur expression. D'autres gènes de l'immunité comme *CXCL14* et *IFN-κ*, sont inhibés au niveau transcriptionnel *via* le recrutement de DNMT1 (DNA-Methyl Transferase 1) qui induit une hyperméthylation de leurs promoteurs (Songock et al., 2017a; Westrich et al., 2017). En modulant l'expression de miRNA (micro RNA), tels les miR-203 et miR-16-1, E7 influence la régulation des gènes au niveau post-transcriptionnel (Jin et Xu, 2015).

Au total, E7 permet aux HPV-HR de modeler à leur guise l'expression de nombreux gènes cellulaires, impactant de fait leurs fonctions respectives. Ainsi, les effets potentiels découlant de cette propriété d'E7 ont une importance capitale dans la carcinogenèse associée aux HPV.

- Immunité

E7 permet aussi de moduler la réponse immunitaire afin de favoriser la propagation virale. C'est en partie *via* la reprogrammation épigénétique qu'E7 régule l'expression des gènes impliqués dans l'immunité (Fig. 10). L'effet d'E7 sur l'immunité a été principalement décrit dans le chapitre 1.5.2. Echappement au système immunitaire et persistance de l'infection.

- Angiogenèse et invasion cellulaire

En augmentant la transcription induite par HIF-1α (Hypoxia Inducible Factor) en réponse à l'hypoxie, E7 favorise l'expression du VEGF (Vascular Endothelial Growth Factor) et de l'IL-8 et l'induction de l'angiogenèse (Bodily *et al*., 2011; Zhang *et al*., 2014). De fait, les tumeurs induites par les HPV sont bien vascularisées avec une faible hypoxie par rapport aux tumeurs non-HPV (Hanns *et al*., 2015). Par ailleurs, *via* la reprogrammation épigénétique, E7 inhibe l'expression de la cadhérine E, marqueur épithélial et induit celle des marqueurs mésenchymateux tels la vimentine et la cadhérine N, promouvant la transition épithélio-mésenchymateuse (Hellner *et al*., 2009; Jung *et al*., 2013; Laurson *et al*., 2010; Liu *et al*., 2017). Le phénotype mésenchymateux favorise l'invasion tumorale et le processus métastatique.

1.6.3.2. Rôle de la protéine E6 des HPV-HR

La protéine E6 est une composée d'environ 150 aa et comprend plusieurs motifs fonctionnels bien conservés au sein des HPV (Vande Pol et Klingelhutz, 2013). Elle présente une localisation majoritairement nucléaire et possède 3 séquences NLS (Nuclear Localisation Signal). E6 comprend deux domaines de liaison au zinc (Zinc Domain, ZD) situés sur les parties N- (ZD1) et C-terminales (ZD2) (Fig. 11). ZD1 et ZD2 permettent la liaison au motif LxxLL (L, Leucine ; x, acide aminé) présent notamment chez le partenaire cellulaire d'E6, E6AP (E6 Associated Protein). Enfin, un motif PBM (PDZ Binding Motif), situé sur la partie C-ter et présent uniquement dans les protéines E6 des HPV-HR, assure la liaison aux protéines à domaines PDZ (PSD-95/DLG/ZO-1 domain) (Fig. 11).

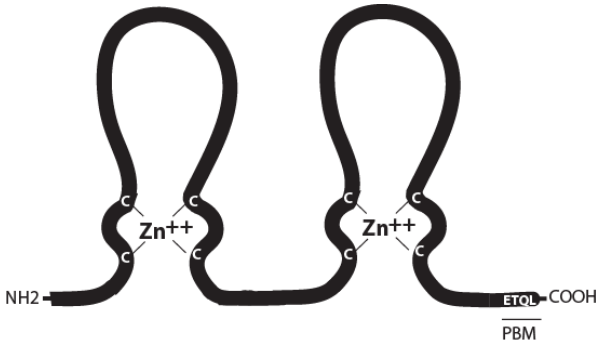


Figure 11. La protéine E6 d'HPV16.

Ces différents domaines permettent l'interaction d'E6 avec plusieurs facteurs cellulaires affectant notamment la division, la différenciation et la mort cellulaire (Fig. 12).

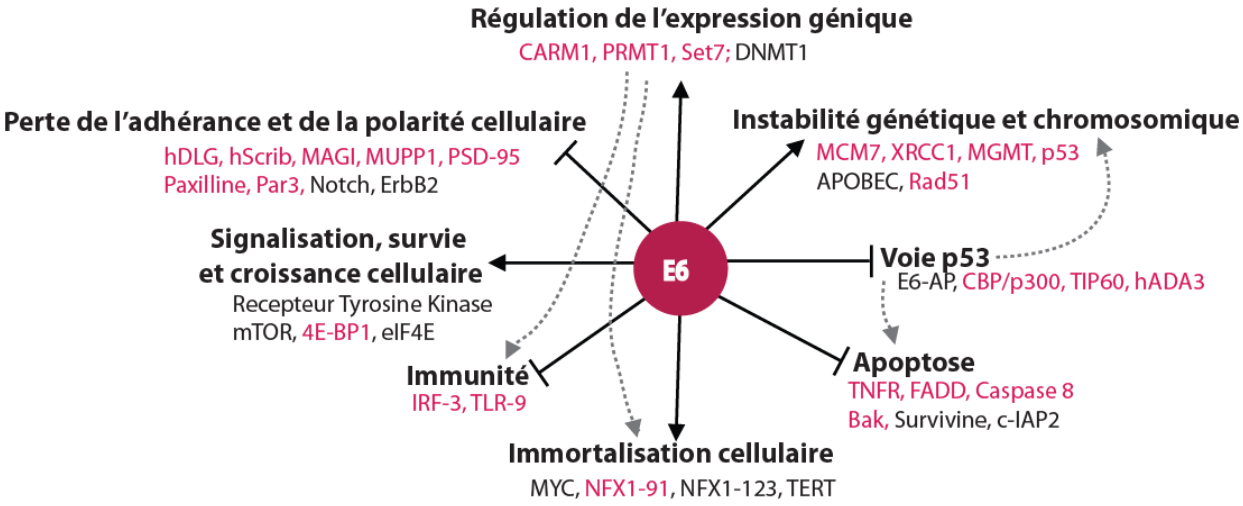


Figure 12. Rôle de l'oncoprotéine E6 HPV-HR.

- Inhibition de la voie p53

La protéine E6 permet la survie et la prolifération cellulaire en inhibant principalement la voie de régulation induite par p53.
En réponse à différents stress, tels l'activation d'oncogène (Ex : prolifération induite par E7), l'infection virale et l'apparition de lésions de l'ADN, la protéine p53 est stabilisée et induit l'arrêt du cycle cellulaire afin de réparer les lésions de l'ADN, ou l'apoptose en cas de dommages trop conséquents. p53 est une protéine « suppresseur de tumeur » capitale, mutée dans plus de 50% des cancers. Dans les cancers associés aux HPV, p53 est le plus souvent sauvage et est inhibée par la protéine E6 par plusieurs mécanismes.

E6 des HPV-HR favorise la dégradation de la protéine p53 *via* l'induction de son ubiquitinylation par l'ubiquitine ligase E6AP (ou UBE3A) (Scheffner *et al*., 1990, 1993). E6 se lie au motif LxxLL d'E6AP *via* son domaine ZD1, recrute p53 pour former un complexe ternaire, lequel après dimérisation d'E6 induit l'ubiquitinylation de p53 (Ansari *et al*., 2012; Martinez-Zapien *et al*., 2016; Zanier *et al*., 2012).

De plus, en empêchant son activation, E6 inhibe l'activité de p53 indépendamment de sa dégradation. En effet, lors d'un stress, p53 subit des modifications post-traductionnelles, notamment des phosphorylations et/ou acétylations, permettant d'augmenter sa stabilité ainsi que son interaction avec ses cibles. Une fois stabilisée, elle induit la transcription de gènes impliqués dans l'arrêt du cycle cellulaire, la réparation de l'ADN et l'apoptose. E6 empêche l'acétylation de p53 par ses co-activateurs, CBP/p300 (CREB Binding protein/Histone acetyl-transferase p300), TIP60 (Tat interacting protein, 60kDa) et hADA3 (human homologue of

yeast Alteration/deficiency in activation 3), inhibant ainsi son activité transactivatrice (Hu *et al*., 2009; Jha *et al*., 2010; Kumar *et al*., 2002; Patel *et al*., 1999; Thomas et Chiang, 2005; Zimmermann *et al*., 1999). Par ailleurs, E6 inhibe non seulement la fixation de p53 à l'ADN *in vitro*, mais la séquestre aussi dans le cytoplasme (Lechner et Laimins, 1994; Mantovani et Banks, 1999; Thomas *et al*., 1995).

De manière intéressante, les protéines E6 des HPV-HR possèdent des activités plus fortes que celles des HPV-BR, notamment une capacité accrue à inhiber p53 et ses coactivateurs. Les protéines E6 des HPV-HR et -BR ciblent aussi des facteurs différents. Par ailleurs, E6 des HPV-HR cible des protéines suppresseur de tumeur de la famille p53, à savoir p63*β* et p73, induisant leur dégradation.

- Inhibition de l'apoptose

E6 inhibe les deux voies d'apoptose, extrinsèque et intrinsèque. Tout d'abord, en inhibant p53, E6 l'empêche d'induire l'expression du récepteur de mort Fas (FasR) et des protéines pro-apoptotiques telles Puma (p53-upregulated modulator of apoptosis), Bax (Bcl2-associated X protein et Noxa (appelé aussi PMAIP1) (Riley *et al*., 2008; Wallace et Galloway, 2015). L'activation de la voie extrinsèque d'apoptose, induite par la fixation du TNF-*α*, Fas et TRAIL à leurs récepteurs, est aussi inhibée par E6 *via* l'interruption de la transduction du signal de mort à trois niveaux (Cai *et al*., 2013; Wallace et Galloway, 2015). (i) En se fixant au récepteur du TNF-*α* et en inhibant le recrutement des protéines adaptatrices FADD (Fas-associated death domain) (Filippova *et al*., 2002). (ii) En induisant la dégradation de FADD, évitant le recrutement et l'activation de la caspase 8 (Filippova *et al*., 2004; Garnett *et al*., 2006; Tungteakkhun *et al*., 2008). (iii) En se liant à la caspase 8, E6 induit sa dégradation ou sa relocalisation au niveau du noyau, limitant son activité cytoplasmique et donc l'activation des caspases effectrices (Filippova *et al*., 2007; Garnett *et al*., 2006; Manzo-Merino *et al*., 2014).

Par ailleurs, E6 inhibe la voie intrinsèque d'apoptose en faisant pencher la balance entre les protéines pro-apoptotiques et anti-apoptotiques en faveur de ces dernières. Ainsi, E6 induit l'augmentation de l'expression des protéines anti-apoptotiques, Bcl-2, Survivine et c-IAP2 (Cellular Inhibitor Of Apoptosis 2) (Borbély *et al*., 2006; Du *et al*., 2004; James *et al*., 2006), tandis qu'elle induit la dégradation de la protéine pro-apoptotique Bak (Bcl2-antagonist killer 1) (Du *et al*., 2004; Underbrink *et al*., 2016).

- Immortalisation cellulaire

En évitant la sénescence réplicative *via* l'induction de l'expression et de l'activité de la télomérase, E6 permet d'immortaliser les cellules. Les cellules somatiques possèdent une capacité proliférative limitée à cause du raccourcissement des télomères au cours des divisions cellulaires successives. Le raccourcissement des télomères continue jusqu'à un point limite où les cellules s'arrêtent de proliférer et subissent la sénescence réplicative, appelée limite de Hayflick. Cette limite correspond à 50 divisions cellulaires, au-delà de laquelle les cellules ne peuvent plus se dupliquer. La télomérase, composée des unités TERC (Telomerase RNA component) et TERT (Telomerase Reverse transcriptase), permet de rallonger les télomères et d'éviter la survenue de la sénescence (Galloway *et al*., 2005).

C'est principalement en induisant la transcription et l'expression de TERT, qu'E6 restaure l'activité télomérase dans les cellules. En effet, E6 est recrutée sur le promoteur de TERT en compagnie d'activateurs de la transcription tel(s) MYC, stimulant ainsi la transcription et l'expression de TERT (Oh *et al*., 2001; Veldman *et al*., 2001, 2003). E6 renforce aussi la transcription de TERT en induisant la dégradation de son répresseur transcriptionnel, NFX1-91 (Nuclear transcription factor, X-box binding 1 isoform 91) (Gewin *et al*., 2004). De plus, en interagissant avec NFX1-123 et PABP4 (poly(A) Binding Protein cytoplasmic 4), E6 permet aussi d'augmenter la stabilité des transcrits TERT (Katzenellenbogen *et al*., 2007, 2009). Par ailleurs, E6 se lie à la télomérase, stimulant son activité (Liu *et al*., 2009).

- Instabilité génétique et chromosomique

L'expression d'E6 permet de contourner les points de contrôle du cycle cellulaire et empêche la réparation de l'ADN, favorisant l'instabilité génétique et génomique (Thompson *et al*., 1997; Wallace et Galloway, 2015). En induisant la dégradation de MCM7 (Minichromosome maintenance complex component 7), E6 altère le contrôle de la réplication de l'ADN, ce qui induit l'apparition d'aneuploïdies (Kühne et Banks, 1998).

E6 fournit un environnement propice à l'apparition de mutations. L'instabilité génétique engendrée s'aggravant de plus en plus au cours des divisions cellulaires, cela se traduit par l'apparition d'instabilités chromosomiques. En interagissant avec XRCC1 (X-ray Repair Cross Complementing 1), E6 inhibe le système de réparation BER (Base Excision Repair) et le système de réparation des cassures simple brin de l'ADN dont l'accumulation peut se traduire par l'apparition de cassures double brins (Iftner *et al*., 2002). De plus, E6 induit la

dégradation de MGMT (O6-methylguanine-DNA methyltransferase), une enzyme responsable de la réparation des dommages liés à l'alkylation de l'ADN (Srivenugopal *et al*i-Osman, 2002).

En induisant l'augmentation de la transcription et de l'expression de l'APOBEC3B et 3C (Apolipoprotein B mRNA Editing Enzyme Catalytic Subunit), deux cytosines désaminases, E6 promeut l'apparition de mutations liées à l'activité de ces APOBEC (Mori *et al*., 2017; Vieira *et al*., 2014). De fait, une signature mutationnelle, des transversions cytosine vers thymidine caractéristiques de l'activité d'APOBEC, est retrouvée dans les cancers du col de l'utérus. De façon intéressante, les protéines E6 des HPV-HR induisent plus fortement l'expression des APOBEC comparativement à celles des HPV-BR (Vieira *et al*., 2014).

En favorisant l'apparition de cassures double brin, E6 promeut les recombinaisons favorables à l'intégration du génome viral au sein de l'ADN cellulaire. De façon intéressante, E6 induit la délocalisation de la protéine RAD51 (RAD51 recombinase) des sites de cassures double brin et induit une diminution de la recombinaison homologue (Wallace *et al*., 2017). Par conséquent, E6 induit la persistance des cassures double brins et leur réparation par le système NHEJ (Non Homologous End Joining), moins fiable, qui favorise l'apparition de mutations.

- Perte de l'adhérence et de la polarité cellulaire

La prolifération et la différenciation cellulaire sont contrôlées au niveau des épithéliums afin de mettre en place et de maintenir une organisation tissulaire importante pour leur fonctions. Le maintien de l'équilibre au sein de l'épithélium requiert des contacts cellule-cellule et cellule-matrice extracellulaire et la mise en place d'une polarité cellulaire. Les carcinomes épidermoïdes et les adénocarcinomes, dont ceux causés par les HPV, sont caractérisés par une perte des contacts et de la polarité cellulaire.

E6 se lie à la paxilline, une protéine intracellulaire associée aux intégrines qui participent à la transduction du signal en provenance de la matrice extracellulaire ; il en résulte une altération de l'organisation des fibres d'actine et de la mise en place de points de contact focaux (Tong et Howley, 1997), modifiant ainsi l'organisation du cytosquelette et l'architecture cellulaire et tissulaire. De plus, *via* son PBM, E6 se lie avec de nombreuses protéines à domaine PDZ impliquées dans l'adhérence et la polarité cellulaire, tels PATJ (PALS1-associated TJ protein), Scrib (Scribbled Planar Cell Polarity Protein), DLG1 (Disc Large protein 1), MUPP1 (Multiple PDZ domain crumbs cell polarity complex component), PTPN13 (Protein tyrosine

phosphatase), MAGI (Membrane associated guanylate kinase interacting protein-like) -1,-2 et -3, induisant leur dégradation (Gardiol *et al*., 1999; Glaunsinger *et al*., 2000; Lee *et al*., 2000; Nakagawa et Huibregtse, 2000; Spanos *et al*., 2008; Storrs et Silverstein, 2007; Thomas *et al*., 2002). Par conséquent, cette perte d'adhérence induite par E6 altère la polarité cellulaire, empêche le contrôle de la prolifération et favorise la croissance cellulaire indépendante de l'ancrage et l'invasion.

- Signalisation, survie et croissance cellulaire

E6 favorise la survie, la croissance, et l'invasion cellulaire en stimulant l'activité des petites protéines G et les voies de signalisation PI3K/AKT, Wnt et Notch.

En se liant à Tip-1 (Tax-Interacting Protein 1) et ARHGEF16 (Rho guanine nucleotide exchange factor 16), une protéine GEF (GTP Exchange Factor), E6 permet de stimuler l'activité de la protéine Cdc42 (Rho Cell division cycle 42) (Oliver *et al*., 2011). De plus, E6 se lie à E6TP1 (E6 Targeted Protein 1), une protéine GAP (GTPase-Activating Protein), induit sa dégradation, ce qui favorise l'activation des protéines G Rap1 (Ras-Related Protein Krev-1) et Rap2 (Gao *et al*., 1999; Singh *et al*., 2003). L'activation de ces 2 protéines G permet le réarrangement du cytosquelette favorisant la migration cellulaire. L'activité de la protéine Rheb (Ras Homolog Enriched In Brain) est aussi augmentée *via* la dégradation d'une autre protéine GAP, TSC2 (Tuberous Sclerosis Complex 2) (Lu *et al*., 2004). Rheb est particulièrement impliquée dans la stimulation de la voie mTOR (mamalian Target Of Rapamycin). Ainsi, l'action d'E6 sur l'activité des petites protéines G permet non seulement de favoriser la motilité cellulaire, mais pourrait aussi induire le renforcement des signaux mitogènes notamment ceux aboutissant à l'activation de mTOR.

De manière intéressante, les HPV-HR stimulent la voie de survie cellulaire PI3K/AKT, permettant de renforcer l'activation de mTOR. En augmentant l'activité des récepteurs à activité tyrosine kinases (EGFR; IGFR, Insulin like growth factor 1 receptor ; INSR, Insulin receptor) et en stabilisant ErbB2 (Erb-b2 receptor tyrosine kinase 2), E6 stimule la voie PI3K/AKT (Narisawa-Saito *et al*., 2007; Spangle et Munger, 2013), laquelle promeut la épithélio-mésenchymateuse et l'invasion cellulaire.
De plus, *via* son interaction avec les protéines à domaines PDZ, E6 inhibe PTEN, un inhibiteur de la voie PI3K/AKT, renforçant ainsi l'activation de mTOR.

La voie mTOR permet de stimuler la synthèse protéique, favorisant la survie et la croissance cellulaire. En outre, E6 induit l'augmentation de l'expression d'eIF4E (eukaryotic Initiation Factor 4E) et de la traduction dépendante de la coiffe (Spangle et Münger, 2010; Wang *et al*., 2013).

D'autres voies de signalisation cellulaire, Wnt/*β*-catenin et Notch, sont aussi stimulées par E6 (Bello *et al*., 2015; Wallace et Galloway, 2015). En outre, l'activation de la voie Wnt/*β*-catenin serait entre autres liée à une stabilisation de la *β*-caténine. Ces voies de signalisation favorisent la prolifération et l'invasion cellulaire.

- Régulation de l'expression génique

En se liant à des facteurs de transcription, tels p53 et MYC, E6 permet respectivement de réprimer ou stimuler l'expression de leurs gènes cibles. E6 peut aussi affecter l'expression de nombreux gènes *via* leur régulation au niveau épigénétique (méthylation de l'ADN et modifications des histones).

En effet, E6 peut interagir et inhiber l'activité de protéines impliquées dans le remodelage de chromatine notamment : (i) les HAT (Histone Acetyl Transferase) CBP/p300, hADA et TIP60 et (ii) les HMT (Histone Methyl Transferase) CARM1 (Coactivator associated arginine methyltransferase 1), PRMT1 (Protein arginine methyltransferase 1) et Set7 (SET domain containing lysine methyltransferase 7) (Durzynska *et al*., 2017). Par ailleurs, E6 induit l'augmentation de l'expression de la DNMT1 et la méthylation des promoteurs des gènes codant l'IFN-κ et la cadhérine E, conduisant par conséquent à leur répression (D'Costa *et al*., 2012; Reiser *et al*., 2011).

La régulation de l'expression génique médiée par E6 peut affecter l'expression de miRNA, contribuant à augmenter l'impact d'E6 sur le transcriptome cellulaire. Les miRNA régulent l'expression de nombreuses cibles au niveau post-transcriptionnel. Ainsi, E6 permet d'augmenter l'expression des miR-9 et miR-21 (Chopjitt *et al*., 2015; Liu *et al*., 2014b). Inversement, elle permet d'inhiber l'expression de miR suppresseur de tumeur tels, miR-375, miR-218 et miR-23b, levant la répression exercée sur les gènes impliqués dans la prolifération et l'invasion cellulaire (Song *et al*., 2015; Wallace et Galloway, 2015).

Par ailleurs, il a été observé qu'E6 pouvait se lier aux acides nucléiques, que ce soit l'ADN et les ARN (Bodaghi *et al*., 2009; Mallon *et al*., 1987; Ristriani *et al*., 2000), suggérant d'autres effets potentiellement plus directs sur l'expression des gènes.

- Immunité

L'effet d'E6 sur l'immunité a été précédemment décrit dans le chapitre 1.5.2. Echappement au système immunitaire et persistance de l'infection.

2. Régulation de l'expression des gènes des HPV-HR

L'expression des différentes protéines virales, à partir du génome d'HPV16, est réalisée par la machinerie de la cellule hôte et est soumise à une régulation à différents niveaux : transcriptionnel, post-transcriptionnel, traductionnel et post-traductionnel.

2.1. Régulation de la transcription d'HPV

Les HPV-HR possèdent plusieurs promoteurs permettant la transcription des gènes viraux. Il existe 2 promoteurs principaux bien décrits chez plusieurs HPV-HR : les promoteurs précoce (p97 chez HPV16) et tardif (p670 chez HPV16) situés sur la LCR et l'ORF E7 respectivement (Graham et Faizo, 2017). D'autres promoteurs ont été décrits notamment pour HPV16 parmi lesquels pE8, qui permet l'expression d'une protéine de fusion E8^E2c régulant négativement la transcription des gènes précoces et la réplication de l'ADN viral ; p14 et les sites TSS (Transcription Start Site) situés au niveau de nucléotides 441, 480 et 542 sur l'ORF E6 HPV16 (Braunstein *et al*., 1999; Fedorova *et al*., 2016; Rosenstierne *et al*., 2003).

2.1.1. Promoteur précoce

Le promoteur p97 fait partie de la LCR et son activité est régulée par celle-ci. La LCR serait en partie responsable du tropisme épithélial des HPV. En effet, placés sous le contrôle de la LCR, E6 et E7 sont uniquement exprimés au niveau de la peau et du col de l'utérus (épithélium) chez la souris, et ne sont pas exprimés dans d'autres tissus non épithéliaux (cerveau, foie, cœur, poumon) (Park *et al*., 2003).

2.1.1.1. Structure de la LCR

La LCR est composée de 850 nucléotides et contient de nombreux motifs, très conservés parmi les HPV, permettant la fixation de facteurs cellulaires et viraux impliqués dans la régulation de la réplication et de la transcription virales. En outre elle contient quatre E2BS et l'ORI (site de liaison de E1, E1BS) (Fig. 11). La LCR d'HPV16, tout comme celle des autres HPV-HR peut être subdivisée en trois régions distinctes (Fig. 13) :

(i) Une région 3' située entre l'E2BS #3 et l'ORF E6. Celle-ci comprend le promoteur p97, notamment une boite TATA permettant la fixation des facteurs généraux de la transcription TFIID et un site de fixation de l'activateur transcriptionnel SP1 (Specificity Protein 1). En plus d'être impliqués dans la reconnaissance de l'ORI, les E2BS notamment #1 et #2, permettent d'exercer un rétrocontrôle transcriptionnel sur p97 *via* le recrutement de la

protéine E2, qui induit un encombrement stérique empêchant la fixation de TFIID et SP1 sur leurs sites respectifs (Fig. 13). D'une façon similaire, la présence d'un site de liaison de YY1 (Yin Yang 1) sur un site chevauchant le TSS, permet d'inhiber l'expression à partir de p97 (Lace *et al*., 2009). Enfin, O'Connor et collaborateurs ont montré que la liaison de Cux1 (Cut like homeobox 1, appelé aussi CDP/cut), recruté au niveau de ses sites de fixation chevauchant l'ORI, réprime p97 (O'Connor *et al*., 2000).

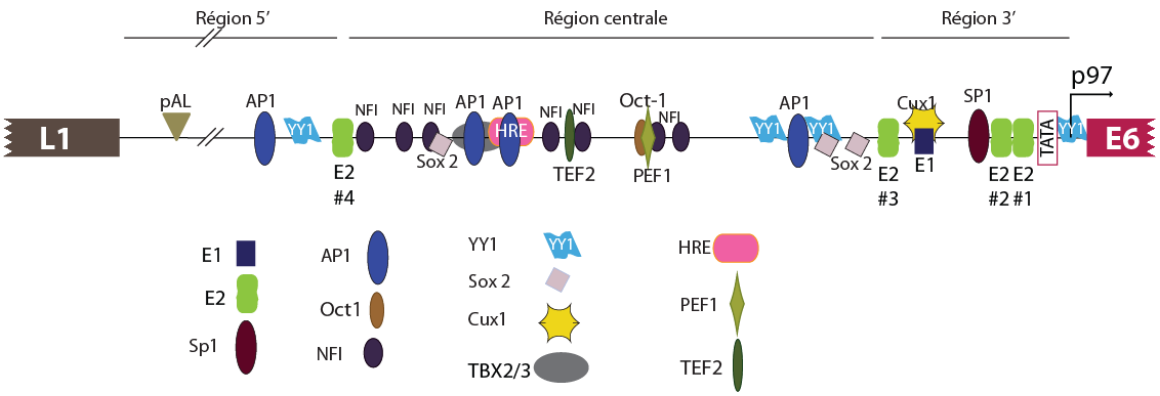


Figure 13. Structure de la LCR d'HPV16.

(ii) Une région centrale située entre les E2BS #3 et #4. Elle contient des enhancers et des silencers impliqués dans la stimulation et la répression de la transcription respectivement (Fig. 11 et Tableau II). En outre, les sites de liaison à NFI (Nuclear Factor I), AP-1 (Activator Protein 1) sont les activateurs ayant le rôle le plus important lors de la stimulation de la transcription à partir de p97. Cette région centrale serait le facteur influant sur le tropisme épithélial des HPV. En effet, l'enhancer contenu dans cette région serait actif au niveau des cellules épithéliales, qui sont permissives à l'expression des gènes viraux. Ainsi, NFI-X, le gène de la famille NFI le plus exprimé dans les cellules mésenchymateuses, inhibe p97 alors les gènes que NFI-A, -B et –C, exprimés dans les cellules épithéliales activent son expression. D'autres facteurs de transcription sont aussi capables de se fixer sur la région centrale de la LCR et d'inhiber ou d'activer p97 (Fig. 11 et Tableau II). Enfin, la présence d'un élément de réponse aux hormones, permet la régulation de l'expression des gènes viraux par des hormones sexuelles telles les œstrogènes qui augmentent l'expression d'E6 et E7 (Park *et al*., 2003).

(iii) Une région 5', délimitée par l'ORF codant L1 et un E2BS #4, qui contient le site pAL (polyadenylation late), des sites de liaison des facteurs de transcription AP-1 et YY1 (Fig. 11). L'activité trans-activatrice de cette région serait redondante avec celle de la région centrale (Bernard, 2013). De plus, les facteurs FOXA1 (Forkhead box A1) et MYC, sont

recrutés au niveau de la région 5'LCR d'HPV16, activant et inhibant l'activité de p97 respectivement (Bernard, 2013; Sichero *et al*., 2012).

Tableau II. Rôle des facteurs de transcription au niveau de la LCR d'HPV16.

Facteur de transcription		Effet
TATA		Activateur
SP1	Specificity Protein 1	Activateur
AP1	Activating Factor 1	Activateur
Oct-1	Octamer-Binding Transcription Factor-1	Activateur
NFI -A, -B, -C	Nuclear Factor I-A, -B, -C	Activateur
HRE	Hormon Responsive Element	Activateur
TEF1	Transcriptional Enhancer Factor 1	Activateur
PEF1	Penta-EF-hand domain containing 1	Activateur
YY1	Ying Yang 1	Represseur
SOX2	SRY (sex determining region Y)-box 2	Represseur
CEBPβ	CCAAT/enhancer binding protein beta	Represseur
TBX2	T-box 2	Represseur
TBX3	T-box 3	Represseur
CUX1	cut like homeobox 1	Represseur
NFI -X	Nuclear Factor 1	Represseur

Par ailleurs, l'utilisation des promoteurs (tels p14 et TSS 441, 480 et 542) par HPV-16 est extrêmement faible (Braunstein *et al*., 1999; Fedorova *et al*., 2016; Rosenstierne *et al*., 2003). De façon intéressante, les TSS situés sur E6 favorisent l'expression d'E7 à partir d'un transcrit monocistronique (Glahder *et al*., 2003). P14 permet l'expression de la protéine E1 (Fedorova *et al*., 2016).

Les HPV-BR possèdent aussi plusieurs promoteurs principaux situés sur la LCR, l'ORF E6 et l'ORF E7, qui permettent la synthèse des protéines E6 et E7 ainsi que les autres protéines précoces et tardives.

2.1.1.2. Régulation par la protéine E2

Via sa fixation sur les E2BS, la protéine E2 permet non seulement de réguler la réplication virale mais aussi la transcription des gènes précoces. Selon son niveau d'expression et son site de fixation, E2 peut avoir des effets activateurs ou inhibiteurs de la transcription du promoteur précoce. Ainsi, une expression faible d'E2 active la LCR alors qu'une expression élevée induit son inhibition (Ottinger *et al*., 2009). En outre, la fixation d'E2 sur les E2BS #3 et #4, ayant une séquence consensus très conservée, active la transcription. A l'inverse, sa fixation de façon moins forte sur les E2BS #1 et #2, qui ont une séquence plus dégénérée, réprime la transcription. Cette répression est particulièrement due à l'encombrement stérique engendrée

par la fixation d'E2, qui empêche la fixation de SP1 et TFIID sur leurs sites respectifs (Dong *et al*., 1994; Tan *et al*., 1994a).

Par ailleurs, l'action d'E2 sur la transcription au niveau de la LCR met aussi en jeu des facteurs tels SMC-C (Structural Maintenance of Chromosomes C) et le complexe TIP60, impliqués dans la modification de la structure de la chromatine (Smith *et al*., 2014). E2 favorise la mise en place de marques répressives de la chromatine telles que la méthylation des histones H3K3me3 (tri-méthylation de la lysine 3 de l'histone H3) au niveau de la LCR. En outre, la perte d'E2 est un facteur clé de la dérégulation de l'expression d'E6 et E7, étant notamment liée à l'intégration du génome viral dans le génome cellulaire.
D'autre part, la perte de la répression transcriptionnelle médiée par E2 pourrait aussi être dûe à la méthylation des E2BS, laquelle empêche la fixation d'E2 sur ses séquences cibles (Kim *et al*., 2003).

2.1.1.3. Régulation épigénétique

Comme l'ADN cellulaire, la transcription à partir du génome viral peut être régulée au niveau épigénétique en étant influencée par la méthylation de l'ADN et la structure de la chromatine. L'ADN des HPV est modifié par les DNMT (DNA methyl transferase), qui méthylent des séquences CpG, situées le long du génome, telles les E2BS. Plusieurs études se sont intéressées à la méthylation des séquences CpG, notamment celles situées sur la LCR et l'ORF L1, qui sont généralement méthylées dans les cancers (Johannsen et Lambert, 2013). C'est notamment lors de l'intégration en plusieurs copies du génome en concatémères (appelée intégration de type II comme dans les cellules Ca Ski), que les CpG de la LCR sont méthylées dans la majorité des copies. La méthylation de l'ADN étant principalement une marque répressive de l'expression, seules les quelques copies non méthylées sont exprimées. En effet, la méthylation au niveau de ces copies permet la mise en place de structure répressive de la chromatine inhibant aussi la transcription (De-Castro Arce *et al*., 2012). A l'inverse, lors d'une intégration en 1 ou 2 copies (intégration de type I comme dans les cellules SiHa), les LCR ne sont pas méthylées et permettent l'expression des gènes.

Le génome d'HPV16 se lie à des histones, non seulement lorsqu'il est intégré mais aussi lorsqu'il se trouve sous une forme épisomale. De façon intéressante, la position de 2 nucléosomes dans la LCR d'HPV16 lors de l'assemblage de la chromatine *in vitro* est similaire avec leurs positions au sein du génome intégré des cellules Ca Ski (Stünkel et Bernard, 1999). En outre, l'assemblage de ces nucléosomes réprime la transcription de p97 *in*

vitro. L'addition de AP-1 et notamment SP1 empêche cette répression induite par les nucléosomes *in vitro*, vraisemblablement en modifiant la structure de la chromatine (Stünkel et Bernard, 1999). En outre, la structure de la chromatine dans la LCR et autour de p742, permet de réguler l'activité des promoteurs viraux (You, 2010). En effet, lors de la différenciation épithéliale, Wooldwidge et Laimins ont observé une augmentation de la méthylation et de l'acétylation des histones H4 et H3 respectivement, lesquelles modifient la structure de chromatine, favorisant l'activité de 97 et p670 (Wooldridge et Laimins, 2008). De plus, la présence de modifications post-traductionnelles des histones est corrélée avec le niveau d'expression des gènes viraux dans les cancers. En effet, des marques de chromatine active enrichies dans la LCR sont corrélées à une expression accrue d'E6 et E7 à partir de génomes intégrés, alors que les marques répressives sont retrouvées dans les génomes ayant une faible expression de ces mêmes oncogènes (Groves *et al*., 2016).

L'expression génique est régulée non seulement par la structure de la chromatine mais aussi par l'organisation de celle-ci, particulièrement au sein de domaines nucléaires *via* sa fixation avec la matrice nucléaire. En outre, c'est *via* les MAR (Matrix Attachement Region) que l'ADN se lie avec la matrice nucléaire, permettant ainsi la réorganisation de la chromatine et impactant l'expression des gènes. Comme l'ADN cellulaire, les génomes des HPV à tropisme génital, possèdent des MAR, situés sur la LCR et le ORF E6 et E5 (Tan *et al*., 1998). De façon intéressante, les MAR d'HPV16 présentent un effet répresseur pour l'expression génique de génomes sous forme d'épisome, alors qu'ils activent la transcription des gènes à partir des génomes intégrés (Stünkel *et al*., 2000).

2.1.2. Promoteur tardif

Contrairement aux promoteurs des gènes précoces et pE8, actifs dans les cellules épithéliales indifférenciées, le promoteur des gènes tardifs n'est activé que lors la différenciation épithéliale. Ainsi, CEBPβ (CCAAT Enhancer Binding Protein Beta) et hSkn-1a/POU2F3 (POU Class 2 Homeobox 3), deux facteurs de transcription impliqués dans l'induction de la différenciation épithéliale, se fixent au niveau de la région E6 et E7 et activent p670 (Kukimoto et Kanda, 2001; Kukimoto *et al*., 2006). D'ailleurs, des expériences de ChIP (Chromatin ImmunoPrecipitation) ont montré une augmentation de la fixation de hSkn-1a au niveau de p670 au cours de la différenciation, laquelle empêche la fixation de YY1, induisant une levée de la répression exercée par ce dernier sur p670 (Kukimoto et Kanda, 2001; Wooldridge et Laimins, 2008). Outre la différenciation cellulaire, l'activité de p670 peut aussi

être régulée par le virus. En effet, Bodily et collaborateurs ont montré que la protéine E7 induit l'activation de p670 (Bodily *et al*., 2013).

Par ailleurs, l'activité de p670 est non seulement régulée au niveau de l'initiation de la transcription mais aussi de son élongation. En effet, Songock et collaborateurs ont montré que la différenciation épithéliale favorise l'élongation de la transcription à partir de p670, en induisant une augmentation de la phosphorylation de l'ARN polymérase II (Songock *et al*., 2017b). Ils ont observé une augmentation des protéines Cdk8 et Brd4 au niveau de la LCR, de p670 et des ORF E1 et E4 lors de la différenciation. Ces protéines permettent la phosphorylation de l'ARN polymérase II (Songock *et al*., 2017b). De plus, non seulement la région E7 mais aussi la LCR sont impliquées dans cette activation du promoteur tardif.

Enfin, une récente expérience menée sur HPV18 a montré que l'orientation de l'origine de réplication affectait l'activité du promoteur tardif (Wang *et al*., 2017). Il se pourrait donc que l'activité du promoteur tardif soit non seulement régulée par des facteurs de transcriptions, mais aussi par le mode de réplication du génome viral, lequel *switch* d'un mode bidirectionnel à un mode en cercle roulant lors de la différenciation cellulaire.

La transcription par le promoteur tardif produit des ARNm (ARN messagers) polycistroniques codant les protéines E1, E4, E5 et parfois L1 et L2, ce qui ne favorise pas la production des protéines de la capside. En effet, l'expression de L2 et L1 n'a lieu que lors de l'utilisation de site pAL (late polyadenylation site) et à l'épissage.

2.2. Régulation post-transcriptionnelle de l'expression gènes viraux

Comme les ARN cellulaires, les ARNm viraux subissent des modifications posttranscriptionnelles telles que l'addition de la coiffe en 5', l'épissage et l'ajout d'une queue poly-adénine. D'ailleurs, les ARNm d'HPV16 sont coiffés, épissés et polyadénylés (Zheng *et al*., 2004). L'épissage et la polyadénylation sont deux processus permettant d'engendrer une certaine diversité au niveau du transcriptome et du protéome cellulaire et virale (Voir les chapitres ci-dessous). Ces modifications influencent aussi le devenir des ARNm : leur stabilité, leur export du noyau vers le cytoplasme ainsi que leur traduction. Par ailleurs, différentes protéines se lient aux ARNm cellulaires et viraux, régulant non seulement leur devenir mais aussi leurs modifications posttranscriptionnelles.

2.2.1. Epissage de l'ARN et polyadénylation

2.2.1.1. L'épissage des ARN

L'épissage est un processus co-transcriptionnel de maturation des pré-ARNm. Il consiste à éliminer les introns (parties non codantes) pour joindre les exons (parties codantes) et former l'ARNm mature. L'épissage des pré-ARNm met en jeu 2 réactions de transestérification catalysées par le spliceosome, un complexe ribonucléoprotéique composé de 5 snRNP (small nuclear Ribonucleoprotein) U1, U2, U4, U5 et U6 et de protéines non-snRNP (Cartegni *et al*., 2002; Wahl *et al*., 2009). Chaque snRNP est composé d'un petit ARN non codant riche en uridine (snRNA U1, U2, U4, U5 ou U6) et de plusieurs protéines.

La reconnaissance des introns et des exons par le spliceosome, se fait au niveau de séquences consensus présentes au sein des pré-ARNm : les sites 5' donneur (SD, Splice Donnor) et 3' accepteur (SA, Splice Acceptor) d'épissage ainsi que le site de branchement (Fig. 14). Le site SD, qui contient un dinucléotide GU, est reconnu par snRNP U1 notamment par appariement de bases avec le snRNA U1. Le site de branchement est lié par snRNP U2, alors que les protéines U2AF35 et U2AF65 se lient aux nucléotides AG du site SA et au bloc de pyrimidine situés en amont respectivement (Fig. 14).

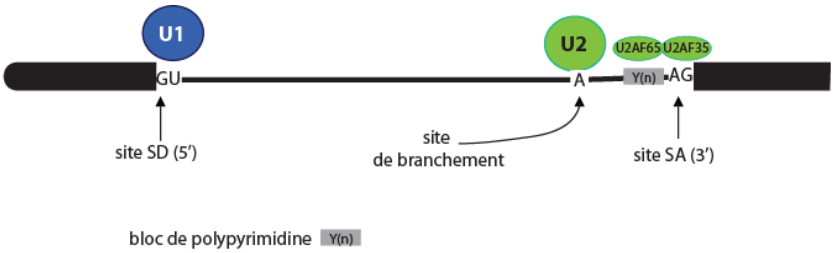


Figure 14. Les sites canoniques d'épissage. (Adaptée de Wahl *et al*., 2009).

La reconnaissance des sites d'épissage et le recrutement des snRNP U1 et U2 au niveau de l'intron forme le complexe A du spliceosome (Fig. 15). Le recrutement subséquent du tri-snRNP U4/U5/U6 permet la formation du complexe pré-catalytique B. Celui-ci subit plusieurs réarrangements dans les liaisons ARN-ARN et ARN protéines, aboutissant à la libération des snRNP U1 et U4, formant ainsi le complexe B*. La 1ère réaction de transestérification, catalysée par le complexe B*, permet la liaison du premier nucléotide de l'intron avec l'adénine du site de branchement libérant l'extrémité 3' de l'exon 1 (Wahl *et al*., 2009; Will et Lührmann, 2011). A la fin de cette réaction, d'autres changements mineurs au niveau du complexe B* aboutissent à la formation du complexe C qui catalyse la 2ème réaction assurant la ligature des exons et la libération de l'intron sous forme de lariat (Fig. 15).

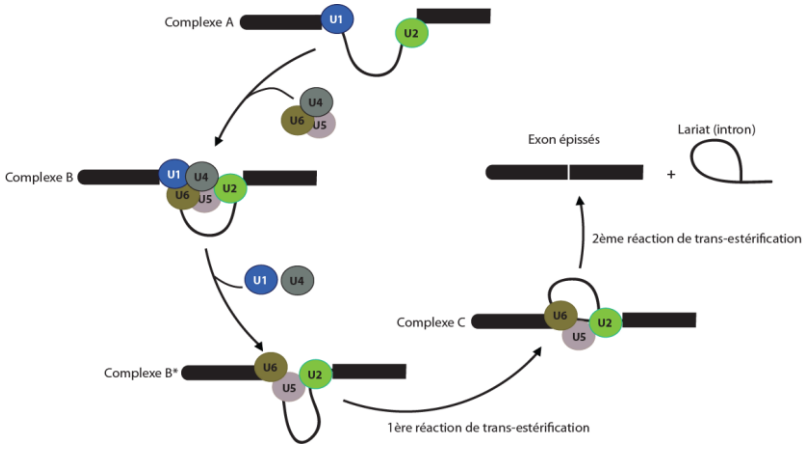


Figure 15. Le processus de l'épissage. (Adaptée de Wahl *et al.*, 2009).

Par ailleurs, la dégénérescence des séquences consensus des sites canoniques d'épissage fait que ceux-ci ne sont pas reconnus ni utilisés avec la même efficacité par le spliceosome. Ainsi, l'utilisation des sites d'épissage de façon différentielle donne lieu à l'épissage alternatif, lequel permet de former plusieurs ARNm à partir d'un seul pré-ARNm, et favorise ainsi l'augmentation de la diversité du transcriptome et du protéome. Différents modes d'épissage alternatif ont été décrits, notamment l'utilisation de sites SD et/ou SA alternatifs, les exons mutuellement exclusifs, l'inclusion/exclusion d'exon et la rétention d'intron (Fig. 16).

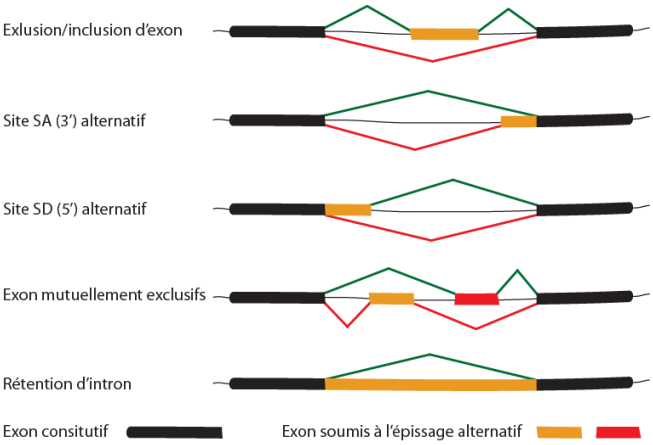


Figure 16. Le processus de l'épissage altérnatif. (Adaptée de Cartegni *et al.*, 2002).

2.2.1.2. Régulation de l'épissage alternatif

Selon le degré de conservation des séquences consensus, les sites d'épissage ne présentent pas la même force quant à leur possible utilisation par le spliceosome. Ainsi, les sites d'épissage sont en concurrence pour le recrutement des éléments du spliceosome. De plus, plusieurs facteurs affectent la façon dont sont reconnus les sites d'épissage. Parmi eux, la présence de séquences *cis*-régulatrices dans les séquences des pré-ARNm, qui sont reconnues par des facteurs régulateurs de l'épissage qui influencent le recrutement du spliceosome (Cartegni *et al.*, 2002; Will et Lührmann, 2011).

Selon qu'ils recrutent des facteurs activateurs ou inhibiteurs de l'épissage, ces éléments sont respectivement nommés ESE (Exonic Splicing Enhancer) et ESS (Exonic splicing Silencer) lorsqu'ils sont localisés dans les exons, ISE (Intronic Splicing Enhancer) et ISS (Intronic Splicing Silencer) correspondants à leur localisation intronique (Fig. 17). Les ESE et ISE recrutent particulièrement des protéines de la famille SR (Serine Argenine Rich). Cette famille de protéines comprend une dizaine de membres tels SRSF1 (Serine and Argenine rich splicing factor 1, appelé aussi SF2/ASF), SRSF2 (appelé aussi SC35), SRSF3, lesquels favorisent la reconnaissance des sites canoniques par le spliceosome et donc l'épissage (Änkö, 2014; Wang et Burge, 2008). En outre, elles participent à la reconnaissance du petit exon entouré par de longs introns, processus appelé « Exon definition » et ce, en permettant une communication entre les sites SD et SA de l'exon. Inversement, les ESS et ISS sont reconnus par des protéines de la famille hnRNP (heterogeneous nuclear Ribonucleoprotein), qui inhibent le recrutement du spliceosome et/ou antagonise l'effet des protéines SR, inhibant ainsi l'épissage ((Wang et Burge, 2008) ; Fig. 17). La famille hnRNP étant plus large que celle des SR, elle comprend plus de 30 protéines telles hnRNP A1 et hnRNP A2.

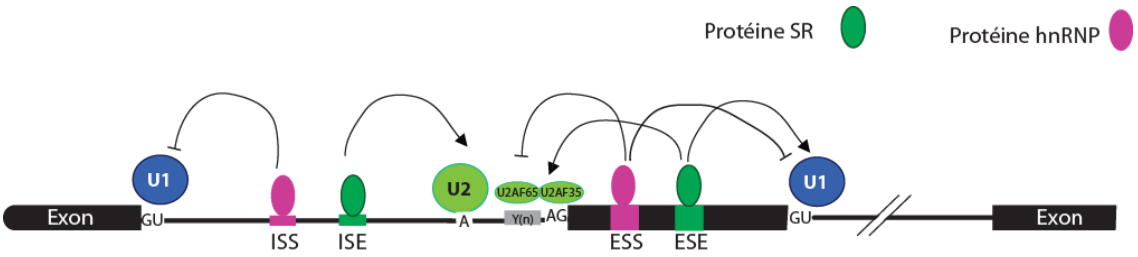


Figure 17. La régulation de l'épissage par les protéines SR et hnRNP. (Adaptée de Cartegni *et al.*, 2002; Wang et Burge, 2008).

Par ailleurs, l'accessibilité de sites *cis*-régulateurs peut être régulée par le degré d'affinité avec différents facteurs *trans*-régulateurs mais aussi par le niveau d'expression de ces derniers. De plus, la structure secondaire des pré-ARNm affecte la fixation de tous types de facteur l'épissage, influençant le mode d'épissage et l'ARNm produit. Enfin, la vitesse de la transcription et la structure de la chromatine affectent également l'épissage (Luco *et al.*, 2011).

2.2.1.3. La polyadénylation

La polyadénylation consiste en l'addition d'une queue poly(A) à l'extrémité 3' des ARNm. Elle se déroule en deux réactions : un clivage de l'ARNm et l'addition de poly(A) par l'enzyme PAP (poly(A) polymérase). En plus de PAP, la polyadénylation fait intervenir plusieurs complexes protéiques, dont le recrutement se fait *via* des séquences consensus

entourant le site clivage. Le complexe CPSF (Cleavage and Polyadenylation Specificity Factor) qui contient l'endonucléase responsable du clivage, reconnait la séquence AAUAAA, la plus conservée parmi les sites pA et située en amont du site de clivage (Di Giammartino et Manley, 2014; Shi et Manley, 2015). Fip1 (Pre-mRNA 3'-end-processing factor FIP1), une sous-unité de CPSF se lie à une séquence riche en U située en amont du motif AAUAAA, renforçant la fixation de CPSF. La fixation de CstF (Cleavage stimulatory Factor), sur la séquence riche en G/U en aval du site de clivage, renforce la fixation de CPSF et stimule le clivage. Enfin, une séquence consensus UGUA présente en amont d'AAUAAA est reconnue par CFI (human Cleavage Factor Im). La polyadénylation est aussi régulée par la fixation d'autres protéines tels hnRNP H et HuR (Hu Antigen R, appelé aussi ELAVL1), qui selon leur site de fixation, stimulent ou inhibent la polyadénylation (Shi et Manley, 2015). Enfin, le recrutement des snRNP au niveau de sites d'épissage inhibe la polyadénylation alors que le snRNP U2 l'active (Shi et Manley, 2015).

2.2.1.4. Epissage et polyadénylation des transcrits d'HPV16

Les transcrits d'HPV16 (et des HPV en général) possèdent 2 sites pA : un site précoce pAE (Early pA) et un site tardif pAL (Late pA), ainsi que plusieurs sites d'épissage dont l'utilisation génère un minimum de 20 transcrits alternatifs (Fig. 18).

Les sites d'épissage situés dans les ORF E1, E2, E4 et L1/L2, sont très conservés au sein de la famille des HPV y compris avec les PV, indiquant l'importance de l'épissage chez ces virus. Cependant, ceux situés dans les ORF E6 et E7 sont uniquement retrouvés chez les HPV-HR. En outre, les 2 sites les plus conservés au sein des HPV-HR, SD226 et SA409 dans le cas d'HPV16, permettent la production d'un ARNm E6*I *via* l'excision de l'intron 1. L'absence d'épissage produisant E6*I chez les HPV-BR suggère son implication dans la carcinogenèse HPV-induite. De façon intéressante, les ARNm E6*I représentent la majorité des transcrits E6 produits par les cellules infectées et/ou transformées, tandis que les transcrits E6Fl (Full length), qui codent la protéine E6, représentent moins de 5% des transcrits viraux (Chen *et al*., 2014a). En somme, l'épissage produisant des ARNm E6*I jouerait un rôle dans le cycle viral et la carcinogenèse, mais ce rôle fait l'objet de débats. Les fonctions potentielles de l'épissage E6*I et des isoformes produites par épissage alternatif au sein de l'ORF E6 sont abordées dans le paragraphe le rôle d'E6*I. Les sites SA526 et SA742 codant potentiellement des protéines E6*II et E6^E7 respectivement, sont uniquement présents chez HPV16. Toutefois, de nombreux HPV-HR contiennent des sites similaires à SA742 et sont capables de produire

E6^E7, tels HPV18 dont l'épissage SD233 et SA791 a été identifié dans les cellules HeLa (HPV18 positive) (Ajiro et Zheng, 2015).

Au niveau des ORF E1, E2 et E4, l'utilisation du site SD880 avec les sites alternatifs SA2582, SA2709 et SA3358, permet de produire les ARNm codant des protéines E1C, E2 et E1^E4 respectivement. En outre, SD880 et SA3358 sont les sites les plus utilisés par les HPV-HR et -BR (Chen *et al*., 2014a; Schmitt *et al*., 2011). Par ailleurs, le site SD1302 est très peu/pas utilisé par les pré-ARNm issus de p97 et p670, vraisemblablement à cause d'une compétition avec le site SD880. D'ailleurs, au sein des transcrits produits à partir de pE8, qui ne contiennent pas SD880, le site SD1302 est utilisé.

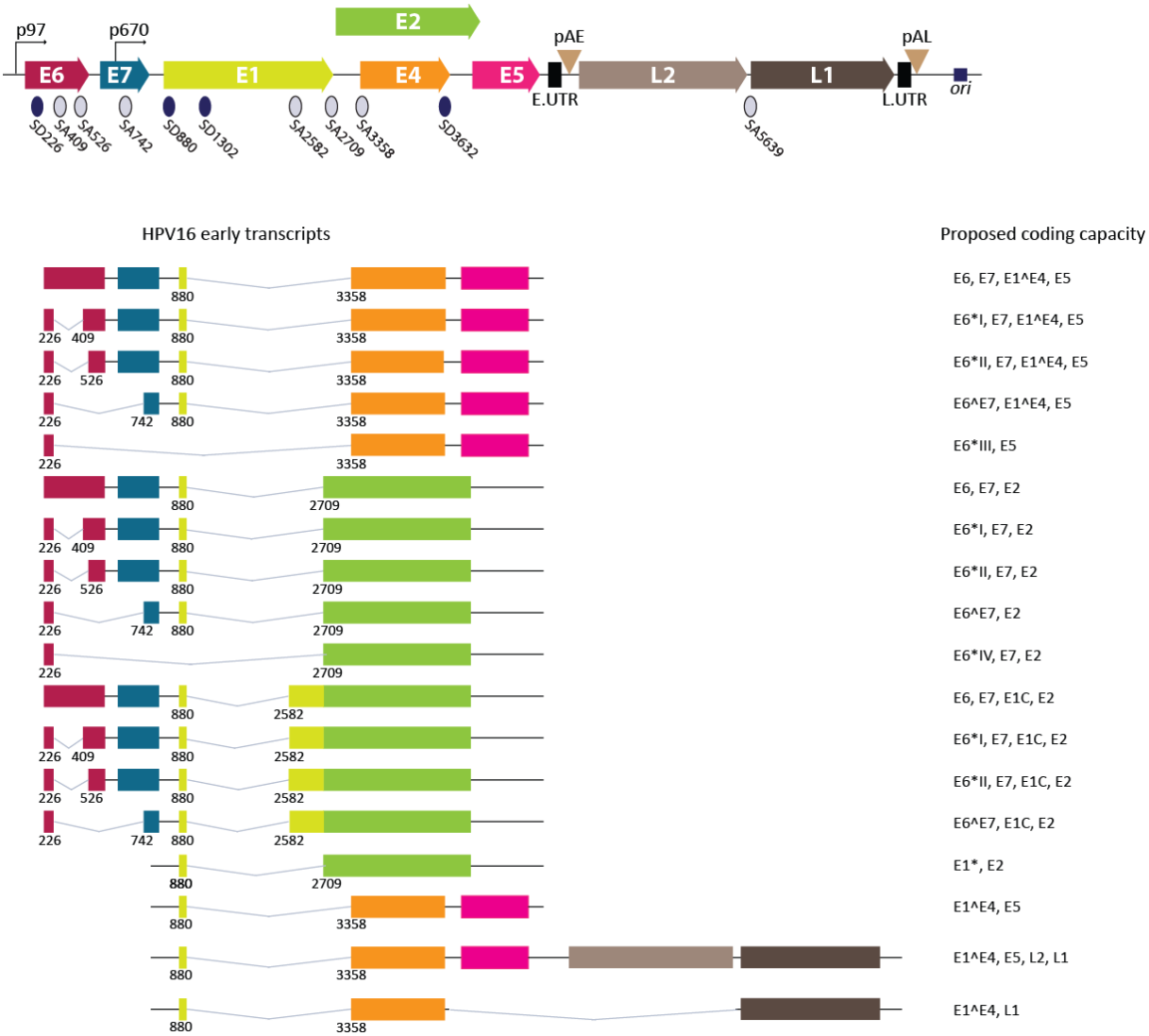


Figure 18. Epissage et polyadénylation alternatifs des ARNm d'HPV16. (Adaptée de Johansson et Schwartz, 2013; Zheng et Baker, 2006).

Enfin, les sites SD3632 et SA5639 sont utilisés uniquement pour l'épissage des pré-ARNm tardifs, qui permettent l'expression de L1 et L2. Ces transcrits sont polyadénylés au site pAL,

qui est utilisé uniquement lors de la dernière phase du cycle viral. L'utilisation du site pAE permet l'expression des gènes précoces des HPV et empêche l'expression de L1 et L2.

2.2.1.4.1. Régulation de l'épissage et polyadénylation des transcrits d'HPV16

La régulation de l'épissage au niveau de l'ORF E6 est moins bien caractérisée que celle des régions E1, E2, L1 et L2. Ainsi, les protéines hnRNP A1 et hnRNP A2 favorisent l'exclusion de l'intron 1, alors que Sam68 et Brm induisent la diminution de la production d'E6*I (Fig.19A ; (Rosenberger *et al*., 2010)). Toutefois, les sites de fixation de hnRNP A1 et Sam68 au niveau des régions E6 et E7 n'ont pas encore été caractérisés. En revanche, chez HPV18, hnRNP A1 se fixe sur un ESS situé dans l'ORF E7 et inhibe l'épissage et la production d'E6*I (Ajiro *et al*., 2016). Rosenberger et collaborateurs ont aussi montré que l'épissage d'E6 était régulé par la voie de signalisation de l'EGF, laquelle favorise la rétention de l'intron et par conséquent la production d'E6Fl (Rosenberger *et al*., 2010). De fait, plusieurs protéines telles Sam68 sont impliquées dans la régulation de l'épissage induite par la transduction du signal (Graham et Faizo, 2017). En outre, la présence de la coiffe, permettant le recrutement des protéines du CBC (Cap binding complex) stimule l'épissage au sein de l'ORF E6 *in vitro* (Zheng *et al*., 2004). Brm est un composant du complexe de remodelage de chromatine SWI/SNF (SWItch/Sucrose Non-Fermentable), qui pourrait modifier l'élongation de la transcription, et affecter l'épissage de la région E6E7. D'ailleurs, CTCF (CCCTC-Binding Factor), un facteur de transcription impliqué dans la régulation de l'épissage et de la transcription, modifierait l'épissage d'E6 d'HPV18. En outre, CTCF induit une pause de l'ARN pol II au niveau de l'ORF E2, favorisant l'épissage d'E6E7 (Paris *et al*., 2015).

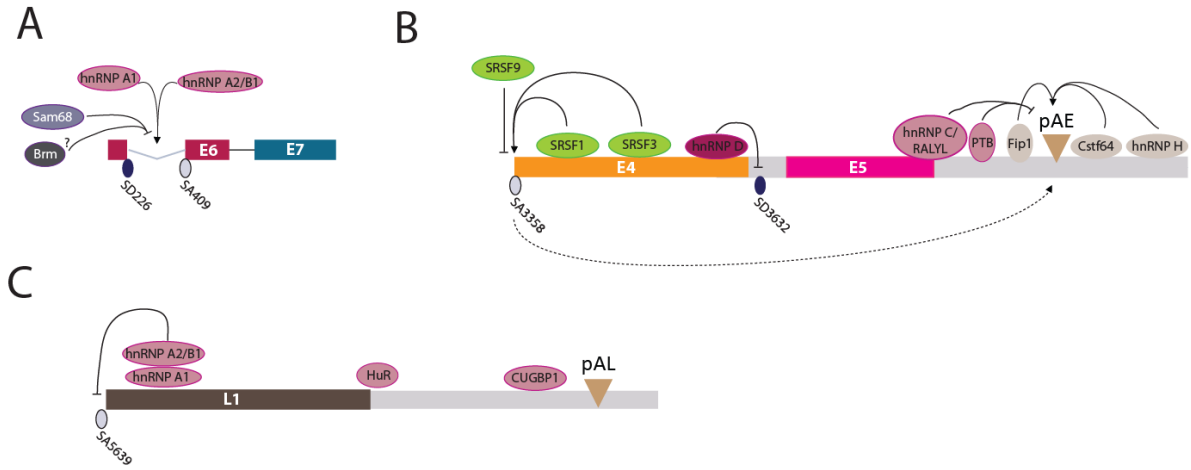


Figure 19. Régulation de l'épissage de l'expression des transcrits précoces et tardifs d'HPV16.

Par ailleurs, avec une approche de chromatographie d'affinité et de spectrométrie de masse, Martinez-Salazar et collaborateurs ont identifié environ 150 protéines se liant au pré-ARNm E6E7. Ils y ont notamment identifié des protéines du spliceosome (snRNPs) mais aussi des protéines hnRNP (hnRNP A1, hnRNP A2/B1, hnRNP C1/C2 …) ainsi que Sam68 (Martínez-Salazar *et al*., 2014). D'autres protéines faisant partie de complexes mRNP (messenger Ribonucleoprotein), tels eIF4A3 (eukaryotic translation initiation factor 4A3), RNPS1 (RNA Binding Protein with Serine Rich Domain 1), DEK (DEK Proto-Oncogene) et EWS (Ewing Sarcoma Breakpoint Region 1 Protein) ont également été identifiées. Ces protéines sont retrouvées dans le complexe B et C du spliceosome, et font partie de l'EJC (Exon Junction Complex), un complexe multiprotéique recrutés sur les ARNm *via* l'épissage et impactant le devenir de ces derniers. Les rôles de l'EJC sont abordés dans le chapitre 3.

Les mécanismes de régulation de l'épissage des transcrits précoces au niveau des ORF E4 et E5 et le site pAE sont assez bien caractérisés. En effet plusieurs ESE et ESS se liant à des protéines de la famille SR et hnRNP ont été identifiés (Fig. 19B). SRSF1 et SRSF3 se lient à l'ORF E4 d'HPV16, activent le site SA3358 et permettent l'expression de la protéine E4 (Jia *et al*., 2009; Li *et al*., 2013a; McPhillips *et al*., 2004; Somberg et Schwartz, 2010). De plus, hnRNP D et SRSF1 inhibent le site SD3632 utilisé par les ARNm codant L1 et L2, régulant ainsi l'expression des protéines tardives (Fig. 19B ; (Li *et al*., 2013b)). En favorisant l'utilisation du site SA3358 et en inhibant SD3632, les protéines SRSF1, SRSF3 et hnRNP D permettent l'utilisation de pAE inhibant de fait l'expression des gènes tardifs L1 et L2 (Jia *et al*., 2009; Li *et al*., 2013a; McPhillips *et al*., 2004; Somberg et Schwartz, 2010). D'un autre côté, SRSF9 qui inhibe le site SA3358 et active SA5639, permet l'expression des gènes tardifs (Somberg *et al*., 2011). En effet, l'utilisation de SD3632, ou l'utilisation d'autres sites SD en amont permet l'expression de L1 et L2 (Dhanjal *et al*., 2015). D'ailleurs, PTB (Polypyrimidin Tract Binding Protein), hnRNP C1, et RALYL (RNA-Binding Raly-Like Protein) se lient à la région 3'UTR (Untranslated Region) précoce et à la région E4 en amont de SD3632, et induisent l'utilisation de SD3632, l'inhibition de pAE permettant ainsi l'expression de L1 et L2 (Fig. 19B ; (Dhanjal *et al*., 2015)).

Par ailleurs, la protéine E2 se lie à des protéines impliquées dans la polyadénylation (CstF64, Fip1, hnRNPH), inhibe l'utilisation du site pAE, favorisant ainsi l'utilisation de pAL et l'expression des ARNm L1 et L2 (Fig. 19B ; (Johansson *et al*., 2012; Zhao *et al*., 2005)). De plus, le facteur hnRNP H qui active pAE en se liant en amont est réprimé durant la différenciation cellulaire, ce qui permet de coupler l'expression des gènes tardifs avec la

différenciation (Oberg *et al*., 2005). Enfin, le site SA5639 permettant l'expression de L1 est régulé par un ESS liant les protéines hnRNP A1 et hnRNP A2/B1, permettant d'affiner le contrôle de l'expression de L1 et L2 (Fig. 19C ; (Zhao *et al*., 2004)).

De façon intéressante, l'infection virale et l'expression subséquente des protéines des HPV-HR modifient l'expression des protéines impliquées dans la régulation des ARNm. En particulier, la protéine E2 augmente l'expression de SRSF1, SRSF2 et SRSF3, permettant d'augmenter l'expression des ARNm E6E7 et E4^L1 (Klymenko *et al*., 2016). De plus, l'expression de SRSF1 et hnRNA A1 lors de la différenciation cellulaire, est augmentée lors de l'infection par HPV16, permettant à ce dernier de mieux contrôler l'expression de ses propres gènes (Cheunim *et al*., 2008). Enfin, une augmentation de l'expression de nombreuses protéines RBP (RNA binding protein) se liant aux ARN, notamment celles de la famille SR et hnRNP, est corrélée avec le développement des cancers du col de l'utérus (Fay *et al*., 2009). Ces résultats pointent le doigt sur **l'importance des RBP dans la régulation de l'expression des protéines virales**, mais aussi sur un **rôle potentiel de la dérégulation des RBP dans les cancers induits par les HPV-HR**.

2.2.2. Régulation de la stabilité des transcrits viraux

Afin de limiter l'expression de leurs protéines les plus immunogènes (notamment L1 et L2), les HPV-HR ont renforcé le contrôle de leur expression par une régulation de la stabilité de leurs transcrits. Ainsi, les HPV-HR limitent l'expression des protéines de capside L1 et L2 en : (i) inhibant le promoteur tardif et (ii) en favorisant l'utilisation de pAE, mais aussi (iii) en régulant la stabilité des transcrits contenant les ORF L1 et L2. En effet, un élément chevauchant la région 3' de l'ORF L1 et la région UTR, déstabilise les transcrits viraux (Fig. 19C). Cet élément permet le recrutement des protéines telles CUGBP (CUG Triplet Repeat RNA-Binding Protein 1) qui favorisent la dégradation des ARNm d'HPV16 et HuR qui induit leur stabilisation (Cumming *et al*., 2009; Goraczniak et Gunderson, 2008; Graham, 2010). De plus, la stabilité des transcrits précoces est aussi régulée par la région 3'UTR précoce (E5_pAE), qui induit leur déstabilisation (Jeon et Lambert, 1995). De façon intéressante, la même équipe a montré qu'une perte de cette région, survenant lors de l'intégration du génome viral, induit une augmentation de la stabilité des transcrits d'HPV16 codant les oncoprotéines E6 et E7 (Jeon et Lambert, 1995).

Par ailleurs, McFarlane et collaborateurs ont observé que la protéine SRSF2, induit une stabilisation des transcrits E6E7 d'HPV16 dans des cellules W12, cellules non tumorigènes

issus d'une lésion CIN2 (McFarlane *et al*., 2015). Or, le développement d'un phénotype tumoral par les W12, lequel lié à une augmentation de l'expression des transcrits E6E7, est corrélé à une augmentation de l'expression de SRSF2. Les auteurs ont suggéré que SRSF2 fournissait une protection aux ARNm E6E7 d'HPV16, face à une dégradation par le NMD (Nonsense-mediated mRNA decay), mécanisme de dégradation des ARNm non-sens dont le mécanisme sera développé dans le chapitre 3.

Enfin, il a récemment été montré que les transcrits E6E7 d'HPV16 étaient ciblés par plusieurs miRNA cellulaires, tels miR-375, miR-122 (He *et al*., 2014; Jung *et al*., 2014; Morel *et al*., 2017). Ces miRs qui inhibent l'expression des protéines E6 et E7 et dégradent les ARNm E6E7, sont sous-exprimés dans les cancers induits par les HPV-HR.

2.2.3. Traduction des ARNm viraux

Les ARNm viraux des HPV-HR sont traduits en protéines par la machinerie de la cellule hôte. Chez les eucaryotes, la traduction est dépendante de la coiffe en 5' *via* laquelle sont recrutés les complexes PIC (pre-initiation complex) de la traduction. Le PIC 43S, composé notamment de la sous-unité ribosomale 40S, d'eIF3 (eucaryotic Initiation Factor 3), eIF1, eIF1A, eIF5 et du complexe ternaire eIF2-GTP-ARNt méthionine, scanne l'ARNm dans le sens 5'-3' à la recherche du codon d'initiation de la traduction, le codon AUG (Hinnebusch et Lorsch, 2012). La reconnaissance de l'AUG (souvent le 1er AUG rencontré) induit un remodelage du complexe PIC-mRNP, lequel abouti à l'hydrolyse du GTP par eIF2, au recrutement de la sous unité ribosomale 80S, à l'initiation de la traduction et à l'insertion de la méthionine en position une dans le polypeptide naissant.

Le recrutement du PIC sur les ARNm et le scan des ARNm requierrent aussi la présence du complexe eIF4F, fixé sur la coiffe des ARNm et composé des facteurs eIF4A, eIF4G, et eIF4E. Globalement, c'est eIF4E qui reconnait la coiffe et qui permet le recrutement des autres composants tandis qu'eIF4A, une ARN hélicase permet de dérouler les structures secondaires présentes dans les ARNm facilitant ainsi la recherche d'AUG. Quant à la protéine d'échafaudage, eIF4G, elle interagit avec PABP et eIF4E pour circulariser l'ARNm.

Ce mode de traduction présent chez les eucaryotes, possédant majoritairement des ARNm monocistroniques, est approprié pour traduire une seule ORF par ARNm. En revanche, il n'est pas approprié pour la traduction d'ARNm polycistroniques tels ceux des HPV-HR et, particulièrement à la traduction des ORF situés en deuxième position, tels E7, E1, E5, L1 et

L2. Dès lors, les virus ont utilisé différents mécanismes afin de contourner ces limites. Par exemple, les HPV-HR utilisent des mécanismes de réinitiation de la traduction et/ou de *leaky scanning* ou le saut du ribosome pour exprimer leurs protéines.

Le *leaky scanning* est l'ignorance d'un codant d'initiation par le PIC lors du scan. Le degré de reconnaissance du codon d'initiation par le PIC est influencé par la présence d'une séquence Kozak G/**A**cc**AUG**G entourant l'AUG (Kozak, 2002). De fait, un AUG situé dans un faible contexte de Kozak aura tendance à être ignoré par le PIC, favorisant de fait la traduction à partir d'un autre AUG, situé en aval et ayant un fort contexte. Par ailleurs, la diversité du transcriptome générée par l'utilisation d'épissage et de sites TSS alternatifs par les HPV-HR, produit des ARNm ayant différentes régions 5', permettant un positionnement différent des ORF et par conséquent la traduction de leurs différentes protéines. De cette manière, il a été montré que le *leaky scanning* de l'AUG d'E6, permet la traduction de l'ORF E7 située en aval et donc l'expression d'E7 à partir des ARNm E6Fl, E6*I et E6*II des HPV-HR (del Moral-Hernández *et al*., 2010; Stacey *et al*., 1995, 2000; Tan *et al*., 1994b). D'après Nilsson et collaborateurs, cela permet aussi l'expression de la protéine E5 à partir des transcrits E6*III épissés de SD226 à SA3358 (Nilsson *et al*., 2018).

D'autre travaux suggèrent au contraire que l'expression de la protéine E7 se fait uniquement à partir des ARNm E6*I *via* la réinitiation de la traduction (Ajiro *et al*., 2012; Tang *et al*., 2006; Zheng *et al*., 2004). En effet, à la terminaison de la traduction, les sous unités ribosomales se séparent et se dissocient de l'ARNm. Dans certains cas, la petite sous-unité 40S reste fixée sur l'ARNm, scanne vers la région en aval et induit une réinitiation permettant la traduction d'une ORF supplémentaire. La réinitiation est en outre favorisée suite à la traduction d'une petite ORF (40aa<) mais aussi par l'augmentation de la taille de la région intercistronique (Kozak, 2001, 2002). Ainsi, la traduction d'une ORF E6*I (43aa), ayant un espace intercistronique de 144 nucléotides avec l'ORF E7, favorise la réinitiation et la traduction d'E7 d'HPV16. A l'inverse, l'expression d'E7 à partir des ARNm E6Fl d'HPV16 (ORF E6 de 150 ; espace intercistronique de 2 nucléotides) ne permet pas/peu l'expression d'E7 *via* la réinitiation de la traduction. Enfin, d'autres TSS permettant d'exprimer des ARNm dont E7 ou L2 sont situés en première position, ont été identifiés par quelque études, suggérant une initiation canonique de leur traduction (Braunstein *et al*., 1999; Graham et Faizo, 2017; Rosenstierne *et al*., 2003).

Par ailleurs, la traduction des protéines des HPV-HR est régulée par la présence de séquences *cis*-régulatrices au sein des ARNm viraux et par la fixation de protéines *trans*-régulatrices. En

effet, il a été montré que des séquences situées dans les régions précoces (ORF E4 et E5) et tardives (ORF L2, 3'UTR), inhibaient la traduction des ARNm viraux (Collier *et al*., 1998; Sokolowski *et al*., 1998; Vinther *et al*., 2005; Wiklund *et al*., 2002). En particulier, il a été montré *in vitro* que les protéines hnRNP K, PCBP1 et PCBP2 (poly(rC) binding protein) se fixent à l'ORF L2 et inhibent sa traduction (Collier *et al*., 1998 ; (Sokolowski *et al*., 1998)).

Enfin, l'expression des protéines virales peut aussi être régulée par l'utilisation de codons rares et la disponibilité des ARNt correspondants. Les ARNm des HPV-HR utilisent souvent des codons rarement utilisés par la cellule hôte et par conséquent, sont traduits avec une efficacité plus faible que les transcrits cellulaires, limitant l'expression des protéines virales (Zhao *et al*., 2003; Zhou *et al*., 1999). De manière intéressante, la différenciation épithéliale est suivie d'un changement en disponibilité des ARNt correspondants aux codons utilisés par les HPV-HR, permettant de fait une meilleure expression des protéines L1 et L2 dans les couches épithéliales apicales (Gu *et al*., 2007). Une fois produites, l'activité et la fonction des protéines virales, notamment E6 et E7, peuvent être entre autre régulées par phosphorylation (Roman et Munger, 2013; Vande Pol et Klingelhutz, 2013).

2.2.4. Rôle de l'épissage alternatif des ORF E6 et E7

L'épissage dans les ORF E6 et E7 aurait une importance dans la carcinogenèse étant donné leur absence chez les HPV-BR. D'ailleurs, le taux de transcrits E6*I et E6*II est corrélé avec les différents stades de carcinogenèse HPV-induite (Cricca *et al*., 2009; Kösel *et al*., 2007; Pastuszak-Lewandoska *et al*., 2014; Schmitt *et al*., 2010). En revanche, les rôles fonctionnels de l'épissage et des transcrits alternatifs, notamment E6*I, dans la carcinogenèse n'ont cessé d'être questionnés.

D'une part, ces ARNm E6*I pourraient permettre l'expression de la protéine E7 (Fig. 20 ; voir ci-dessus), impliquant une régulation de la balance E6/E7 *via* l'épissage. Le site SD226 (dans le cas d'HPV16), utilisé pour produire les différentes protéines isoformes, a été aussi montré comme essentiel à l'expression des protéines E1 et E5 (Lace *et al*., 2008; Nilsson *et al*., 2018). D'autre part, l'épissage au sein des ORF E6 et E7, permettrait de générer des ARN fonctionnels et/ou des ORF alternatives fonctionnelles. En outre, l'intron éliminé de l'ARNm E6*I des HPV-HR, présente une taille de 183 nucléotides similaire à celle des introns produisant des miRNA (150-250 nucléotides), et pourrait de fait permettre la synthèse de miRNAs viraux (Fig. 20). Toutefois, et même si plusieurs miRNAs encodés potentiellement par HPV16 et HPV18 ont été prédits dans plusieurs régions génomiques (LCR, E6, E7, E1,

L2), aucun miRNA encodé par les ORF E6 et E7 n'a pu être validé (Gu *et al*., 2011; Weng *et al*., 2017).

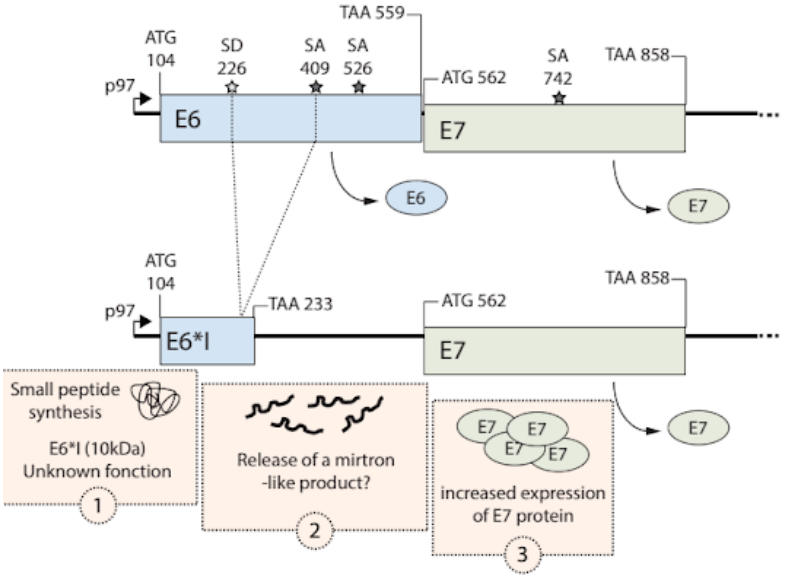


Figure 20. Rôles potentiels des ARNm E6*I.

Du reste, les ARNm E6*I, E6*II et E6^E7 pourraient produire des protéines fonctionnelles (Fig. 20). L'existence des protéines E6*I, E6*II et E6^E7 reste assez discutée, car elles n'ont été observées que lors d'expériences de surexpression et non dans des conditions physiologiques. Seule la protéine E6*I d'HPV18 a été observée dans des xénogreffes chez la souris de cellules HeLa (Schneider-Gädicke *et al*., 1988). Pourtant la surexpression de ces protéines a tout de même montré qu'elles possédaient certaines capacités affectant l'expression de plusieurs protéines cellulaires. Il a été ainsi montré que l'expression ectopique d'E6 et/ou E6*I d'HPV16 dans des cellules VADS (Voies aéro-digestives supérieures) HPV-augmente la radiosensibilité de ces cellules alors qu'E6*II n'a montré aucun effet (Pang *et al*., 2011). E6^E7 a la capacité de se lier à HSP90 (Heat Shock 90kDa Protein 1) et GRP78 (Glucose-regulated protein, 78kDa), qui induisent une augmentation de la demi-vie des protéines E6, E7 et E6*I.

Il est à noter que la majorité des expériences se sont focalisées sur la protéine E6*I, l'ARNm correspondant étant la forme majoritaire de transcrits retrouvés chez les HPV-HR, mais aussi la plus conservée. La protéine E6*I, composée des 43 premiers acides aminés d'E6, contient un motif C-x-x-C (Fig. 21A). La surexpression de la protéine E6*I a montré qu'elle était localisée, de façon diffuse, dans le noyau et le cytoplasme (Vaeteewoottacharn *et al*., 2005; Wanichwatanadecha *et al*., 2012). D'après ces expériences, la protéine E6*I présenterait

particulièrement une activité anti-oncogénique (Fig. 21B). En effet, la surexpression d'E6*I inhibe la prolifération cellulaire et l'apoptose des cellules HPV16+ mais pas des cellules HPV- (Pim et Banks, 1999; Pim *et al*., 1997). Pim et collaborateurs ont aussi montré qu'E6*I d'HPV18 empêche la dégradation de p53 par E6 (Pim *et al*., 1997). Cet effet apparait dépendant de la liaison d'E6*I avec la protéine E6 d'HPV18 (Pim et Banks, 1999). L'expression d'E6*I est de surcroit augmentée durant les phases G2/M du cycle cellulaire et est inversement corrélée à l'expression de la protéine p53 (Guccione *et al*., 2004). Ces résultats suggèrent qu'E6*I intervient dans la régulation de l'activité d'E6.

Par ailleurs, les effets anti-oncogéniques d'E6*I d'HPV16 sont à la fois dépendants et indépendants de l'expression de la protéine E6. Ainsi, l'expression d'E6*I dans des cellules SiHa induit l'augmentation de p53 et de la procaspase 8 (Filippova *et al*., 2014; Tungteakkhun *et al*., 2010). La stabilité de la procaspase 8 étant aussi augmentée par l'expression d'E6*I, indépendamment de l'expression d'E6 (Filippova *et al*., 2007). Par conséquent, l'expression ectopique d'E6*I dans des cellules HPV- et/ou HPV+, favorise la mort cellulaire et inhibe la croissance tumorale dans des xénogreffes chez la souris nude (Filippova *et al*., 2014). En outre, Filippova et collaborateurs ont montré qu'E6*I contrecarrait, de façon dose-dépendante, l'effet protecteur de la protéine E6 vis-à-vis de la mort cellulaire induite par Fas et le TNFα (Filippova *et al*., 2009).

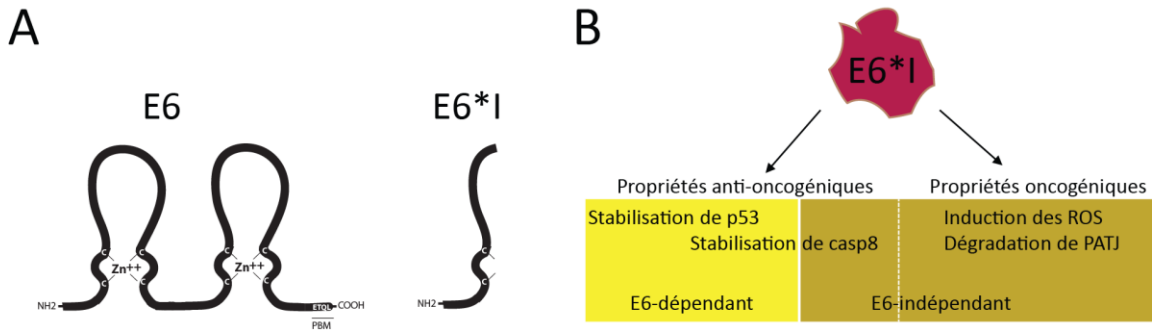


Figure 21. Structure et rôle de la protéine E6*I.

D'autre travaux ont aussi montré qu'E6*I avait des propriétés pro-oncogéniques dont la plupart sont indépendantes de l'expression de la protéine E6. En effet, l'expression d'E6*I d'HPV18 induit la dégradation de protéines cibles d'E6 (DLG1) *et al*tère l'expression d'AKT (Pim *et al*., 2009). E6*I d'HPV18 tout comme E6, diminue la stabilité de protéines impliquées dans le maintien de la polarité épithéliale telles que PATJ, DLG1 et MAGI-1, ainsi que la protéine kinase AKT (Pim *et al*., 2009; Storrs et Silverstein, 2007). Toutefois, alors qu'E6*I se lie avec PATJ suggérant un effet direct sur cette protéine, aucune interaction

avec DLG1, MAGI-1 et AKT n'a été observée, suggérant un effet indirect d'E6*I. Jha et collaborateurs ont aussi montré qu'E6*I d'HPV16 induit la dégradation de TIP60, une protéine à activité HAT (Histone Acetyl Transferase) impliquée dans la régulation de la transcription (Jha *et al*., 2010). D'ailleurs, la transfection d'E6*I, mais pas d'E6 ni d'E7, induit la transcription de nombreux gènes impliqués dans le métabolisme cellulaire tels AKR1C1 (Aldo-keto reductase family 1 member C1), AKR1C2, CD36, FGF1 (Fibroblast growth factor 1), FGF9, IGFBP3 (Insulin like growth factor binding protein 3), PLCB4 (Phospholipase C beta 4), RGS5 (Regulator of G protein signaling 5), SIPR1 (Sphingosine-1-phosphate receptor 1) et THBS1 (Thrombospondin 1) (Wanichwatanadecha *et al*., 2012). En particulier, E6*I induit une chimiorésistance *via* l'inhibition de l'aldo-kéto réductase codée par AKR1C. E6*I dérégule aussi le métabolisme cellulaire, en induisant une diminution de l'expression des protéines antioxydantes SOD1 (Superoxide Dismutase 1), SOD2 et GPX1 (Glutathione Peroxidase 1), engendrant de fait une augmentation des niveaux d'espèce réactives de l'oxygène ROS (Reactive-Oxygen Species) (Williams *et al*., 2014). Les ROS pourraient alors engendrer des lésions de l'ADN et une instabilité génomique favorisant l'intégration d'ADN épisomal dans le génome cellulaire. Enfin, alors qu'E6*I contient la moitié d'un domaine en doigt de zinc (uniquement 2 cystéines), Heer et collaborateurs ont observé une liaison avec le zinc (Heer *et al*., 2011). Ils ont suggéré qu'E6*I pourrait se dimériser autour d'un atome de zinc. Cette éventuelle oligomérisation pouvant dépendre de la concentration cellulaire en E6*I, pourrait expliquer certains effets dose-dépendants d'E6*I observés par Filippova et collaborateurs (Filippova *et al*., 2009).

3. Le complexe de jonction des exons

L'épissage des pré-ARNm permet non seulement de former les ARNm matures et de générer une diversité au sein du transcriptome, mais aussi d'influencer le devenir de ces mêmes ARNm. Il est en effet bien connu que l'historique nucléaire des ARNm, notamment l'épissage, stimule l'expression des protéines et dans certains cas, inhibe l'expression des ARNm. A l'issu de l'épissage, un complexe protéique est déposé sur les ARNm à 20-24 nucléotides en amont de la jonction exon-exon générée (Le Hir *et al*., 2000a, 2000b). Ce complexe, appelé EJC, reste fixé tout au long de la vie des ARNm et n'est éliminé que lors de leur traduction (Le Hir *et al*., 2016).

3.1. Cycle de l'EJC

3.1.1. Composition de l'EJC

L'EJC est composé de plusieurs protéines dont 4 protéines cœur : eIF4A3 appelée aussi DDX48, Y14 appelée aussi RBM8A (RNA-binding protein Y14 ou RNA binding motif protein 8A), MAGOH (MAGO homolog, EJC core component) et MLN51 (Metastatic Lymph Node 51) appelée aussi CASC3. eIF4A3 fait partie de la famille des eIF4A (avec eIF4A1 et eIF4A2), des protéines à activité ARN hélicase ATP-dépendante (Adenosine TriPhosphate) à boite DEAD (Aspartate-Glutamate-Alanine-Aspartate). Ces protéines possèdent deux domaines structuraux RecA-like impliqués dans l'activité ATPase de ces hélicases. Toutefois, seul eIF4A3 est incorporée dans l'EJC alors qu'eIF4A1 et eIF4A2 interviennent dans l'initiation de la traduction (Ballut *et al*., 2005). MAGOH et Y14 forment un hétérodimère fonctionnel très conservé au cours de l'évolution (Gong *et al*., 2014). En outre, MAGOH se lie au domaine RRM (RNA Recognition Motif) d'Y14 et l'empêche d'interagir avec l'ARN (Fribourg *et al*., 2003). Chez les mammifères, ils existent deux gènes paralogues *MAGOH* et *MAGOHB*, lesquels expriment des protéines MAGOH très similaires, ayant les mêmes fonctions et compensant l'absence de l'un des gènes (Singh *et al*., 2013). Enfin, MLN51, une protéine surexprimée dans des cancers du sein, contient plusieurs domaines conservés notamment le motif SELOR (Speckles Localizer and RNA binding motif), lequel est nécessaire et suffisant à la formation de l'EJC (Ballut *et al*., 2005; Degot *et al*., 2002, 2004; Woodward *et al*., 2017). MLN51 est la protéine la moins conservée des protéines du cœur de l'EJC. Elle présente des caractéristiques distinctes des trois autres protéines de l'EJC. En effet, les quatre protéines du cœur de l'EJC sont des protéines navettes, cependant eIF4A3, Y14 et MAGOH présentent majoritairement une localisation nucléaire

alors que MLN51 est principalement cytoplasmique (Daguenet *et al*., 2012; Degot *et al*., 2002, 2004). De plus, eIF4A3, Y14 et MAGOH peuvent former un complexe fonctionnel indépendant de MLN51 (Voir ci-dessous). En outre, la déplétion de chaque gène de l'EJC est létale au niveau embryonnaire ; Par ailleurs, les altérations assez similaires induites par la déplétion d'eIF4A3, Y14 ou MAGOH sont très différentes de celles induites par la déplétion de MLN51 (Mao *et al*., 2017).

Mis à part les fonctions liées à l'EJC (Voir Chapitre 3.4), MLN51 possèdent des fonctions cytoplasmiques, indépendante de l'EJC. Elle est impliquée dans la formation des granules de stress et induit le désassemblage des P-bodies (Processing bodies ou corps GW) (Baguet *et al*., 2007; Cougot *et al*., 2014). eIF4A3 et Y14 possèdent aussi des fonctions indépendantes de l'EJC, intervenant respectivement dans la maturation des ARN ribosomaux chez la levure *S. Cerevisiae* et dans le *decapping* des ARNm (Alexandrov *et al*., 2011; Chuang *et al*., 2013). La libération des composants du cœur de l'EJC lors du désassemblage de celui-ci, pourrait donc permettre à chacun d'entre eux d'effectuer une activité supplémentaire, sans rapport direct avec l'EJC.

3.1.2. Assemblage de l'EJC

In vitro, lorsque ces quatre protéines (eIF4A3, Y14, MAGOH et MLN51) sont mises en présence d'ARNsb (ARN simple brin) et d'ATP, l'EJC s'assemble spontanément sur l'ARN et ce, indépendamment de la séquence nucléotidique ou de l'épissage (Fig. 22). L'EJC est fixé sur l'ARN *via* eIF4A3, dont la forme fermée fait office de pince s'agrippant sur le squelette de ribose-phosphate de 6 nucléotides de l'ARN (Andersen *et al*., 2006; Ballut *et al*., 2005; Bono *et al*., 2006). C'est la liaison avec l'ATP qui permet à eIF4A3 d'avoir sa forme fermée, où ses 2 domaines RecA sont associés, alors que l'hydrolyse de l'ATP induit leur séparation et le passage à une forme ouverte incapable de se lier à l'ARN. De fait, l'hétérodimère Y14-MAGOH se fixe sur eIF4A3 et stabilise le complexe en inhibant sa fonction ATPase (Ballut *et al*., 2005; Nielsen *et al*., 2009). Enfin, MLN51 entoure eIF4A3 et interagit avec un nucléotide supplémentaire, renforçant la stabilité de l'EJC (Fig. 22 ; (Andersen *et al*., 2006; Ballut *et al*., 2005; Bono *et al*., 2006)).

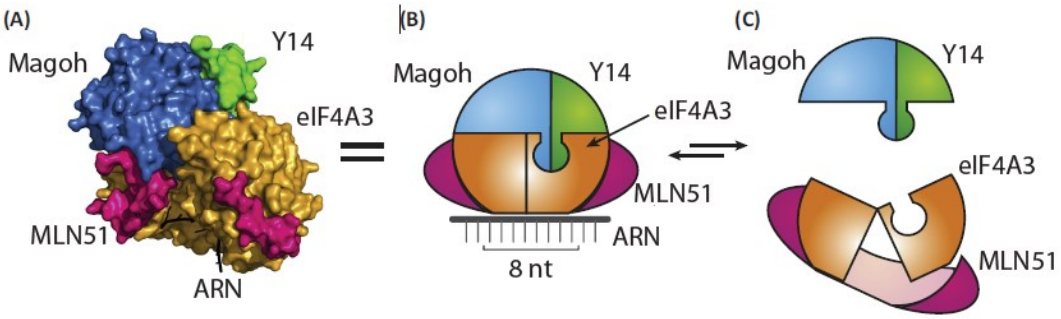


Figure 22. Structure de l'EJC. (Adaptée de Boehm et Gehring, 2016).

In vivo, l'EJC se forme quasiment *via* l'épissage du pré-ARNm. Le recrutement d'eIF4A3 sur le pré-ARNm est assuré par la protéine impliquée dans l'épissage, CWC22 (Alexandrov *et al*., 2012; Steckelberg *et al*., 2012, 2015). Celle-ci se fixe sur eIF4A3, la stabilise sous sa forme ouverte empêchant ainsi sa liaison avec l'ARN ou Y14-MAGOH (Barbosa *et al*., 2012; Buchwald *et al*., 2013). Pendant l'épissage, le spliceosome subit des changements et forment plusieurs complexes notamment B et C (Voir ci-dessus). Il a été montré qu'eIF4A3, Y14 et MAGOH co-précipitaient avec le complexe C. En outre, ces 3 protéines forment un complexe trimérique appelé pré-EJC, après la 1ère réaction de trans-estérification. MLN51 est recrutée après la ligature des exons pour former l'EJC (Gehring *et al*., 2009a). Par ailleurs, Daguenet et collaborateurs ont montré que l'assemblage de l'EJC a lieu préférentiellement dans les perispeckles, structures nucléaires situées à la périphérie des speckles. Les perispeckles sont riches en ARNm (polyA+), en facteurs de l'EJC et en ARN polymérase active alors que les speckles sont riches en facteurs d'épissage tels SC35 (Daguenet *et al*., 2012; Spector et Lamond, 2011).

3.1.3. Remodelage de l'EJC

L'EJC formé par les quatre protéines cœur eIF4A3, Y14, MAGOH et MLN51 fait office de plateforme à laquelle sont recrutées plusieurs protéines périphériques permettant de réguler l'expression des ARNm (Tableau III). Dès son assemblage dans le noyau jusqu'à son désassemblage dans le cytoplasme (Voir ci-dessous), l'EJC est soumis à un remodelage constant quant à sa composition en facteurs périphériques (Fig. 23).

Les premiers facteurs périphériques à être recrutés sur l'EJC sont vraisemblablement les protéines impliquées dans l'épissage et l'export des ARNm, telles RNPS1, SAP18 (Sin3A associated protein 18), Acinus (Apoptotic chromatin condensation inducer in the nucleus), Pinin, Aly/REF (Aly/REF export factor) et UAP56 (U2AF65-associated protein 56). D'ailleurs ces facteurs sont présents dans les complexes spliceosomaux et dans les

perispeckles (Daguenet *et al*., 2012; Kota *et al*., 2008; Teng et Wilson, 2013). De plus, il a été montré que RNPS1, Aly/REF et UAP56 pouvaient être recrutés sur l'ARNm de façon indépendante d'eIF4A3 et du cœur de l'EJC (Zhang et Krainer, 2007). En outre, RNPS1 et SAP18 forment les complexes ASAP (Apoptosis and splicing-associated protein) et PSAP avec Acinus et Pinin respectivement, qui sont impliquées dans la régulation de l'épissage. Les modalités de fixation des complexes ASAP et PSAP sur l'EJC sont encore peu connus. Par ailleurs, Aly/REF et UAP56, facteurs d'export faisant partie du complexe TREX (TRanscription and Export) et pouvant être recrutés sur les ARNm indépendamment de l'épissage, permettent le recrutement des protéines d'export NXT1 (Nuclear transport factor 2 like export factor 1) et NXF1 (Nuclear RNA export factor 1). Le complexe NXT1-NXF1 interagit avec les nucléoporines et promeut l'export des ARNm vers le cytoplasme. Enfin, les facteurs périphériques UPF3a (Up-Frameshift Suppressor 3 Homolog A) ou UPF3b et SKAR (S6K1 Aly/REF-like target) sont aussi recrutés sur l'EJC dans le noyau.

Tableau III. Les composants de l'EJC. (Adapté de Le Hir *et al*., 2016).

Protéine (Alias)	Localisation cellulaire	Fonctions cellulaires	Présence dans le spliceosome	Liaison stable aux ARNm épissés	Liaison direct aux facteurs de l'EJC
Cœur de l'EJC					
eIF4A3 (DDX48, *eIF4AIII*)	Protéine navette; principalement nucléaire	Biogénèse des ribosomes et expression des selenoprotéines	oui	Noyau et cytoplasme	Y14, Magoh, MLN51, UPF3b, SMG6
Y14(RBM8A, *Tsunagi*)	Protéine navette; principalement nucléaire	Localisation des ARNm et décapping	oui	Noyau et cytoplasme	Magoh,eIF4A3, UPF3b, SMG6
MAGOH (*Mago nashi*)	Protéine navette; principalement nucléaire	Localisation des ARNm et décapping	oui	Noyau et cytoplasme	Y14, eIF4A3, MLN51, UPF3b, SMG6
MAGOHB	Protéine navette; principalement nucléaire	Localisation des ARNm et décapping	oui	Noyau et cytoplasme	Y14, eIF4A3, MLN51, UPF3b, SMG6
MLN51 (CASC3, *Barentz*)	Protéine navette; principalement cytoplasmique	Traducation, assemblage des granules de stress, Localisation des ARN	Non	Noyau et cytoplasme	eIF4A3, Magoh, eIF3A, eIF3D
Protéines périphériques					
RNPS1	Protéine navette; principalement nucléaire	Epissage	oui	Noyau et cytoplasme	Acinus, Pinin, SAP18, UPF3b
Acinus	Nucléaire	Transcription, épissage et apoptose	oui	noyau	RNPS1, SAP18
Pinin	Nucléaire	Epissage	oui	noyau	RNPS1, SAP18
SAP18	Protéine navette; principalement nucléaire	Transcription et épissage	oui	Noyau et cytoplasme (?)	Acinus, Pinin, SAP18
ALYREF	Protéine navette; principalement nucléaire	Transcription et export des ARNm	oui	noyau	NXF1, UAP56, Y14
UAP56 (DDX39B)	Protéine navette; principalement nucléaire	Export et épissage des ARNm	oui	-	ALYREF
NXF1 (TAP)	Protéine navette; principalement nucléaire	Export des ARNm	Non	noyau	ALYREF, NXT1
NXT1 (p15)	Protéine navette; principalement nucléaire	Export des ARNm	Non	noyau	NXF1
UPF3a (RENT3a)	Protéine navette; principalement nucléaire	NMD	Non	Noyau et cytoplasme	UPF2
UPF3b (RENT3b)	Protéine navette; principalement nucléaire	NMD	Non	Noyau et cytoplasme	eIF4A3, Y14, Magoh, UPF2, RNPS1
UPF2 (RENT2)	Cytoplasmique, périnucléaire	NMD	Non	Cytoplasme	UPF3b, UPF3a, UPF1
UPF1 (RENT1)	Cytoplasmique	NMD	Non	Cytoplasme	UPF2
SMG6	Protéine navette; principalement cytoplasmique	NMD	Non	?	eIF4A3, Y14, Magoh
SKAR	Protéine navette; principalement cytoplasmique	Traduction et export des ARNm	?	Noyau et cytoplasme	?

La plupart des protéines périphériques recrutées sur l'EJC sont nucléaires, se détachant de l'EJC dans le noyau et surtout au niveau du cytoplasme. Ainsi, après l'export, les facteurs Aly/REF, UAP56, SAP18, Acinus et Pinin ne sont plus détectés sur les ARNm, tandis que les facteurs impliqués dans les fonctions cytoplasmiques de l'EJC RNPS1, UPF3b/a et SKAR restent fixés (Fig. 23). Par ailleurs, des facteurs périphériques de l'EJC, tels UPF1 et UPF2, sont recrutés sur les ARNm au niveau du cytoplasme, et permettent la dégradation des ARNm aberrants par le NMD (voir ci-dessous). Enfin, l'endonucléase SMG6 (human suppressor with *morpho*- genetic effect on *genitalia 6*) se lie aussi au cœur de l'EJC. En outre, SMG6 et UPF3b/a utilisent un même motif EBM (EJC Binding Motif) pour se lier à l'EJC, et pourraient donc rentrer en compétition quant à leur fixation sur l'EJC (Kashima *et al*., 2010). UPF3b ayant une meilleure affinité pour l'EJC, SMG6 serait recruté après la dissociation d'UPF3b dans le cytoplasme.

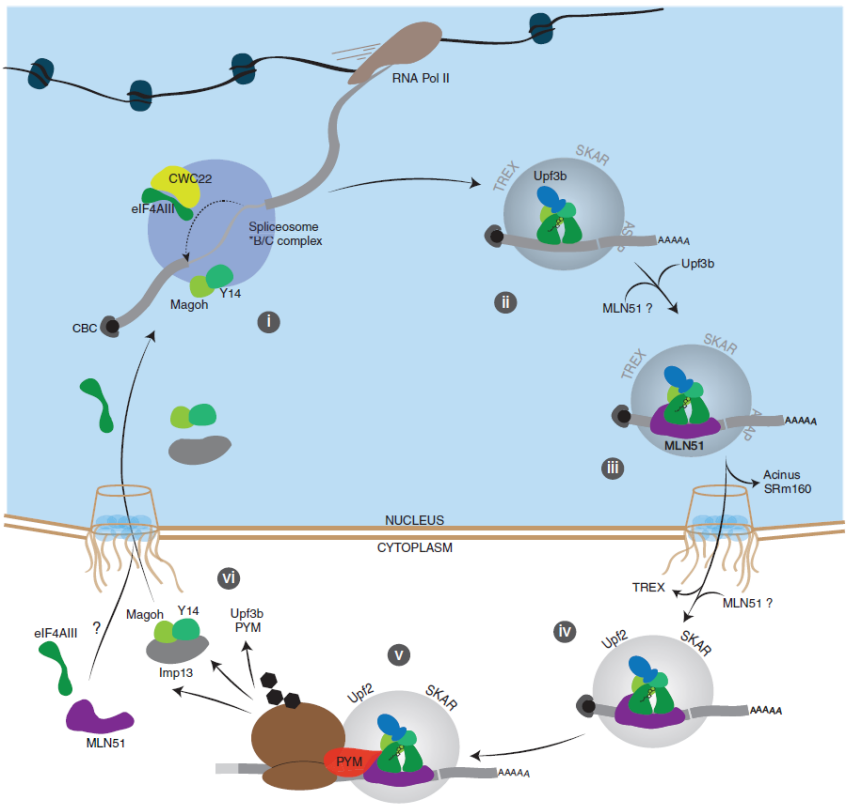


Figure 23. Le cycle de l'EJC. (Adaptée de Woodward *et al*., 2017).

3.1.4. Désassemblage et recyclage de l'EJC

Dès la découverte de l'EJC, il a été montré que celui-ci était éliminé des ARNm pendant le 1er cycle de traduction (Dostie et Dreyfuss, 2002; Lejeune *et al*., 2002). En effet, ce 1er cycle concerne souvent les ARNm liés à CBC, lesquels présentent des EJC, alors que les ARNm

liés à eIF4E (remplace CBC pour se lier à la coiffe) ne possèdent plus d'EJC (Lejeune *et al*., 2002). En outre, le déplacement du ribosome le long de l'ARNm permet d'enlever les EJC présents dans l'ORF et la région 5'UTR (Boehm et Gehring, 2016). De plus, l'EJC est aussi désassemblé par PYM1 (PYM homolog 1, exon junction complex associated factor) qui cible le dimère Y14-MAGOH (Diem *et al*., 2007; Gehring *et al*., 2009b). *Via* son domaine N-ter, PYM1 se lie à l'interface créée par l'association de Y14 et MAGOH (Bono *et al*., 2004). De fait, PYM1 ne se lie pas à Y14 ou MAGOH seule, mais exclusivement au dimère Y14-MAGOH (Diem *et al*., 2007). PYM1 se lie aussi avec les ribosomes et le PIC de traduction *via* son domaine C-ter, couplant ainsi sa fonction avec la traduction (Diem *et al*., 2007; Gehring *et al*., 2009b). L'interaction PYM-Y14-MAGOH empêche potentiellement la reformation de l'EJC, et favorise le recyclage des éléments de l'EJC. Ce recyclage est d'autant plus important que le nombre de molécules d'eIF4A3 (12 000/cellule), Y14-MAGOH (40 000/cellule) est assez négligeable comparé au nombre de jonction exon-exon cellulaire (400 000/cellule) (Gehring *et al*., 2009b). En effet, l'inhibition de PYM1 réduit l'expression des protéines à partir d'un plasmide rapporteur contenant un intron (Diem *et al*., 2007).

Par ailleurs, le recyclage des éléments de l'EJC Y14-MAGOH vers le noyau est réalisé *via* l'importine 13 (Boehm et Gehring, 2016; Bono *et al*., 2010; Le Hir *et al*., 2016). Le désassemblage de l'EJC libère aussi eIF4A3 et MLN51, lesquelles peuvent former un hétérodimère (Chamieh *et al*., 2008), dont le mécanisme d'import vers le noyau n'a pas encore été décrit.

3.2. Dépôt de l'EJC au sein du transcriptome

La différence entre le nombre de composants de l'EJC et de jonctions exon-exon cellulaires (Voir ci-dessus), suggère que l'EJC ne puisse pas être présent sur toutes les jonctions exon-exon. Ainsi, plusieurs études ont montré que l'occupation des jonctions exon-exon par l'EJC ne dépassait pas les 80% (Hauer *et al*., 2016; Saulière *et al*., 2010, 2012; Singh *et al*., 2012). Cette fixation différentielle de l'EJC peut avoir un impact sur le rôle de celui-ci dans la régulation de l'expression des ARNm. Le dépôt de l'EJC sur une jonction exon-exon peut dépendre du contexte nucléotidique entourant son site de fixation. Ainsi, la présence de structure secondaire au lieu de fixation de l'EJC (20-30 nucléotides en amont de la jonction exon-exon) fait office de contrainte au dépôt de l'EJC (Mishler *et al*., 2008). De plus, le dépôt a lieu préférentiellement aux alentours de séquences GAAGA, reconnues par les protéines

SRSF1 et SRSF7 (Saulière *et al*., 2012). Par ailleurs, les exons utilisant l'épissage alternatif montrent une haute proportion à la fixation de l'EJC, suggérant que les protéines SR, intervenant lors de l'épissage alternatif, favoriserait le dépôt de l'EJC.

D'autre part, il a été montré qu'à part son dépôt à la position préférentielle (20-24 nucléotides en amont de la jonction exon-exon), l'EJC se fixe aussi à d'autres positions non canoniques, appelé ncEJC (non canonical EJC) (Hauer *et al*., 2016; Saulière *et al*., 2012; Singh *et al*., 2012). Le dépôt de l'EJC à une position non canonique peut être induit par la présence de structure secondaire de l'ARN à la position canonique (Mishler *et al*., 2008). Le cEJC et le ncEJC étant formés sur des ARNm épissés, la proportion de ce dernier a été assez diverse entre les études (Hauer *et al*., 2016; Saulière *et al*., 2012; Singh *et al*., 2012). En effet, lorsque la position de l'EJC a été révélée par eIF4A3, ou d'autres facteurs périphériques de l'EJC tels UPF3b et RNPS1, la proportion de ncEJC est plus élevée comparativement à l'utilisation de MLN51 (Hauer *et al*., 2016).

En somme, ce positionnement différentiel de l'EJC pourrait induire des variations dans la régulation des gènes par l'EJC. En outre, la présence potentielle d'EJC devrait être prise avec précaution quant à la prédiction de l'impact de l'épissage sur un ARNm donné.

3.3. Les fonctions de l'EJC

3.3.1. Régulation de l'épissage

Identifié tel une empreinte de l'épissage, l'EJC a été longtemps décrit comme impliqué dans la régulation des événements post-épissage affectant le devenir des ARNm. Toutefois, plusieurs études ont montré l'importance de l'EJC dans l'épissage des pré-ARNm chez plusieurs organismes (Le Hir *et al*., 2016). Chez la drosophile, il a été montré que l'inhibition de facteurs de l'EJC retrouvés dans le spliceosome tels eIF4AIII, Y14, Mago et RNPS1 altérait l'épissage des transcrits de la voie de signalisation MAPK (Mitogen-activated protein kinase) (Ashton-Beaucage *et al*., 2010; Roignant et Treisman, 2010). L'inhibition de MLN51 n'a par contre induit aucune altération de l'épissage des transcrits de la voie MAPK. La déplétion des composants du pré-EJC (eIF4AIII, Y14, Mago) et/ou RNPS1 altérait particulièrement l'épissage des transcrits contenant de longs introns (>1000 nucléotides) (Ashton-Beaucage *et al*., 2010; Roignant et Treisman, 2010). L'inhibition des facteurs du pré-EJC altèrent aussi l'épissage des transcrits piwi, induisant la rétention de l'intron 4 (Hayashi *et al*., 2014; Malone *et al*., 2014).

Par ailleurs, Michelle et collaborateurs ont montré que l'inhibition des facteurs du pré-EJC (eIF4A3 et Y14) et du complexe ASAP (RNPS1, SAP18 et Acinus) mais pas de MLN51, dans des lignées cellulaires humaines, engendrait des altérations de l'épissage des gènes de l'apoptose tels Bcl-X (Michelle *et al*., 2012). Ces altérations favorisant l'expression des isoformes pro-apoptotiques, induisent par conséquent l'apoptose des cellules déplétées en ces facteurs de l'EJC. De façon similaire, Fukumura et collaborateurs ont montré l'implication de l'EJC dans l'épissage de plusieurs gènes impliqués dans la régulation du cycle cellulaire (Fukumura *et al*., 2016). Ils ont en outre montré que l'inhibition de l'EJC favorisait la rétention d'introns de plusieurs gènes tels MDM2 (Mouse double minute 2 homolog). Les introns retenus sont relativement de petite taille comme l'intron 4 de piwi chez la drosophile (Fukumura *et al*., 2016; Hayashi *et al*., 2014; Malone *et al*., 2014). Une autre étude menée par Wang et collaborateurs a aussi montré l'impact de l'EJC sur l'épissage de nombreux transcrits. En séquençant le transcriptome des cellules HeLa dont l'expression d'eIF4A3, Y14 ou MLN51 a été inhibée par RNA-seq (RNA-sequencing), ils ont identifié une centaine d'ARNm dont l'épissage est affecté par les 4 protéines cœur de l'EJC (Wang *et al*., 2014). Selon leur étude, la longueur des introns adjacents n'explique pas la dépendance de l'épissage de ces transcrits à l'EJC, contrairement à ce qui a été observé chez la drosophile (Ashton-Beaucage *et al*., 2010; Roignant et Treisman, 2010; Wang *et al*., 2014). Le mode de régulation de l'épissage par l'EJC pourrait être différent entre la drosophile et les mammifères, notamment l'Homme.

3.3.2. Export nucléo-cytoplasmique des ARNm

Les ARNm sont exportés du noyau vers le cytoplasme afin d'être traduits en protéines. Le mécanisme principal d'export des mRNP vers le cytoplasme est médié par le complexe NXF1/NXT1 (appelé aussi TAP/p15), lequel interagit avec les composants des pores nucléaires (Sloan *et al*., 2016). NXF1/NXT1 est recruté sur les mRNP *via* sa liaison à des protéines adaptatrices, telles les protéines du complexe TREX et les protéines SR. L'EJC pourrait stimuler le recrutement de TREX et NXF1/NXT1 sur les ARNm épissés et donc leur export. D'ailleurs, l'épissage promeut l'export nucléaire d'ARNm rapporteurs et augmente la cinétique d'export vers le cytoplasme (Le Hir *et al*., 2001; Valencia *et al*., 2008; Zhou *et al*., 2000). Cet effet est notamment dû au recrutement des protéines du complexe TREX, Aly/REF et UAP56 et de fait NXF1/NXT1 sur l'EJC (Le Hir *et al*., 2001; Luo *et al*., 2001; Zhou *et al*., 2000). Ainsi, en modulant l'efficacité d'export des ARNm, l'EJC peut donc réguler les niveaux d'expression des protéines. Toutefois, des expériences menées chez la drosophile et

le ver *C.elegans*, ont montré que l'inhibition de l'EJC n'affecte l'export des transcrits cellulaire chez la drosophile que de façon très modéré (Gatfield et Izaurralde, 2002; Longman *et al*., 2003). Ces transcrits cellulaires, beaucoup plus grands que les ARNm rapporteurs, disposeraient potentiellement de plusieurs sites de liaison à des protéines adaptatrices, tels SRSF1, SRSF6 et SRSF7 permettant de recruter les protéines de l'export TREX et/ou NXF1/NXT1 (Le Hir *et al*., 2016). De plus, TREX et donc NXF1/NXT1 peuvent être recrutées sur les ARNm *via* le CBC sur la coiffe, comme c'est le cas dans les ARNm sans intron (Cheng *et al*., 2006; Nojima *et al*., 2007). Enfin, l'export d'une certaine catégorie de transcrits pourrait avoir une dépendance plus importante vis-à-vis de l'EJC (Le Hir *et al*., 2016).

3.3.3. Stimulation de la traduction

Il a été montré que la présence d'introns améliore l'expression des protéines, non seulement en affectant leur export du noyau, mais aussi en augmentant leur traduction (Nott *et al*., 2004). En particulier, la présence d'un intron dans la région 5'UTR, potentiellement d'un EJC, stimule la traduction à partir de plasmides rapporteurs. Cet effet pourrait être lié à la fonction hélicase d'eIF4A3. Il a été ainsi montré par Choe et collaborateurs qu'eIF4A3 optimise la traduction en déroulant les structures secondaires de l'ARNm (Choe *et al*., 2014). Cette équipe a montré qu'en plus de se lier aux ARNm épissés, eIF4A3 pouvait se lier à des ARNm ne contenant pas d'introns.

Par ailleurs, les protéines du cœur de l'EJC interagissent avec de nombreuses protéines impliquées dans la traduction telles SKAR et eIF3 (Chazal *et al*., 2013; Ma *et al*., 2008). D'une part, Ma et collaborateurs ont montré que SKAR serait un facteur périphérique de l'EJC (Ma et collaborateurs), se liant avec les ARNm contenant CBC (Ma *et al*., 2008). SKAR permettrait de recruter la protéine S6K1 (Ribosomal Protein S6 Kinase B1) sur les ARNm épissés, entrainant la phosphorylation des facteurs de traduction et des protéines ribosomales et par conséquent, faciliterait l'initiation de la traduction sur ces ARN.

D'autre part, Chazal et collaborateurs ont montré que MLN51 se liait avec le facteur du PIC, eIF3 et stimule ainsi la traduction des ARNm épissés et non épissés dans une moindre mesure (Chazal *et al*., 2013). En outre, l'effet de MLN51 sur la traduction peut être à la fois dépendant et indépendant de l'EJC. L'EJC favoriserait uniquement la traduction des ARNm liés à CBC, alors qu'un rôle potentiel de MLN51 dans la traduction dépendante d'eIF4E est à ne pas négliger. De façon intéressante, MLN51 et eIF4A3 peuvent interagir en dehors de

l'EJC, et pourraient jouer un rôle potentiel après le désassemblage de l'EJC (Chamieh *et al*., 2008).

3.3.4. Rôle dans la surveillance des ARNm par le NMD

L'EJC est impliqué dans la reconnaissance et la dégradation des ARNm contenant un PTC (Premature Termination Codon) par le NMD. En outre, l'EJC a été identifié principalement *via* des études sur le NMD.

3.4. Le Nonsense-mediated mRNA decay

Le NMD est un mécanisme de contrôle de qualité des ARNm ciblant les ARNm aberrants, notamment ceux ayant un codon PTC. Il permet ainsi d'éviter la production de protéines tronquées potentiellement délétères pour les cellules. Chez les mammifères, la reconnaissance et la dégradation des ARNm contenant un PTC met en jeu plusieurs protéines dont celles de l'EJC. Alors que le NMD est majoritairement dépendant de l'EJC, chez les mamifères, plusieurs observations suggèrent que le NMD puisse avoir lieu indépendamment de celui-ci, à commencer par les modèles de levures et drosophiles où l'EJC n'intervient pas dans le NMD.

3.4.1. Le NMD dépendant de l'EJC

La présence d'un intron en aval d'un codon stop a longtemps été décrite comme un signal induisant la dégradation de l'ARNm par le NMD. Des études ont montré que la présence d'intron épissé (et donc d'EJC) en aval d'un codon Stop (à 50-55 nucléotides) est le signal principal induisant la reconnaissance des ARNm par le NMD.

Les EJC sont enlevés des ARNm au cours du 1er tour de traduction, dépendant de GTP. L'arrêt du ribosome au niveau d'un codon PTC, à plus de 50 nucléotides en amont de la dernière jonction exon-exon, permet la persistance d'un EJC sur l'ARNm. En outre, le positionnement du ribosome sur le PTC permet de recruter les facteurs de terminaison eRF1 (eucaryotic Release Factor 1) et eRF3, mais aussi les protéines UPF1, SMG1, SMG8 et SMG9 formant le complexe SURF (SMG1-UPF1-Release Factor) (Fig. 24 ; (Kashima *et al*., 2006; Yamashita *et al*., 2009). La protéine UPF1 se lie ensuite avec des protéines de l'EJC, UPF2 et UPF3, conduisant à la formation du complexe DECID (Decay-inducing complex) (Fig. 24 ; (Kashima *et al*., 2006)). La formation du complexe DECID permet la phosphorylation d'UPF1 par SMG1, et conduit à l'inhibition de la traduction et à la dégradation des ARNm (Isken *et al*., 2008). La phosphorylation d'UPF1 induit la dissociation

d'eRF1 et eRF3 du complexe DECID, mais aussi le recrutement des protéines SMG5, SMG6 et SMG7 qui vont médier la dégradation des ARNm (Fig. 24 ; (Isken *et al*., 2008; Okada-Katsuhata *et al*., 2012)). SMG5 et SMG7, permettent la déphosphorylation d'UPF1 mais aussi le recrutement de protéines impliquées dans le *decapping* et la deadénylation des ARNm, ce qui induit leur dégradation par des exonucléases (Cho *et al*., 2013; Durand *et al*., 2007; Loh *et al*., 2013; Ohnishi *et al*., 2003). Quant à l'endonucléase SMG6, elle induit un clivage endonucléolytique des ARNm, libérant des extrémités 5' et 3' accessibles à la dégradation par les exonucléases (Eberle *et al*., 2009; Huntzinger *et al*., 2008).

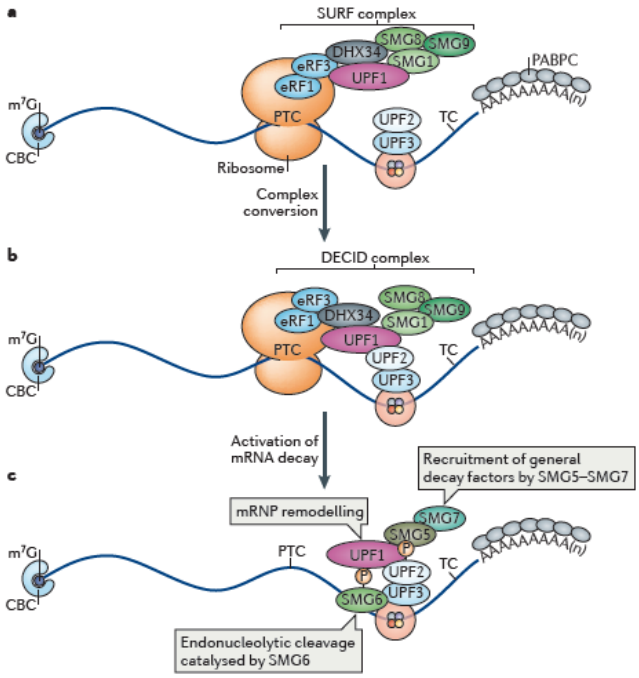


Figure 24. La fonction de l'EJC dans le NMD. (Adaptée de Le Hir *et al*., 2016).

Le NMD est principalement dépendant de l'EJC chez les mammifères. Toutefois, certains ARNm ciblés par le NMD ne contiennent pas d'intron dans la région 3'UTR (Amrani *et al*., 2004; Bühler *et al*., 2006; Chan *et al*., 1998; Matsuda *et al*., 2007; Zünd *et al*., 2013). Inversement, d'autres ARNm contenant des introns en 3'UTR sont résistants au NMD. Il est à noter que le profil de dépôt de l'EJC pourrait expliquer certaines de ces différences. D'une part, la présence de ncEJC dans certaines régions 3'UTR (donc sans intron), pourrait induire le NMD, d'une façon dépendante de l'EJC (Saulière *et al*., 2012; Singh *et al*., 2012). D'autre part, les EJC n'étant pas présents sur toutes les jonctions exon-exon, l'induction du NMD par la présence d'intron en aval d'un codon stop est dépendante de la présence de l'EJC et non de l'épissage en lui-même (Saulière *et al*., 2012; Singh *et al*., 2012). Il est aussi possible que ces ARNm aient développé des propriétés leur permettant d'échapper au NMD (Voir ci-dessous).

Par ailleurs, la dégradation d'ARNm en l'absence d'intron en 3'UTR, pourrait être effectuée par un mécanisme NMD indépendant de l'EJC, comme c'est le cas chez les eucaryotes inférieurs (Ex. la levure *S. Cerevisiae*).

3.4.2. Signaux de dégradation par le NMD

Un PTC peut être induit par des mutations de l'ADN, d'épissage alternatif ou d'erreurs de transcription ; il peut aussi correspondre à des caractéristiques déjà existantes au sein des transcrits. En effet, plusieurs situations physiologiques peuvent correspondre à un PTC (Fig. 25). (i) La présence d'une uORF (upstream ORF) en amont de l'ORF principale dont la terminaison de la traduction peut être assimilée à un PTC. En effet, le codon stop est situé en amont de l'ORF principale (de son codon physiologique), et donc d'un ou plusieurs EJC. (ii) L'épissage alternatif au sein de la région 3'UTR, permettant de déposer un EJC en aval du codon stop physiologique, lui conférant la caractéristique de PTC. (iii) Une longue région 3'UTR peut activer le NMD d'une façon indépendante ou dépendante (ncEJC) de l'EJC. (iv) la présence d'une séquence de décalage du cadre de lecture du ribosome. Ces séquences permettent de décaler le ribosome aux alentours d'un codon stop afin de pouvoir continuer la traduction. L'efficacité de ce décalage étant faible, la traduction se termine pour certains ARN au codon stop, ayant des caractéristiques de PTC.

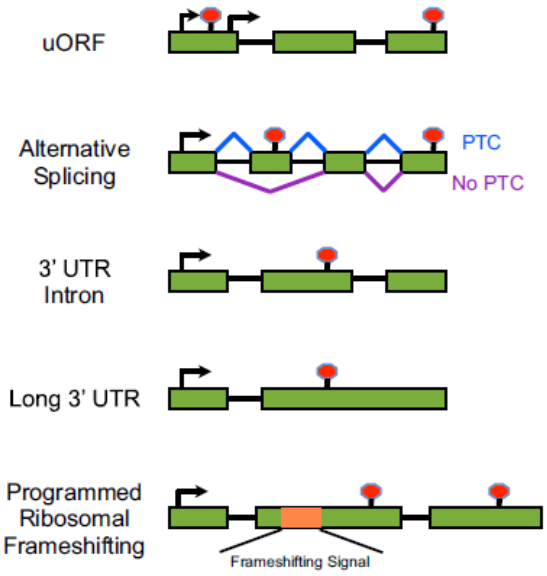


Figure 25. Les signaux de dégradation par le NMD. (D'après Nickless *et al.*, 2017).

Ces signaux qui induisent la dégradation des ARNm par le NMD, sont retrouvés dans plusieurs ARNm physiologiques, ce qui suggère que le NMD puisse cibler des ARN naturels au sein des cellules. De plus, la terminaison de traduction qui peut avoir lieu au niveau des

codons, la sélénocystéine, (qui correspond au codon stop UGA), induit aussi une dégradation par le NMD. L'expression des sélénoprotéines est régulée par la balance entre l'efficacité du NMD et la capacité de ses transcrits à échapper à ce mécanisme, en outrepassant le codon Stop. De fait, l'expression des sélénoprotéines est régulée par le niveau de Selenium cellulaire qui permet d'incorporer la sélénocystéine.

3.4.3. Cas particuliers : le NMD indépendant de l'EJC

L'activation du NMD indépendamment de l'EJC est le modèle utilisé chez les non-mammifères (Levure, drosophile). Certaines données suggèrent sa conservation chez les mammifères (Amrani *et al*., 2004; Bühler *et al*., 2006; Chan *et al*., 1998; Matsuda *et al*., 2007; Zünd *et al*., 2013). Le mécanisme d'activation du NMD de façon indépendante de l'EJC n'a pas encore été bien caractérisé. En outre, les caractéristiques permettant la reconnaissance des PTC seraient : la présence d'une séquence DSE (Downstream Sequence Element) qui lie des protéines et/ou d'une longue région 3'UTR (Lejeune, 2017; Rebbapragada et Lykke-Andersen, 2009).

La DSE permet de lier des protéines, comme Hrp1p/Nab4p (Nuclear polyadenylated RNA-binding protein 4), laquelle active le NMD, agissant de fait d'une façon similaire à un EJC (Rebbapragada et Lykke-Andersen, 2009). Chez la levure, la majorité des ARNm présente une région 3'UTR de 100 nucléotides. Un PTC qui induit l'allongement de la région 3'UTR provoque la dégradation des ARNm par le NMD. Des ARNm qui ont une région 3'UTR plus longue de façon naturelle sont aussi ciblés par le NMD. La reconnaissance d'un PTC par le NMD serait liée aux conditions de terminaison de la traduction, notamment de la distance entre PABPC et le codon stop. PABPC, un stimulateur de la terminaison de la traduction, entre en compétition avec UPF1 pour la liaison avec eRF3 et agit comme un facteur inhibiteur du NMD. Dans le cas d'un PTC, la distance entre le ribosome et PABPC est plus grande, empêche de fait celle-ci de stimuler la terminaison de la traduction et permet la fixation d'UPF1 avec eRF3. En outre, le recrutement de PABPC à une distance suffisamment proche d'un PTC, inhibe la dégradation d'ARNm jusque-là ciblé par le NMD.

De plus, chez les eucaryotes supérieurs, il a été montré que certains ARNm possédant une région 3'UTR longue (> 1 kb), sont dégradés par le NMD. Toutefois, une région 3'UTR assez longue ne signifie pas nécessairement une dégradation par le NMD, étant donné le nombre de gènes ayant une longue 3'UTR insensibles au NMD. De plus, il a été montré que plusieurs transcrits, sauvages (ARNm Thrombopoietin) ou mutés (ARNm Fibrinogen alpha chain,

Fibrinogen gamma chain *et al*pha-1,3- mannosyltransferase), possédant potentiellement des EJC en 3'UTR n'étaient pas dégradés par le NMD (Asselta *et al*., 2001; Denecke *et al*., 2004; Neerman-Arbez *et al*., 2004; Stockklausner *et al*., 2006). Certains de ces transcrits pourraient posséder des caractéristiques leur permettant d'échapper au NMD. Des structures secondaires de l'ARN peuvent potentiellement rapprocher PABPC du codon stop ; des séquences *cis*- permettant de recruter des facteurs inhibiteurs du NMD. En somme, la présence d'un PTC est reconnue *via* sa distance vis-à-vis de la queue polyA, tandis que la présence d'EJC (DES chez la levure) en aval renforce l'activité du NMD. D'ailleurs, la dégradation des ARNm par le NMD est plus efficace lorsqu'un EJC est présent dans la région 3' UTR.

Du reste, la majorité des études sur le NMD mettent en œuvre l'inhibition du NMD *via* UPF1 afin d'identifier des signaux et des cibles du NMD (comme ci-dessous). Il est important de ne pas omettre l'implication d'UPF1 dans plusieurs systèmes de dégradation des ARNm : le SMD (Staufen-mediated mRNA Decay) ; la dégradation régulée des ARNm des histones dépendant de la réplication ((He et Jacobson, 2015; Imamachi *et al*., 2012). Dans le SMD, un mécanisme de dégradation dépendant de la traduction, UPF1 est recruté sur l'ARN par la protéine Staufen 1. Le SMD cible les ARN possédant une séquence de liaison de Staufen 1.

3.4.4. Rôle du NMD dans le contrôle de l'expression des gènes

Bien que le NMD ait d'abord été identifié comme un mécanisme de contrôle de la qualité des ARNm, plusieurs études ont montré qu'il est aussi impliqué dans la régulation de nombreux transcrits naturels (Nickless *et al*., 2017) (Fig. 26). En effet, des études transcriptomiques ont montré que l'inhibition des facteurs du NMD, UPF1, UPF2, SMG6 ou SMG7, induit une augmentation de l'expression de nombreux ARNm (Chapin *et al*., 2014; Colombo *et al*., 2017; McGlincy *et al*., 2010; Mendell *et al*., 2004; Wittmann *et al*., 2006). De fait, en régulant l'expression des transcrits cellulaires, le NMD est impliqué dans la régulation de nombreux processus cellulaires physiologiques ou pathologiques (Fig. 26).

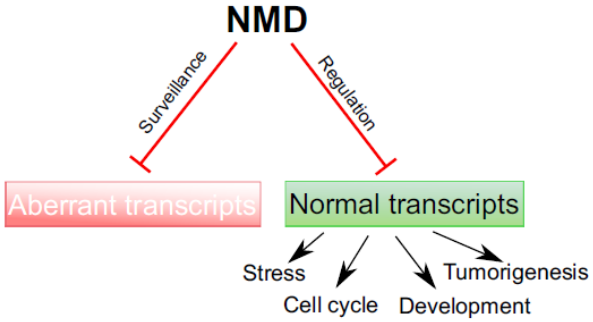


Figure 26. Rôle du NMD dans les cellules. (D'après Goetz et Wilkinson, 2017).

Ces gènes cibles « physiologiques » du NMD sont donc régulés selon l'activité du NMD, qui diffère selon le type cellulaire et le contexte physiologique (Bateman *et al*., 2003; Goetz et Wilkinson, 2017; Linde *et al*., 2007; Pereverzev *et al*., 2015). Par exemple, une diminution de l'activité du NMD, accompagnée par l'expression accrue de ses cibles, a été observée lors du développement du cerveau chez la souris (Bruno *et al*., 2011). Dans ce cas, la diminution de l'activité du NMD est expliquée par l'inhibition d'UPF1 et MLN51 par le miR-128 (Bruno *et al*., 2011). De plus, l'activité du NMD est aussi modifiée par la présence de stress cellulaire, permettant l'expression des protéines de réponse aux stress, régulées par le NMD (Gardner, 2008; Goetz et Wilkinson, 2017; Li *et al*., 2017). L'activité du NMD est entre autre régulée par le niveau de Calcium (Ca^{2+}) intracellulaire et de ROS (Nickless *et al*., 2014). En outre, l'oncogène MYC induit la production des ROS, qui conduit à la phosphorylation d'eIF2α et à l'inhibition du NMD (Wang *et al*., 2011a). La régulation de l'activité du NMD est donc primordiale pour contrôler l'expression de ses gènes cibles et vraissemblablement pour le maintien des fonctions physiologiques assurées par leurs produits. Une altération de l'activité du NMD et de l'expression des gènes cibles, pourrait participer à l'apparition de processus pathologiques. D'ailleurs, le NMD est dérégulé dans de nombreux cancers (Voir ci-dessous).

De façon intéressante, le NMD permet la régulation de l'expression des protéines impliquées dans le NMD, mais aussi des facteurs régulateurs d'épissage *via* un mécanisme de rétrocontrôle négatif. En effet, l'expression des transcrits codant les facteurs du NMD (UPF1, UPF2, SMG1, SMG5, SMG6 et SMG7) est liée à l'activité du NMD dans les cellules et *vice-versa* (Huang *et al*., 2011; Yepiskoposyan *et al*., 2011). Ainsi, l'activité du NMD est corrélée à l'expression de ses facteurs tel RNPS1 ou SMG1 (Ahn *et al*., 2013; Viegas *et al*., 2007). Selon leur niveau d'expression au niveau protéique, les facteurs d'épissage comme SRSF2 et hnRNP A2 induisent un épissage alternatif sur leur propres transcrits, générant des isoformes ciblés par le NMD (McGlincy *et al*., 2010; Sureau *et al*., 2001). En somme, le NMD contribue à réguler l'expression de nombreuses cibles, directes ou indirectes, impliquant ainsi ce mécanisme dans plusieurs processus cellulaires. Les transcrits ciblés directement par le NMD présentent des signaux de reconnaissance par le NMD (Voir ci-dessus).

De façon intéressante, différentes voies du NMD ont été identifiées selon leur dépendance aux facteurs du NMD. En effet, alors qu'UPF1 et SMG1 sont impliqués dans chacune des voies du NMD identifiées, UPF2, UPF3 ou les composants de cœur de l'EJC semblent altérer l'expression de différents transcrits (Chan *et al*., 2007; Gehring *et al*., 2005; Metze *et al*., 2013). Gehring et collaborateurs ont observé que le NMD induit par la fixation de protéines

de l'EJC (MLN51) sur la région 3'UTR provoque la dégradation des ARNm, d'une façon dépendante de l'EJC, UPF1 et UPF3, mais pas d'UPF2 (Gehring *et al*., 2005). Inversement, UPF2 et RNPS1 induisent le NMD de façon indépendante du cœur de l'EJC. Il a aussi été montré qu'UPF3b était dispensable pour la dégradation de certains ARNm, constituant des branches de NMD dépendantes ou indépendantes d'UPF3b (Chan *et al*., 2007). L'implication de ces différentes voies du NMD au niveau cellulaire n'a pas encore été bien caractérisée, mais on suppose déjà que cela permet d'affiner le niveau de régulation de l'expression des gènes par le NMD.

3.4.5. Le NMD dans la défense antivirale

Les ARN viraux, génomiques et messagers, présentent un ou plusieurs signaux atypiques, qui peuvent activer le NMD lors de leur traduction (voir ci-dessous). De fait, il a été montré que le NMD restreint l'expression d'ARN de virus infectant des animaux mais aussi des plantes, suggérant un rôle antiviral de ce système (Balistreri *et al*., 2014; Garcia *et al*., 2014; LeBlanc et Beemon, 2004; Mocquet *et al*., 2012; Nakano *et al*., 2013).

Ce sont notamment les ARN génomiques des virus à ARNsb positif, PVX (Potato Virus X) chez les plantes et SFV (Semliki Forest Virus) et SINV (Sindbis Virus) chez les mammifères, traduits et présentant de longues 3'UTR, qui sont dégradés par le NMD (Balistreri *et al*., 2014; Garcia *et al*., 2014).

De plus, les ARNm traduits à partir des génomes des rétrovirus, RSV (Rous Sarcoma Virus) et HTLV-1 (Human T-cell Leukemia Virus type 1), sont aussi ciblés par le NMD (LeBlanc et Beemon, 2004; Mocquet *et al*., 2012; Nakano *et al*., 2013). En outre, la dégradation de ces ARN viraux est principalement dépendante d'UPF1, tandis que l'implication d'UPF2, UPF3 ou des facteurs de l'EJC est moins évidente/étudiée. En effet, Balistreri et collaborateurs ont montré l'implication de SMG5 et SMG7 dans la dégradation des ARN du SFV, mais pas d'UPF2 et UPF3 (Balistreri *et al*., 2014).

De façon inattendue, il a été montré que UPF1 participe à l'expression des protéines virales du HIV-1 (Human Immunodefiscience Virus type 1), favorisant l'infection par ce virus (Ajamian *et al*., 2008). Le rôle inattendu joué par UPF1 dans le cycle du HIV-1 pourrait être causé par un détournement de fonction orchestré par le virus afin de favoriser l'expression des gènes viraux. En effet, UPF1 favorise l'export des ARNm viraux vers le cytoplasme (Ajamian *et al*., 2015). Pour ce faire, la protéine virale Rev recrute UPF1, mais exclut UPF2, vers les

ARNm viraux afin de former le complexe responsable de l'export des ARN viraux vers le cytoplasme (Ajamian *et al*., 2015).

Par ailleurs et à l'instar de HIV-1, les virus à ADN ou à ARN pourraient altérer ou corrompre le NMD afin d'établir leurs infections. Plusieurs mécanismes mis au point par des virus afin de protéger leur ARN de la dégradation par le NMD, ont ainsi été identifiés (Quek et Beemon, 2014).

Certains virus comme HCV (Hepatitis C Virus), HTLV-1 et PVX inhibent l'activité du NMD, affectant par la même occasion l'expression de gènes cellulaires, cibles du NMD (Garcia *et al*., 2014; Mocquet *et al*., 2012; Nakano *et al*., 2013; Ramage *et al*., 2015). En outre, HTLV-1 inhibe le NMD *via* l'action des protéines Tax et Rex tandis que la protéine de capside du HCV se lie à la protéine de l'EJC, PYM1 (Mocquet *et al*., 2012; Nakano *et al*., 2013; Ramage *et al*., 2015). En particulier, la protéine Tax d'HTLV-1 inhibe la déphosphorylation d'UPF1 (Mocquet *et al*., 2012). De fait, l'expression des ARN viraux, génomiques et cytoplamiques respectifs de HTLV-1 et PVX corrèle inversement avec l'activité du NMD. De façon intéressante, les souches de PVX les plus virulentes sont celles qui inhibent le plus le NMD (Garcia *et al*., 2014).

Quant au RSV, celui-ci induit la stabilisation spécifique de ses transcrits, s'affranchissant propablement de l'obligation d'altérer le NMD (LeBlanc et Beemon, 2004). Les ARNm du RSV susceptibles d'être reconnus par le NMD contiennent un RSE (RNA Stability Element), qui les protègent du NMD *via* le remodelage des mRNP (Ge *et al*., 2016). En effet, le RSE permet de recruter entre autre la protéine PTB, qui inhibe la fixation d'UPF1 à la région 3'UTR, inhibant ainsi le NMD (Ge *et al*., 2016).

3.4.6. L'expression des facteurs de l'EJC et du NMD est altérée dans les cancers

Les composants de l'EJC et du NMD sont importants dans le développement et dans le maintien des fonctions cellulaires. Il a été ainsi montré qu'une baisse de l'expression d'un des gènes impliqués dans l'EJC et le NMD, conduisait à des défauts du développement engendrant des maladies génétiques voire à une létalité embryonnaire en cas de perte totale de leur expression (Hwang et Maquat, 2011; Nguyen *et al*., 2014; Palacios, 2013). Une altération de la fonction des facteurs du NMD au niveau constitutionnel est donc assez rare. En revanche, une altération somatique d'un ou plusieurs de ces facteurs a été observé dans des cancers. *UPF1* présente des mutations somatiques dans 18 sur 23 échantillons des cancers du

pancréas, lesquelles inhibent son expression protéique (Liu *et al*., 2014a). L'expression d'*UPF1* est diminuée dans les cancers hépatiques, ce qui favorise la croissance tumorale alors que l'expression d'UPF1 est de bon pronostic (Chang *et al*., 2016). Une dérégulation de l'EJC et du NMD participerait à la dérégulation de l'expression des gènes observée dans les cancers (Chang *et al*., 2016).

L'inhibition du NMD pourrait favoriser la carcinogenèse mais aussi la progression tumorale (Popp et Maquat, 2018). En induisant la transition épithélio-mésanchymateuse, l'inhibition du NMD favorise la survie, l'invasion et la dissémination des cellules cancéreuses (Cao *et al*., 2017). De plus, une altération du NMD, qui régule des gènes impliqués dans la réponse aux stress, pourrait fournir un avantage aux cellules cancéreuses, telle une meilleure capacité d'adaptation (Goetz et Wilkinson, 2017). Le NMD pourrait donc être vu comme une voie anti-oncogénique. Par ailleurs, le microenvironnement tumoral peut inhiber le NMD, permettant une meilleure réponse aux stress, et une augmentation de la survie cellulaire favorisant la tumorigenèse (Wang *et al*., 2011b).

Objectifs

Aujourd'hui, le rôle étiologique des HPV-HR dans le développement de nombreux cancers est largement reconnu. Le développement des cancers suite à une infection par un HPV-HR est un processus lent, nécessitant la persistance de l'infection. En particulier, le cancer du col de l'utérus résulte de l'apparition et du développement de lésions précancéreuses CIN1, CIN2 et CIN3 dont la majorité régresse, n'atteignant donc pas le stade cancer. La prévention du cancer du col de l'utérus est réalisée par la vaccination, mais surtout par le dépistage précoce des lésions précancéreuses et leur traitement. Ces mêmes approches peuvent potentiellement être utilisées pour diminuer l'incidence des autres cancers induits par les HPV-HR. Toutefois, il y a une tendance à user de sur-traitement de certaines lésions précancéreuses, au regard de leur potentiel à évoluer en cancer. Le traitement des lésions précancéreuses, par conisation, peut provoquer un risque accru de fauches couches ultérieures, de même que des sur-traitements peuvent représenter des coûts considérables pour la santé publique. Il est donc nécessaire d'avoir une bonne information sur les risques représentés par les lésions CIN identifiées et ce, afin de connaitre la balance bénéfices/risques de chaque patiente. Une meilleure connaissance des mécanismes de carcinogenèse induite par les HPV-HR pourrait permettre d'identifier des marqueurs pronostiques de l'évolution des lésions précancéreuses, permettant ainsi de fournir un meilleur traitement personnalisé à chaque patiente.

La persistance d'une infection par un HPV-HR est l'un des marqueurs les plus associés au développement de lésions de haut grade, mais l'évolution de celle-ci en cancer est loin d'avoir été bien caractérisée. En outre, l'évolution des lésions précancéreuses semble être corrélée à l'intégration du génome viral mais aussi au niveau d'expression des gènes viraux E6 et E7. Ces 2 gènes codent les oncoprotéines E6 et E7 qui sont responsables du pouvoir carcinogène de HPV-HR. D'ailleurs, les propriétés et les fonctions d'E6 et E7 des HPV-HR divergent de celles des HPV-BR, qui sont eux non carcinogènes. En outre E6 et E7 de HPV-HR possèdent non seulement une forte activité sur les cibles partagées avec les HPV-BR, mais aussi d'autres propriétés propres, dont la liaison aux acides nucléiques. L'une des caractéristiques les plus remarquable concernant les gènes E6E7 des HPV-HR réside dans la présence des sites d'épissage permettant de produire des isoformes E6*I, E6*II et E6^E7 dont les rôles sont plus ou moins décrits, qui sont absents chez les HPV-BR.

Afin de mieux appréhender les mécanismes de la carcinogenèse induite par les HPV-HR, il est intéressant d'étudier non seulement les rôles d'E6 et de ses isoformes, mais aussi le mécanisme de régulation de leur expression. Comme décrit dans le chapitre 2, la régulation post-transcriptionnelle revêt un rôle important quant à la régulation de l'expression des

oncogènes viraux E6 et E7, notamment les différentes isoformes. L'épissage est très important dans la régulation de l'expression de la balance E6Fl/E6*I ou E6Fl/E7, mais pourrait aussi s'avérer important dans la régulation du devenir de tous les autres transcrits viraux. En effet, l'épissage permettrait le recrutement de l'EJC sur les ARNm viraux, qui régulera leur métabolisme à plusieurs niveaux. Des nombreux rôles de l'EJC, son rôle dans le NMD représente un intérêt particulier, étant donné ses propriétés récentes de processus anti-viral (voir chapitre 3). D'ailleurs, les transcrits d'HPV16 présentent des caractéristiques inhabituelles (multiples ORF, longues 3'UTR), et peuvent donc être reconnus et dégradés par le NMD.

Notre travail s'est donc focalisé sur :

(i) L'impact des protéines E6 et E6*I sur le transcriptome

(ii) L'étude du rôle d'E6 en tant que protéine de liaison aux ARN

(iii) L'investigation du rôle de l'EJC et du NMD sur la régulation de l'expression des gènes E6 et E7

et (iv) L'étude de l'épissage de l'ORF E6 dans des tumeurs des VADS.

Résultats

1. Rôle de l'isoforme E6*I

1.1. Résumé de la publication 1

Dans cette publication, nous nous sommes intéressés au rôle de l'épissage au sein de l'ORF E6, qui forme des ARNm E6*I. De façon intéressante, seuls les HPV-HR produisent l'isoforme E6*I qui représentent les transcrits majoritaires au sein des ORF E6 et E7. L'absence des sites d'épissage permettant la production d'E6*I chez HPV-BR, non carcinogènes, suggère l'implication de cet épissage particulier dans la carcinogenèse induite par les HPV-HR. D'ailleurs, le taux de transcrits E6*I est corrélé au développement des lésions précancéreuses. Toutefois, le rôle biologique de l'épissage d'E6*I est sujet à débat. En effet, il permettrait l'expression de la protéine E7 à partir des transcrits polycistroniques E6*I-E7, par réinitiation de la traduction. D'autres études suggèrent que ces transcrits permettent la production d'une protéine E6 tronquée, E6*I, laquelle possèdent des activités controversées. D'une part, elle présente des propriétés anti-oncogéniques, inhibant entre autre la dégradation de p53 par E6. De plus, E6*I inhibe la croissance tumorale *in vivo*. D'autre part, en induisant la dégradation de protéines à domaine PDZ, DLG1, PATJ et MAG-1, mais aussi des protéines impliquées dans la réponse au stress, GPX1 et SOD2, E6*I promeut la tumorigenèse. La variation des fonctions d'E6*I selon le contexte cellulaire, rend le rôle de cette protéine encore plus controversée. Afin de mieux appréhender la carcinogenèse liée aux HPV-HR, nous avons voulu étudier le rôle des transcrits E6*I.

Pour cela, nous avons utilisé le RNA-seq afin d'identifier des cibles dérégulées par l'expression ectopique d'E6*I d'HPV16. Nous avons généré des cellules U2OS transfectées stablement avec des vecteurs exprimant E6 et E6*I ou E6*I seule, puis procédé à l'étude de leurs transcriptomes respectifs. Nous avons observé que l'expression des protéines E6*I, seule ou en présence d'E6, n'a pas d'effet sur la prolifération des cellules U2OS, tandis qu'elle a induit une baisse des propriétés clonogéniques de ces cellules. L'expression d'E6 avec E6*I, mais pas d'E6*I seule, augmente la radiosensibilité des cellules. Enfin, l'analyse des transcriptomes par RNA-seq a montré que l'expression ectopique stable des isoformes E6 et E6*I a dérégulé l'expression de 419 gènes. Ces gènes sont notamment impliqués dans des voies relatives à l'expression des gènes viraux, la carcinogenèse virale, la transduction du signal et la traduction. Nous avons aussi trouvé que l'expression de l'isoforme E6*I seule a dérégulé l'expression de 53 gènes dont 11 sont communs avec ceux dérégulés par l'expression d'E6 et E6*I. Les gènes dérégulés par E6*I sont impliqués dans l'organisation de

la matrice extracellulaire, des voies de signalisation cellulaire (MAPK) et d'adhérence cellulaire. De plus, ces gènes dérégulés par l'expression d'E6*I sont communément affectés par le niveau intracellulaire de ROS. De façon intéressante, plusieurs études ont montré l'implication d'E6*I avec la voie des ROS, suggérant que l'effet d'E6*I passe par la production de ROS. E6*I pourrait donc induire un stress oxydatif indépendamment d'E6. Les ROS produits provoquent donc l'activation des voies de signalisation et la réorganisation tissulaire, contribuant à la carcinogenèse. Le stress oxydatif lié à l'augmentation des ROS peut induire une instabilité génomique, favorisant l'intégration du génome viral à celui de la cellule hôte. En somme, un rôle oncogénique d'E6*I qui interviendrait dans la carcinogenèse associée aux HPV-HR serait plus que probable.

Comparative RNA-sequencing reveals that HPV16 E6*I, an E6-truncated isoform, modulates whole gene expression independently of E6

Paget-Bailly Philippe[1,2,§], Meznad Koceila[1,2,§], Perrard Jérôme[1,2], Bruyère Diane[4], Herfs Michael[4], Jung Alain[5], Mougin Christiane[1,2], Prétet Jean-Luc[1,2,3], Baguet Aurélie[1,2*]

[1] EA3181, LabEx LipSTIC ANR-11-LABX-0021, UFR SMP, 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] Centre régional de lutte contre le cancer Paul STRAUSS, University of Strasbourg, Strasbourg, France.

[§] These authors contributed equally to this work

[*]Corresponding author

Aurélie Baguet, UFR SMP, 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

Running title: RNA-sequencing analysis of ectopic expression of HPV16 E6*I isoform

Abstract: 190 words. Text: 4087 words.

Abstract

The high-risk Human Papillomavirus (hr-HPV) infections are responsible of anogenital cancers and a rising number of oropharyngeal cancers cases. Among hr-HPV, HPV16 is the most prevalent type. HPV16 genome expression is tightly regulated at transcriptional and post-transcriptional levels. Modulation in the splice pattern leads to polycistronic HPV16 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 used RNA-sequencing to identify cellular targets deregulated by HPV16 E6*I in an HPV-negative context. The expression of all HPV16 E6 isoforms (E6 and E6*) in HPV-negative cells deregulates the expression of 419 transcripts involved in biological processes such as viral carcinogenesis and protein translation. We found that the expression of E6*I affects the overall transcriptome profile with a deregulation of genes associated with extracellular matrix organization, MAPK signaling and cancer pathways, all related to oxidative stress-mediated signaling events. Taken together, these data suggest that E6*I can act independently of the presence of the full length E6 protein to promote oxidative stress conditions.

Importance

This study is the first report assessing the impact of HPV16 E6*I on cellular whole gene expression using RNA-sequencing. Our findings show that ectopic expression of E6*I is sufficient to modulate the cell transcriptome. Interestingly, we identified some protein-coding genes involved in oxygen deprivation which are deregulated by E6*I. These results confirm the implication of E6*I in cellular-stress response which potentially favor HPV16 DNA integration.

Key words: HPV16, E6-protein isoforms, viral carcinogenesis, whole gene expression

Introduction

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 humans (1). Based on their oncogenic potential, these viruses are classified in high-risk HPV (hr-HPV) including HPV16 and HPV18, and low-risk HPV (lr-HPV) 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. Among a dozen hr-HPV, with sufficient evidence of carcinogenicity to humans, HPV16 is the most prevalent type in HPV-associated cancers (2). 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 ORF (3). 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 two major viral oncoproteins, E6 and E7 but also E6-truncated proteins, E6*I, E6*II and E6^E7 (4) (5). hr-HPV 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 (8) (9) (6). 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 hr-HPVs harbor these cryptic splice sites indicating that E6 ORF splicing events could be relevant for HPV-driven carcinogenesis (10). Moreover, several studies and unpublished data from our laboratory reported increased levels of spliced E6*I RNA correlating

with the severity of cervical lesions (12); (13); Chen et al., 2014(14). Concomitantly, it has been proposed that the ratio of E6*I/E6*II spliced variants can be used as a predictive marker of clinical outcome in HPV-related cervical lesions (13) and oropharyngeal cancers (15). Even if HPV16 transcript detection can be consider as a tool in screening and investigation of HPV-related neoplasia, the biological function of the E6*I protein remains elusive. Nevertheless it has been reported that HPV18 E6*I inhibits E6-mediated degradation of p53 (16), causes the degradation of some PDZ proteins such as Dlg, PATJ and MAG-1 (17) (18) and modulates the expression of a subset of cellular factors involved in stress response as aldo-keto reductase genes (19), superoxide dismutase isoform 2 (SOD2) and glutathione peroxidase 1 (Gpx1) (20). *In vivo* studies showed that ectopic E6*I reduces tumor growth of both HPV-positive and HPV-negative cells in a xenograft nude mouse model (21). Depending on HPV status, E6*I seems implicated in different cellular pathways, and could act independently of the full length E6 protein. To better understand 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 E6*I expression in U-2 OS HPV-negative stable cell lines

HPV-negative U-2 OS cell line was transfected with vectors encoding either all E6 isoforms (pXJ40-E6All) or exclusively E6*I (pXJ40-E6*I) or with an empty vector (Supplementary Figure 1). Western blot analysis was 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. Consequently, the 3F8 antibody does not recognize the E6*I protein, while 6F4 recognizes both isoforms. The full length E6 protein was detected with both E6-directed antibodies, in cells transiently expressing the pXJ40-E6All vector (Figure 1.A). No E6*I protein was observed in pXJ40-E6All or pXJ40-E6*I expressing cells. However MG-132 treatment, a proteasome inhibitor, revealed a specific band with the 6F4 antibody, corresponding to the E6*I isoform (Figure 1.A). After selection of resistant cells to G418, cellular clones stably transfected with the empty vector (U-CTL: clones #1 and #3), pXJ40-E6All (U-E6All: clones #1, #6 and #7) and pXJ40-E6*I (U-E6*I: clones #1, #6 and #14) were isolated and tested for E6 and E6*I expression. The E6 protein is specifically observed in U-E6All stable cell lines and decreased levels of p53 in these cells, containing wild type p53, confirm the expression of a functional E6 protein (Figure 1.B). Unsuccessfully, western blotting analysis failed to detect E6*I protein in both U-E6All and U-E6*I stable cell lines. Because E6*I protein was hardly detectable, E6 and E6*I RNA levels were simultaneously monitored by RT-PCR. 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 RNAs) in comparison with a high amount of E6*I RNA, (SD226/SA409 spliced RNAs) (Figure 1.C). 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), do not express E6 or E6*I. Similar

E6*I RNA levels were detected in both U-E6All and U-E6*I stable cell lines. Overall, splicing efficiency of E6 transcripts in U-E6All cell lines is comparable to those observed in HPV16-positive cancer cell lines (SCC90, SiHa, Ca Ski) and head and neck squamous cell carcinoma (HNSCC) tumor samples (22) (Figures 1.D).

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 assessed by measuring the number of cells once every day during 4 days, show no change in the rate of cell growth between cell lines (Figure 2.A). 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 colony for U-E6All and U-E6*I compared to U-CTL (Figures 2.B and C). It has been previously observed that HPV16 E6 increases the radio sensitivity of HPV-negative cancer cells (23). To confirm this in our cellular models, the radiation sensitivity was measured *via* clonogenic assay following irradiation (Figure 2.D). In all cell lines, survival curves decrease with increasing irradiation dose. 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 (Figures 2.D and E). Although E6 or E6*I had no impact on cell proliferation and seems to reduce the clonogenic survival of tumor cells, we observed that only E6 enhances radio sensitivity of cancer cells. These data suggest that E6 and E6*I can have different isoform-specific functions in HPV-negative cells.

Transcriptome profiles reveal that E6*I regulates cellular targets independently of E6

To assess cellular targets potentially deregulated by E6*I, whole transcriptome of stable cell lines: U-E6All (clones #1, #6 and #7), U-E6*I (clones #1, #7 and #14) and U-CTL (clones #1, #3 and parental U-2 OS cell line named #par) were analyzed by RNA-sequencing. First, unsupervised hierarchical clustering analysis based on RNA-sequencing datasets revealed that the U-E6All gene expression profiles share a great homology and form a cluster whereas U-E6*I gene expression profiles are more heterogeneous (Figure 3.A). Concomitantly, the heat-map representation of U-E6All deregulated genes shows a homogeneous pattern while U-E6*I display a heterogeneous pattern between cellular clones (Supplementary Figures 2.A and B). Venn diagram representation of deregulated genes in U-E6All and U-E6*I compared to U-CTL shows that 419 transcripts are significantly deregulated (FDR<0.05) in U-E6All of which 264 are upregulated and 155 are down-regulated (Figure 3.B and supplementary Table 1). The expression of 53 transcripts are deregulated in U-E6*I cell lines of which 45 are up-regulated and 8 are down-regulated (Figure 3.B and supplementary Table 2). The overlap of 11 deregulated transcripts is observed between U-E6All and U-E6*I cell lines.

Classification using KEGG pathway analysis shows an over-representation of U-E6All deregulated transcripts in several processes such as metabolic pathway, ribosome, pathways in cancer, micro-RNAs in cancer and viral carcinogenesis (Figure 3.C). If U-E6*I deregulated transcripts are also classified in pathway in cancer many of them are implicated in others pathways: MAPK signaling pathway, focal adhesion, cytokine-cytokine receptor interaction, regulation of actin cytoskeleton, IL-17 signaling pathway and ras signaling pathway (Figure 3.D).

These results show that E6*I has a moderate impact on cellular transcriptome but suggest that E6*I could regulate gene expression independently of E6.

RNA-sequencing validation of transcripts deregulated by E6*I

To confirm RNA-seq data, RT-qPCR analyses were performed on the 11 transcripts deregulated in both U-E6All and U-E6*I cell lines and on 8 transcripts homogenously deregulated in each U-E6*I clone. For all transcripts tested, Pearson correlation coefficients (r) were calculated. Except for the two long intergenic non-coding (linc) RNAs, PICSAR (r=0.697) and CH507-210p18.1 (r=0.747), high r-values (from 0.815 to 0.995) were observed, indicating a good concordance between RNA-seq and RT-qPCR data (Table 1).

Based on protein-coding transcripts deregulated in U-E6*I cell line, a protein interaction network was constructed using the STRING database (Figure 4A). Interestingly, the highest connectivity within the network are proteins involved in extracellular matrix organization: [A2M] (*Alpha-2-macroglobulin*), [MMP3] (*Matrix Metalloproteinase 3*), [P4HA1] (*Prolyl 4-hydroxylase, alpha polypeptide I*), [COL22A1] (*Collagen, type XXII, alpha 1*), [ITGA11] (*Integrin, alpha 11*), in response to decreased oxygen levels: [CCL2] (*C-C Motif Chemokine Ligand 2*), [CRYAB] (*CRYstallin Alpha B*) also named HSPB5 (*Heat Shock Protein Beta-5*), [ARNT2] (Aryl-hydrocarbon receptor nuclear translocator 2), [STC1] (*Stanniocalcin* 1) or in both processes: [CTGF] (*Connective Tissue Growth Factor*) also known as IGFBP8 (*Insulin-like Growth Factor Binding Protein 8*), [PDGFB] (*Platelet-derived growth factor beta polypeptide*), [MMP13] (*Matrix Metalloproteinase 13*).

Using RT-qPCR, we confirmed the upregulation of oxidative stress response-related genes in U-E6*I cell lines such as A2M (up to 13-fold increase), ADGRL4 (*Adhesion G Protein-*

182 *Coupled Receptor L4*) also known as ELTD1 (*EGF Latrophilin and seven Transmembrane*
183 *Domain 1*) (up to 7 fold increase), CCL2 (up to 35 fold increase), CRYAB (up to 7-fold
184 increase), CNN1 (*Calponin 1*) (up to 7-fold increase), CTGF (up to 5-fold increase) and MMP3
185 (*Matrix Metalloproteinase 3*) (up to 5-fold increase) (Figure 4B). Thus, we show that RT-qPCR
186 results are in agreement with the RNA-seq analysis and that E6*I expression could deregulate
187 some genes involved in extracellular matrix organization and oxidative stress response.
188
189
190

191 ## Discussion

192 To our knowledge, this study is the first report assessing the impact of HPV16 E6*I on
193 cellular gene expression using RNA-sequencing. The aim of this study was to uncover target
194 genes specifically modulated by E6*I independently of the full length E6 protein. To this end,
195 stable cell lines expressing either all E6 protein isoforms (U-E6All) or exclusively E6*I (U-
196 E6*I), were generated from U-2 OS. The U-2 OS cell line constitutes an adequate model for
197 studying HPV gene expression and HPV-associated carcinogenesis because it expresses a wild-
198 type p53 which can be targeted by E6. Furthermore, it has been shown that viral gene expression
199 of both alpha and beta papillomaviruses in U-2 OS is very similar to that described in
200 keratinocytes ((24) (25). Immunoblotting experiments allowed to confirm the expression of a
201 functional E6 by the loss of p53 in U-E6All cell lines (26). Moreover, E6 early transcripts were
202 efficiently spliced in U-E6All cell lines, yielding high amount of E6*I transcripts as observed in
203 HPV16-positive cell lines and HNSCC clinical samples. However, although the E6*I protein was
204 detected in transient transfection, no protein has been observed in stable cell lines probably due to

the extremely short half-life of E6*I isoform. This has already been documented in a previous study which reported the detection of an E6*I fusion protein following transient transfection but failed to detect E6*I in stable cell lines (27). Emerging roles of non-coding RNAs in gene expression regulation and the noticeable difficulties to detect the E6*I protein by western blotting invite us to consider that the transcriptome deregulation observed in the present study could be mostly mediated by HPV16 E6*I RNA rather than the E6*I protein. 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. As previously referred, we confirmed that E6 isoforms or E6*I expression in HPV-negative cells has no impact on cell proliferation (16); (28). As observed in a previous study (29), E6 expression inhibits colony formation. Here, we showed that E6*I has a stronger inhibitory effect on colony formation. 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 partially be explained by the *TP53* wild-type status of our cellular model (U-2 OS). Indeed, in contrast to E6*I, E6 oncoprotein was shown to disrupt several cellular functions, including the p53-mediated cellular responses to DNA damage (30). Moreover, E6 from various HPV genotypes directly binds XRCC1 inducing altered single-strand break repair and genetic instability (31). Altogether, our findings support the higher radio sensitivity displayed by HPV-positive neoplasms compared to their virus-unrelated counterparts (32).

Using RNA-sequencing, we found 419 significantly deregulated transcripts in U-E6All cells. Among them, a substantial number of transcripts was already described in the literature as deregulated in HPV16 cellular models or HPV16-positive clinical samples. Indeed, the down-regulation of transcripts encoding ribosomal subunits has been previously reported in cervical cancers (33). Hence KEGG pathway and GO enrichment analyses show viral gene expression

229 and protein synthesis related terms enriched. In addition, we confirm the upregulation of SYCP2,
230 Nup210 and DICER1 whose expression has been previously shown to be increased in HPV16-
231 positive clinical samples from cervical (33), head and neck cancers (34), and HPV-negative cell
232 lines expressing both E6 and E7 proteins (35) (36).

233 Comparative RNA-sequencing analysis reveals that E6*I expression is sufficient to
234 modulate the cellular transcriptome. However, this modulation is moderated since 53 transcripts
235 were deregulated including only 11 transcripts found deregulated in both U-E6All and U-E6*I,
236 indicating that E6*I can act independently of E6. KEGG pathway and GO enrichment analyses
237 show that E6*I expression is associated with cancer pathway, mitogen activated protein kinase
238 (MAPK), cytokines, focal adhesion, growth factor binding and extracellular matrix organization.
239 In order to validate RNA-sequencing analysis, RT-qPCR were performed on some identified
240 genes and allowed to demonstrate that fold changes estimated from RNA-seq had high
241 correlation with that from the RT-qPCR.

242 Recently, a whole transcriptome study showed that HPV16 modulates genes implicated in
243 the cellular metabolism, inflammatory response and immune signaling pathway (37). Using
244 STRING analysis, we identified some protein-coding genes, deregulated by E6*I involved in
245 oxygen deprivation. These pathways are closely associated with the level of reactive oxygen
246 species (ROS). Indeed, ROS impact several processes such as cellular signaling leading to the
247 modulation of cell proliferation and apoptosis, but also the regulation of cytokines, chemokines
248 and growth factors. Antioxidant defense system exists to limit ROS and maintain cellular
249 homeostasis. Viruses are known to induce oxidative stress by modulating the level of ROS (38)
250 (39)). Previous studies have linked high HPV16 E6*I expression to increased level of ROS,
251 changes in the activity of antioxidant enzymes and DNA damage accumulation (40); (20); (41).
252 In our study, no change was observed in SOD and Gpx mRNA levels, probably due to a post-

translational regulation of these enzymes. But several of E6*I-deregulated genes are directly involved in ROS metabolism or are induced by ROS: A2M (42), MMP3 (43); (44), CRYAB ((45), CCL2 (46), CTGF (47), PTGFB (48); (49) and RAC2 (50) ; (51).

A2M, frequently increased during viral infections, is a key regulator of the inflammatory response, regulating MMPs and inflammatory cytokines (52); (53) ; (54) ; (55). MMP3 and MMP13 gene loci amplifications are strongly associated with HPV-positive cancers (56) (57). Interestingly, MMP3 can act as a transcription activator of both CTGF (58) and CRYAB (59), also found upregulated in our model expressing E6*I. CTGF, a growth factor is upregulated in HNSCC (60). While CRYAB, a small heat shock protein (HSPB5) promote stressed cell survival by strongly inhibiting apoptotic inducers and effectors (Aggeli et al, 2008), and has been associated with HPV16 viral life cycle (61) and metastatic progression in HNSCC (62). CCL2, a well-known cytokine produced by inflammatory cells is associated with poor survival in cervical cancers (63). Interestingly, the oxidative stress associated factors upregulated by E6*I involved in the neoplasic progression, inflammation, extracellular matrix reorganization could favor cell invasion.

Oxidative stress plays crucial roles in cancer because of its role in mutagenesis, initiation and neoplastic progression in the context of HPV-driven carcinogenesis. Oxidative stress, induced by E6*I, could also facilitate the integration of HPV genomes into the host cell genome (64); (65)). This hypothesis is consistent with high levels of spliced E6*I RNA which correlate with the severity of cervical lesions (12); (13); Chen et al., 2014; (14). Consistently, it has been reported that non-carcinogenic lr-HPVs, whose genome doesn't integrate into the host cell genome, do not contain E6 ORF splice sites allowing the production of E6*I. E6 ORF alternative splicing could be used by the virus to (1) allow the production of E6-truncated protein isoforms (66), (2) release of the intron 1 which could be process as a mirtron-like product, corresponding

to an unconventional miRNA production generated from an excised intron (67) and/or (3) regulate both E6 and E7 expression (68) (Figure 5). These strategies could be, temporally or simultaneously, used by the virus to promote viral life cycle, genome integration or to respond to various environmental stress conditions (irradiation, chemotherapy, inflammation, oxygen deprivation, …).

In conclusion, our study reports cellular transcriptome modulation induced by HPV16 E6*I using RNA-sequencing. We identified cellular genes involved in ROS stress response deregulated by the ectopic expression of E6*I. In agreement with previous studies, we proposed that E6*I, could promote viral genome integration into host genome by inducing ROS and subsequently increased genome instability. Subsequent studies are needed to fully understand how oxidative stress conditions mediated by E6*I could affect HPV-driven carcinogenesis.

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). Cell lines were incubated at 37°C in a humidified atmosphere with 5% of CO_2.

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 pXJ40-E6*I vector, lacking the intron 1 corresponding to the nucleotides from 227 to 408 was generated from the pXJ40-E6All vector, using Q5Site-Directed Mutagenesis Kit (New England Biolabs) and the following primers: MutHPV16E6*I-F: GTGTATTAACTGTCAAAAGCC and MutHPV16E6*I-R: CTCACGTCGCAGTAACTG. 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 with in the presence or absence of 10µM of proteasome inhibitor MG132 for 16h and harvested after 48h 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 analysis and clonogenic assays. Cell proliferation assay was done with 10,000 cells in 12-well plate by counting cells each day for 4 days using the EVE^{TM} automatic cell counter (Nano EnTek). Clonogenic assay was performed by seeding 500 cells of each U-CTL, U-E6All and U-E6*I cell lines in 10cm dishes. After 14 days of culture, colony were counted after crystal violet staining. Clonogenic radio sensitivity assay was performed by seeding 300 cells in 6-well plate 24h 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), two wells were analyzed for colonies counting.

Immunoblot assays. U-2 OS cells were lysed using RIPA buffer (50 mMTris-HCl pH7.4, 150 mM NaCl, 1% NP40, 0.5% Na deoxycholate, 1mM EDTA) supplemented with 30 µg/mL anti-protease mixture (Roche Diagnostics) followed by sonication. After centrifugation at 10,000 X *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. Twenty or 40 µg 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 Amersham^{TM}Hybond^{TM}-PVDF membrane (GE Healthcare) and blocking 1h with 5% nonfat milk in Tris-buffered Saline-Tween20, membranes were probed with primary antibodies for 2h then with HRP-conjugated secondary antibodies (see below) for 1h. Chemiluminescent signal was detected using PierceECL 2 Western Blotting Substrate (ThermoScientific).

Antibodies. Mouse anti-E6 antibodies 1E-6F4 (E6 N-ter) and 2E-3F8 (E6 C-ter) were purchased from Euromedex. Anti-p53 (DO-1) and anti-β-actin (AC-15) monoclonal mouse antibodies were obtained from Cell Signaling and Sigma-Aldrich respectively. 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 µg were retro transcribed in 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). Primer sequences used are listed in Supplementary Table 1. RPLP0 expression levels was used as internal normalization standards. The $2^{-\Delta\Delta Ct}$ method was used for relative quantification.

RNA sequencing. RNA-seq was carried out to compare 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 1µg of total RNA using TrueSeq Stranded mRNA Sample

preparation kit (Illumina, Inc.). Sequencing was carried on 75 bp in single read on the NextSeq 500 platform using the NextSeq 500 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) (69) 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 transcript assembly and to assess transcript abundance which was normalized between all nine samples using the fragments per kilobase per million mapped reads (FPKM) method to correct transcript length and library size bias (70). Differentially expressed transcripts between control condition U-CTL and U-E6All or U-E6*I conditions were finally identified using Cuffdiff (Trapnell Laboratory, University of Washington, USA). Transcripts with a false discovery rate (FDR)<0,05 were considered significantly deregulated.

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 analysis were performed using Gene Ontology (GO) database v1.2 on PANTHER v13.0 (71), Kyoto Encyclopedia of Genes and Genomes (KEGG) release 81.0 (72) and STRING v10.5 (73).

388 **Statistics.** Statistical analyses were carried out using XLSTAT v2017.4 (Addinsoft, Paris,
389 France). Pearson's correlation coefficients were used to investigate correlation between the RNA-
390 sequencing and RT-qPCR analyses. Mann-Whitney tests were used to analyze clonogenic assays
391 data and determine significant differences in the number of colonies formed (*: $p<0.05$; **:
392 $p<0.01$; ***: $p<0.001$). All qPCR and clonogenic experiments were done in triplicate and the
393 error bars represent the standard deviation.
394

395 ## Acknowledgements

396 The authors wish to thank Muriel Masson for kindly providing the pXJ40-E6All vector,
397 Julie Frejaville and Anne Peigney for the technical assistance. K.M. and P.P. had fellowships
398 from the French ministry of research and technology. J.P. had fellowships from the Conseil
399 Régional de Franche-Comté. Funds were from Cancéropôle du Grand-Est (CGE), La Ligue
400 Contre le Cancer (Comité du Doubs and CCIR-GE), Université de Franche-Comté (UFC) and
401 Région Franche-Comté.
402

403 ## References

404 1. Schiffman M, Boyle S, Raine-Bennett T, Katki HA, Gage JC, Wentzensen N, Kornegay
405 JR, Apple R, Aldrich C, Erlich HA, Tam T, Befano B, Burk RD, Castle PE. 2015. The role of
406 human papillomavirus (HPV) genotyping in cervical cancer screening: A large-scale evaluation
407 of the cobas HPV test. Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored
408 Am Soc Prev Oncol 24:1304–1310.

409 2. de Martel C, Plummer M, Vignat J, Franceschi S. 2017. Worldwide burden of cancer
410 attributable to HPV by site, country and HPV type. Int J Cancer 141:664–670.

411 3. Graham SV. 2017. Keratinocyte Differentiation-Dependent Human Papillomavirus Gene
412 Regulation. Viruses 9:245.

413 4. Johansson C, Schwartz S. 2013. Regulation of human papillomavirus gene expression by
414 splicing and polyadenylation. Nat Rev Microbiol 11:239–251.

5. Chen J, Xue Y, Poidinger M, Lim T, Chew SH, Pang CL, Abastado J-P, Thierry F. 2014. Mapping of HPV transcripts in four human cervical lesions using RNAseq suggests quantitative rearrangements during carcinogenic progression. Virology 462:14–24.

6. Moody CA, Laimins LA. 2010. Human papillomavirus oncoproteins: pathways to transformation. Nat Rev Cancer 10:550.

7. McLaughlin-Drubin ME, Münger K. 2009. The human papillomavirus E7 oncoprotein. Virology 384:335–344.

8. Dyson N, Howley PM, Münger K, Harlow E. 1989. The Human Papilloma Virus--16 E7 Oncoprotein is Able to Bind to the Retinoblastoma Gene Product. Science 243:934–937.

9. Werness BA, Levine AJ, Howley PM. 1990. Association of human papillomavirus types 16 and 18 E6 proteins with p53. Science 248:76–80.

10. Smotkin D, Prokoph H, Wettstein FO. 1989. Oncogenic and nononcogenic human genital papillomaviruses generate the E7 mRNA by different mechanisms. J Virol 63:1441–1447.

11. Cornelissen MTE, Smits HL, Briet MA, van den Tweel JG, Struyk APHB, van der Noordaa J, 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.

12. Kösel S, Burggraf S, Engelhardt W, 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.

13. Cricca M, Venturoli S, Leo E, Costa S, Musiani M, Zerbini M. 2009. 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.

14. Lin K, Lin K, Lu X, Lu X, Chen J, Chen J, Zou R, Zou R, Zhang L, Zhang L, Xue X, Xue X. 2015. E6-associated transcription patterns in human papilloma virus 16-positive cervical tissues. Oncol Lett 9:478–482.

15. Hong A, Zhang X, Jones D, Zhang M, Lee CS, Lyons JG, Veillard A-S, Rose B. 2016. E6 viral protein ratio correlates with outcomes in human papillomavirus related oropharyngeal cancer. Cancer Biol Ther 17:181–187.

16. Pim D, Massimi P, Banks L. 1997. Alternatively spliced HPV-18 E6* protein inhibits E6 mediated degradation of p53 and suppresses transformed cell growth. Oncogene 15:257–264.

17. Storrs CH, Silverstein SJ. 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.

18. Pim D, Tomaić V, 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.

19. Wanichwatanadecha P, Sirisrimangkorn S, Kaewprag J, 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.

20. Williams VM, Filippova M, Filippov V, Payne KJ, Duerksen-Hughes P. 2014. Human Papillomavirus Type 16 E6* Induces Oxidative Stress and DNA Damage. J Virol 88:6751–6761.

21. Filippova M, Evans W, Aragon R, Filippov V, Williams VM, Hong L, Reeves ME, Duerksen-Hughes P. 2014. The small splice variant of HPV16 E6, E6□, reduces tumor formation in cervical carcinoma xenografts. Virology 450–451:153–164.

22. Jung AC, Briolat J, Millon R, de Reyniès A, Rickman D, Thomas E, Abecassis J, Clavel C, Wasylyk B. 2010. Biological and clinical relevance of transcriptionally active human papillomavirus (HPV) infection in oropharynx squamous cell carcinoma. Int J Cancer NA-NA.

23. Shin H-J, Kim J-Y, Hampson L, Pyo H, Baek HJ, Roberts SA, Hendry JH, Hampson IN. 2010. 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.

24. Toots M, Männik A, Kivi G, Jr MU, Ustav E, Ustav M. 2014. The Transcription Map of Human Papillomavirus Type 18 during Genome Replication in U2OS Cells. PLOS ONE 9:e116151.

25. Isok-Paas H, Männik A, Ustav E, Ustav M. 2015. The transcription map of HPV11 in U2OS cells adequately reflects the initial and stable replication phases of the viral genome. Virol J 12:59.

26. Scheffner M, Werness BA, Huibregtse JM, Levine AJ, Howley PM. 1990. The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. Cell 63:1129–1136.

27. Filippova M, Filippov VA, Kagoda M, Garnett T, Fodor N, Duerksen-Hughes PJ. 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.

28. Sedman SA, Barbosa MS, Vass WC, Hubbert NL, Haas JA, Lowy DR, Schiller JT. 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.

29. Ristriani T, Fournane S, Orfanoudakis G, Travé G, 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.

30. Kessis TD, Slebos RJ, Nelson WG, Kastan MB, Plunkett BS, Han SM, Lorincz AT, Hedrick L, Cho KR. 1993. Human papillomavirus 16 E6 expression disrupts the p53-mediated cellular response to DNA damage. Proc Natl Acad Sci 90:3988–3992.

31. Iftner T, Elbel M, Schopp B, Hiller T, Loizou JI, Caldecott KW, Stubenrauch F. 2002. Interference of papillomavirus E6 protein with single- strand break repair by interaction with XRCC1. EMBO J 21:4741–4748.

32. Mirghani H, Amen F, Tao Y, Deutsch E, Levy A. 2015. Increased radiosensitivity of HPV-positive head and neck cancers: Molecular basis and therapeutic perspectives. Cancer Treat Rev 41:844–852.

33. Boon JA den, Pyeon D, Wang SS, Horswill M, Schiffman M, Sherman M, Zuna RE, Wang Z, Hewitt SM, Pearson R, Schott M, Chung L, He Q, Lambert P, Walker J, Newton MA, Wentzensen N, Ahlquist P. 2015. Molecular transitions from papillomavirus infection to cervical precancer and cancer: Role of stromal estrogen receptor signaling. Proc Natl Acad Sci 112:E3255–E3264.

34. Slebos RJC, Yi Y, Ely K, Carter J, Evjen A, Zhang X, Shyr Y, Murphy BM, Cmelak AJ, Burkey BB, Netterville JL, Levy S, Yarbrough WG, Chung CH. 2006. Gene Expression Differences Associated with Human Papillomavirus Status in Head and Neck Squamous Cell Carcinoma. Clin Cancer Res 12:701–709.

35. Pyeon D, Newton MA, Lambert PF, den Boon JA, Sengupta S, Marsit CJ, Woodworth CD, Connor JP, Haugen TH, Smith EM, Kelsey KT, Turek LP, Ahlquist P. 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.

36. Harden ME, Munger K. 2017. Perturbation of DROSHA and DICER expression by human papillomavirus 16 oncoproteins. Virology 507:192–198.

37. Klymenko T, Gu Q, Herbert I, Stevenson A, Iliev V, Watkins G, Pollock C, Bhatia R, Cuschieri K, Herzyk P, Gatherer D, Graham SV. 2017. 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.

38. Paracha UZ, Fatima K, Alqahtani M, Chaudhary A, Abuzenadah A, Damanhouri G, Qadri I. 2013. Oxidative stress and hepatitis C virus. Virol J 10:251.

39. Reshi ML, Su Y-C, Hong J-R. 2014. RNA Viruses: ROS-Mediated Cell Death. Int J Cell Biol. Research article.

40. Mouret S, Sauvaigo S, Peinnequin A, Favier A, Beani J-C, 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.

41. Evans W, Filippova M, Filippov V, Bashkirova S, Zhang G, Reeves ME, 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.

42. Chen X, Kong X, Zhang Z, Chen W, Chen J, Li H, Cao W, Ge Y, Fang S. 2014. Alpha-2-macroglobulin as a radioprotective agent: a review. Chin J Cancer Res 26:611–621.

526 43. Radisky DC, Levy DD, Littlepage LE, Liu H, Nelson CM, Fata JE, Leake D, Godden EL,
527 Albertson DG, Nieto MA, Werb Z, Bissell MJ. 2005. Rac1b and reactive oxygen species mediate
528 MMP-3-induced EMT and genomic instability. Nature 436:123–127.

529 44. Choi D-H, Kim J-H, Seo J-H, Lee J, Choi WS, Kim Y-S. 2014. Matrix Metalloproteinase-
530 3 Causes Dopaminergic Neuronal Death through Nox1-Regenerated Oxidative Stress. PLoS ONE
531 9.

532 45. van de Schootbrugge C, Schults EM, Bussink J, Span PN, Grénman R, Pruijn GJ,
533 Kaanders JH, Boelens WC. 2014. Effect of hypoxia on the expression of αB-crystallin in head
534 and neck squamous cell carcinoma. BMC Cancer 14:252.

535 46. Li X, Xu Q, Wu Y, Li J, Tang D, Han L, Fan Q. 2014. A CCL2/ROS autoregulation loop
536 is critical for cancer-associated fibroblasts-enhanced tumor growth of oral squamous cell
537 carcinoma. Carcinogenesis 35:1362–1370.

538 47. Murphy-Marshman H, Quensel K, Shi-wen X, Barnfield R, Kelly J, Peidl A, Stratton RJ,
539 Leask A. 2017. Antioxidants and NOX1/NOX4 inhibition blocks TGFβ1-induced CCN2 and α-
540 SMA expression in dermal and gingival fibroblasts. PLoS ONE 12.

541 48. Chen KC-W, Zhou Y, Zhang W, Lou MF. 2007. Control of PDGF-induced reactive
542 oxygen species (ROS) generation and signal transduction in human lens epithelial cells. Mol Vis
543 13:374–387.

544 49. Adachi T, Togashi H, Suzuki A, Kasai S, Ito J, Sugahara K, Kawata S. 2005. NAD(P)H
545 oxidase plays a crucial role in PDGF-induced proliferation of hepatic stellate cells. Hepatology
546 41:1272–1281.

547 50. Diebold BA, Bokoch GM. 2001. Molecular basis for Rac2 regulation of phagocyte
548 NADPH oxidase. Nat Immunol 2:211–215.

549 51. Nieborowska-Skorska M, Kopinski PK, Ray R, Hoser G, Ngaba D, Flis S, Cramer K,
550 Reddy MM, Koptyra M, Penserga T, Glodkowska-Mrowka E, Bolton E, Holyoake TL, Eaves CJ,
551 Cerny-Reiterer S, Valent P, Hochhaus A, Hughes TP, Kuip H van der, Sattler M, Wiktor-
552 Jedrzejczak W, Richardson C, Dorrance A, Stoklosa T, Williams DA, Skorski T. 2012. Rac2-
553 MRC-cIII–generated ROS cause genomic instability in chronic myeloid leukemia stem cells and
554 primitive progenitors. Blood 119:4253–4263.

555 52. Gangadharan B, Antrobus R, Dwek RA, Zitzmann N. 2007. Novel Serum Biomarker
556 Candidates for Liver Fibrosis in Hepatitis C Patients. Clin Chem 53:1792–1799.

557 53. Ho A-S, Cheng C-C, Lee S-C, Liu M-L, Lee J-Y, Wang W-M, Wang C-C. 2010. Novel
558 biomarkers predict liver fibrosis in hepatitis C patients: alpha 2 macroglobulin, vitamin D binding
559 protein and apolipoprotein AI. J Biomed Sci 17:58.

560 54. Brasier AR, Zhao Y, Wiktorowicz JE, Spratt HM, Nascimento EJM, Cordeiro MT,
561 Soman KV, Ju H, Recinos A, Stafford S, Wu Z, Marques ETA, Vasilakis N. 2015. Molecular
562 classification of outcomes from dengue virus -3 infections. J Clin Virol 64:97–106.

55. Chen C-H, Zhang X-Q, Lo C-W, Liu P-F, Liu Y-T, Gallo RL, Hsieh M-F, Schooley RT, Huang C-M. 2010. The essentiality of α-2-macroglobulin in human salivary innate immunity against new H1N1 swine origin influenza A virus. PROTEOMICS 10:2396–2401.

56. Bhosale PG, Pandey M, Cristea S, Shah M, Patil A, Beerenwinkel N, Schäffer AA, Mahimkar MB. 2017. Recurring Amplification at 11q22.1-q22.2 Locus Plays an Important Role in Lymph Node Metastasis and Radioresistance in OSCC. Sci Rep 7.

57. Grozdanov P, Hadjidekova S, Dimova I, Nikolova I, Toncheva D, Ganchev G, Zlatkov V, Galabov AS. 2016. Characterization of genomic changes in the cervical pre-cancerous lesions and tumors induced by different types of human papillomaviruses. VirusDisease 27:271–276.

58. Eguchi T, Kubota S, Kawata K, Mukudai Y, Uehara J, Ohgawara T, Ibaragi S, Sasaki A, Kuboki T, Takigawa M. 2008. Novel transcription-factor-like function of human matrix metalloproteinase 3 regulating the CTGF/CCN2 gene. Mol Cell Biol 28:2391–2413.

59. Eguchi T, Calderwood SK, Takigawa M, Kubota S, Kozaki K. 2017. Intracellular MMP3 Promotes HSP Gene Expression in Collaboration With Chromobox Proteins. J Cell Biochem 118:43–51.

60. Mullis TC, Tang X, Chong KT. 2008. Expression of connective tissue growth factor (CTGF/CCN2) in head and neck squamous cell carcinoma. J Clin Pathol 61:606–610.

61. Wan F, Miao X, Quraishi I, Kennedy V, Creek KE, Pirisi L. 2008. Gene expression changes during HPV- mediated carcinogenesis: A comparison between an in vitro cell model and cervical cancer. Int J Cancer 123:32–40.

62. van de Schootbrugge C, Bussink J, Span PN, Sweep FC, Grénman R, Stegeman H, Pruijn GJ, Kaanders JH, Boelens WC. 2013. αB-crystallin stimulates VEGF secretion and tumor cell migration and correlates with enhanced distant metastasis in head and neck squamous cell carcinoma. BMC Cancer 13:128.

63. Zijlmans HJM a. A, Fleuren GJ, Baelde HJ, Eilers PHC, Kenter GG, Gorter A. 2006. The absence of CCL2 expression in cervical carcinoma is associated with increased survival and loss of heterozygosity at 17q11.2. J Pathol 208:507–517.

64. Williams VM, Filippova M, Soto U, Duerksen-Hughes PJ. 2010. HPV-DNA integration and carcinogenesis: putative roles for inflammation and oxidative stress. Future Virol 6:45–57.

65. Camini FC, Caetano CC da S, Almeida LT, Magalhães CL de B. 2017. Implications of oxidative stress on viral pathogenesis. Arch Virol 162:907–917.

66. Pim D, 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.

67. Czech B, Hannon GJ. 2011. Small RNA sorting: matchmaking for Argonautes. Nat Rev Genet 12:19.

68. Tang S, Tao M, McCoy JP, 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.

69. Trapnell C, Pachter L, Salzberg SL. 2009. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25:1105–1111.

70. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5:621–628.

71. Mi H, Huang X, Muruganujan A, Tang H, Mills C, Kang D, Thomas PD. 2017. PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. Nucleic Acids Res 45:D183–D189.

72. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. 1999. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res 27:29–34.

73. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P, Jensen LJ, von Mering C. 2017. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res 45:D362–D368.

Figures Legends

Figure 1. Ectopic expression of HPV16 E6 and E6*I in U-2 OS HPV-negative cell line. A. 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) or pXJ40-E6*I vector (E6*I). After 48h incubation, cells were treated either with DMSO or 10µM MG132 proteasome inhibitor for 16h. The membrane was then incubated with two E6 antibodies: the 6F4 antibody which recognizes the N-terminal part of the E6 protein and the 3F8 antibody which targets the C-terminal part of E6 protein. 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 11kDa corresponding to a cleavage product observed in U-E6All cells treated with MG132. B. Western blotting analysis showing p53, E6 and β-actin protein levels in U-2OS 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. 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-2OS 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. 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 400bp and could correspond to heteroduplexes between the different amplicons generated during PCR reaction.

Figure 2. 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 means +/- S.D. of three 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 means +/- S.D. of three independent experiments.D/E. 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).

Figure 3. 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 two pairwise comparisons U-E6All vs. U-CTL and U-E6*I vs. U-CTL. We observed that 11 transcripts were deregulated in both U-E6All and U-E6*I cell lines compared to U-CTL. C. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways terms for deregulated genes in U-E6All vs. U-CTL. D. GO and KEGG pathways terms for deregulated genes in U-E6*I vs. U-CTL. Red and blue bars indicate GO and KEGG pathways respectively.

Figure 4. Interactome and RNA levels of deregulated genes in U-E6*I cell lines compared to U-CTL. A. Protein interaction network corresponding to the protein-encoded genes deregulated in U-E6*I cell lines using STRING database. Line thickness indicates the strength of data support. In blue and red proteins involved in extracellular matrix organization and response to decreased oxygen levels, respectively. B. Quantification by RT-qPCR of A2M, ADGRL4, CCL2, CNN1, CRYAB, CTGF and MMP3 RNA levels. Values are represented as mean +/- S.D. of 3 independent experiments. RPLP0 is used as housekeeping gene.

Table 1. Validation of RNA-sequencing data by RT-qPCR. Significant changes in gene expression (Log_2 fold change) in (A) U-E6All and U-E6*I compared to U-CTL and (B) in U-E6*I compared to U-CTL. Pearson coefficient (r) represent the correlation between RNA-seq and qPCR data.

Figure 5: Potential significance of HPV16 E6*I splicing. The p97 promotor directs the expression of HPV16 E6/E7 transcripts submitted to highly orchestrated alternative splicing.

E6/E7 RNA splicing depends of the recognition of several sequence motifs by the cellular spliceosome, including 1 splice donor (SD226) and 3 splice acceptor sites (SA409, SA526 and SA742). The absence of splicing leads to the production of the full length E6 protein while the recognition of SD226/SA409 leads to the elimination of a part of the E6 ORF (intron 1) generating a frameshift and a premature termination codon within the same exon (TAA233). We propose that [SD226/SA409] alternative splicing could mediate (1) the production of a truncated E6 protein named E6*I and/or (2) the release of the intron 1 which could be process as a mirtron-like product and/or (3) the up-regulation of E7 expression.

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. 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 $\log_2$ 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 $\log_2$ scale of fold change relative to the average of each transcript across all 6 clones.

Table S1. Differential expression analysis of transcripts between U-E6All vs. U-CTL cell lines.

Table S2. Differential expression analysis of transcripts between U-E6*I vs. U-CTL cell lines.

Table S3. List of RT-PCR and RTqPCR primers

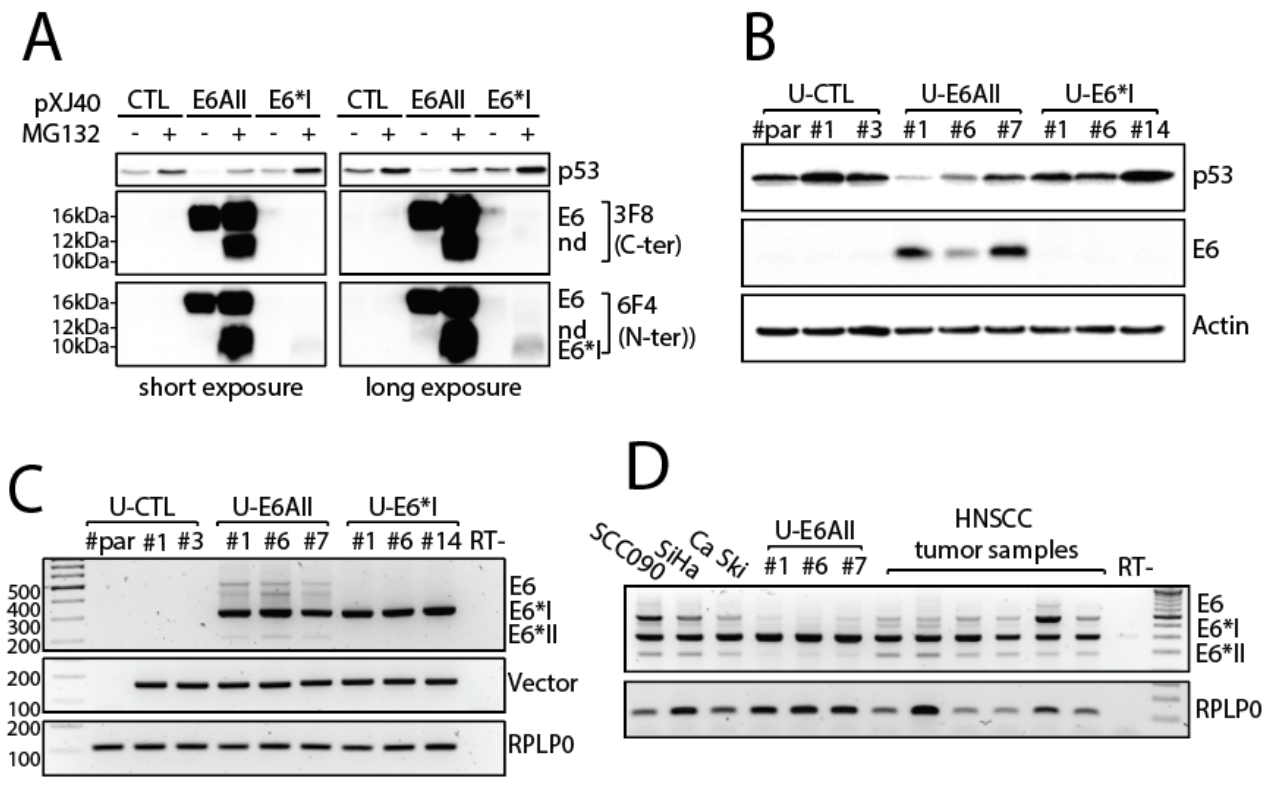


Fig 1

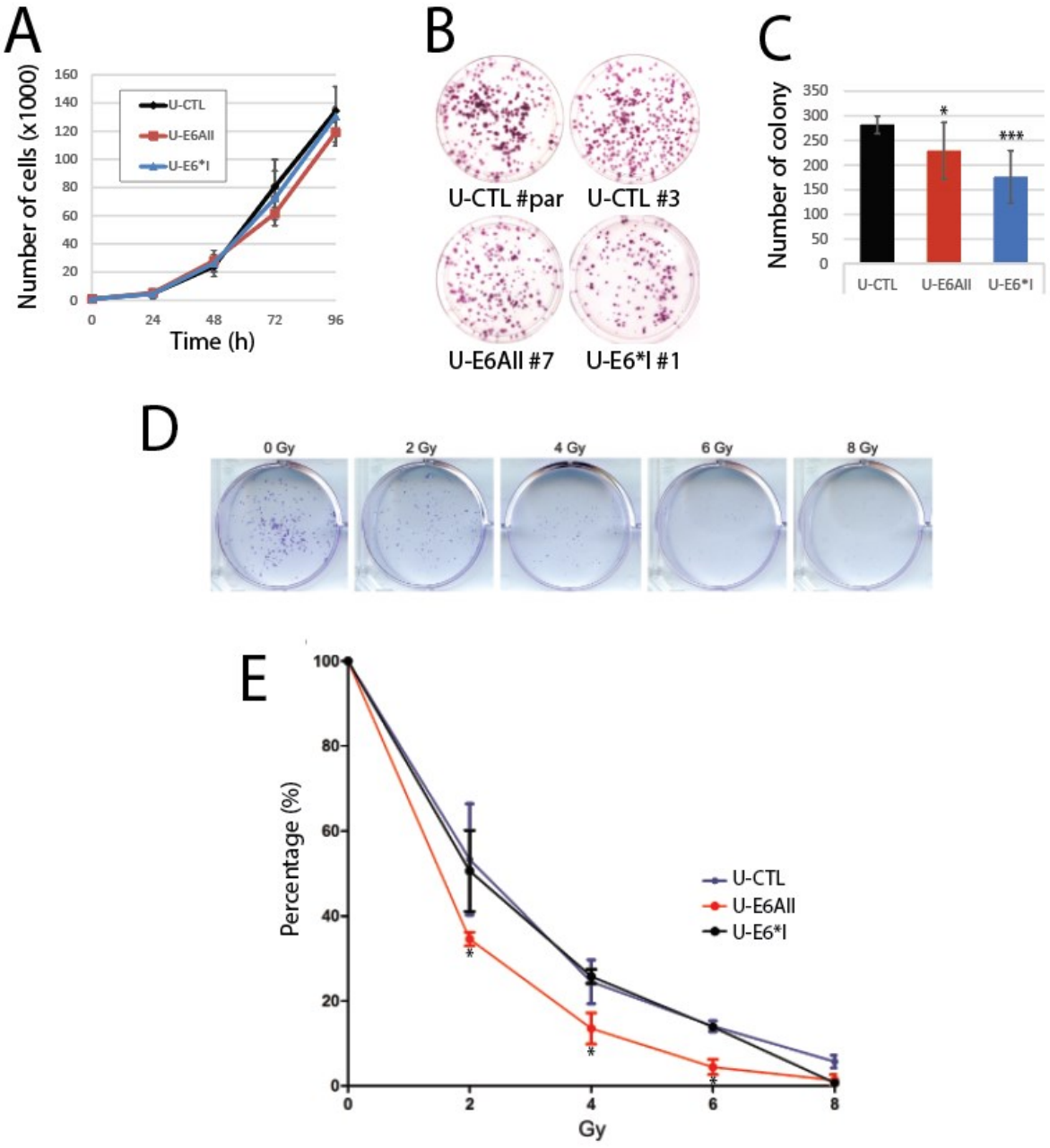

A
Number of cells (x1000)
160
140
120
100
80
60
40
20
0
0
24
48
72
96
Time (h)
U-CTL
U-E6AII
U-E6*I
B
U-CTL #par
U-CTL #3
U-E6AII #7
U-E6*I #1
C
Number of colony
350
300
250
200
150
100
50
0
*
***
U-CTL
U-E6AII
U-E6*I
D
0 Gy
2 Gy
4 Gy
6 Gy
8 Gy
E
Percentage (%)
100
80
60
40
20
0
0
2
4
6
8
Gy
U-CTL
U-E6AII
U-E6*I

Fig 2

708

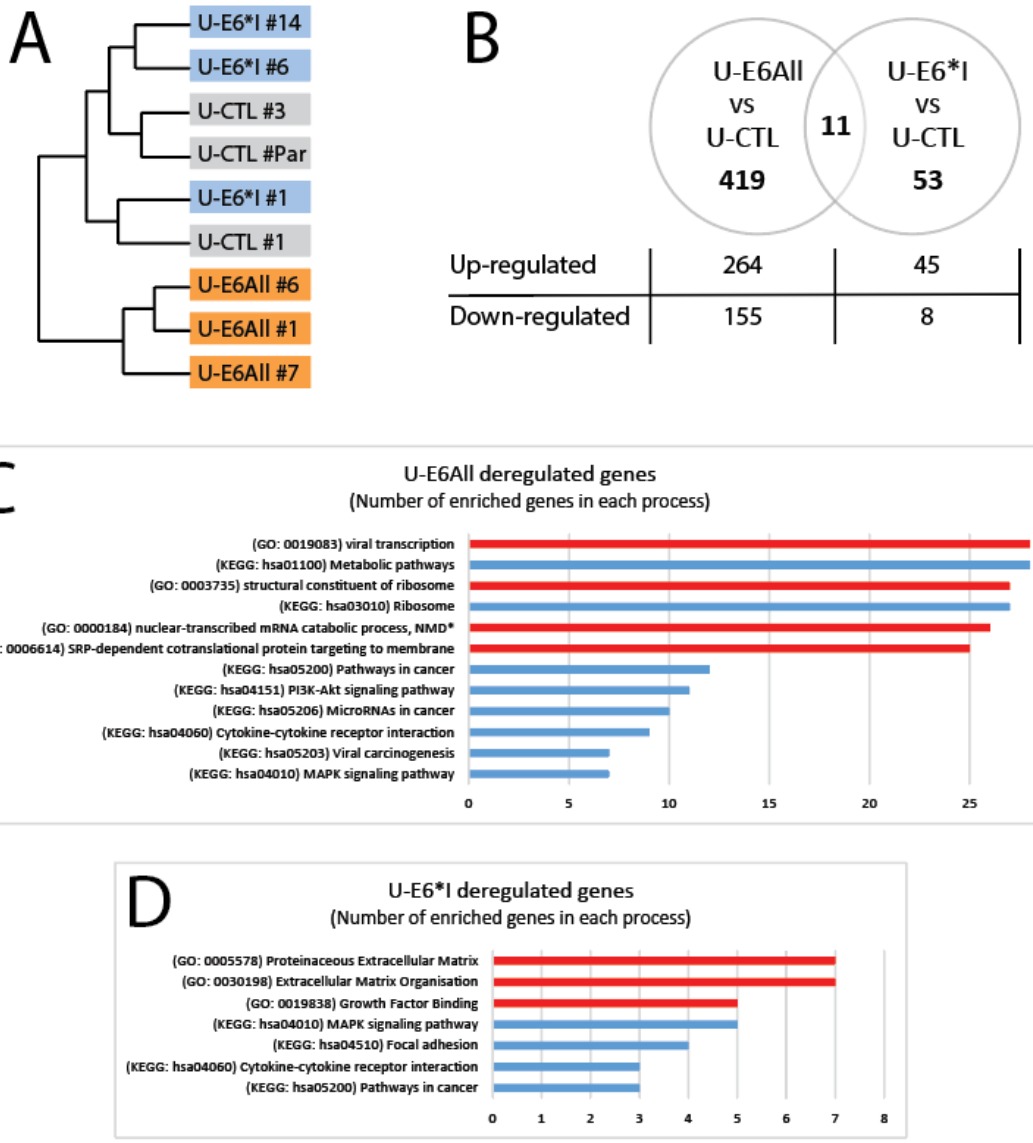

A
U-E6*I #14
U-E6*I #6
U-CTL #3
U-CTL #Par
U-E6*I #1
U-CTL #1
U-E6All #6
U-E6All #1
U-E6All #7
B
U-E6All vs U-CTL
419
11
U-E6*I vs U-CTL
53
Up-regulated 264 45
Down-regulated 155 8
C
U-E6All deregulated genes
(Number of enriched genes in each process)
(GO: 0019083) viral transcription
(KEGG: hsa01100) Metabolic pathways
(GO: 0003735) structural constituent of ribosome
(KEGG: hsa03010) Ribosome
(GO: 0000184) nuclear-transcribed mRNA catabolic process, NMD*
(GO: 0006614) SRP-dependent cotranslational protein targeting to membrane
(KEGG: hsa05200) Pathways in cancer
(KEGG: hsa04151) PI3K-Akt signaling pathway
(KEGG: hsa05206) MicroRNAs in cancer
(KEGG: hsa04060) Cytokine-cytokine receptor interaction
(KEGG: hsa05203) Viral carcinogenesis
(KEGG: hsa04010) MAPK signaling pathway
0 5 10 15 20 25 30
D
U-E6*I deregulated genes
(Number of enriched genes in each process)
(GO: 0005578) Proteinaceous Extracellular Matrix
(GO: 0030198) Extracellular Matrix Organisation
(GO: 0019838) Growth Factor Binding
(KEGG: hsa04010) MAPK signaling pathway
(KEGG: hsa04510) Focal adhesion
(KEGG: hsa04060) Cytokine-cytokine receptor interaction
(KEGG: hsa05200) Pathways in cancer
0 1 2 3 4 5 6 7 8

Fig 3

709

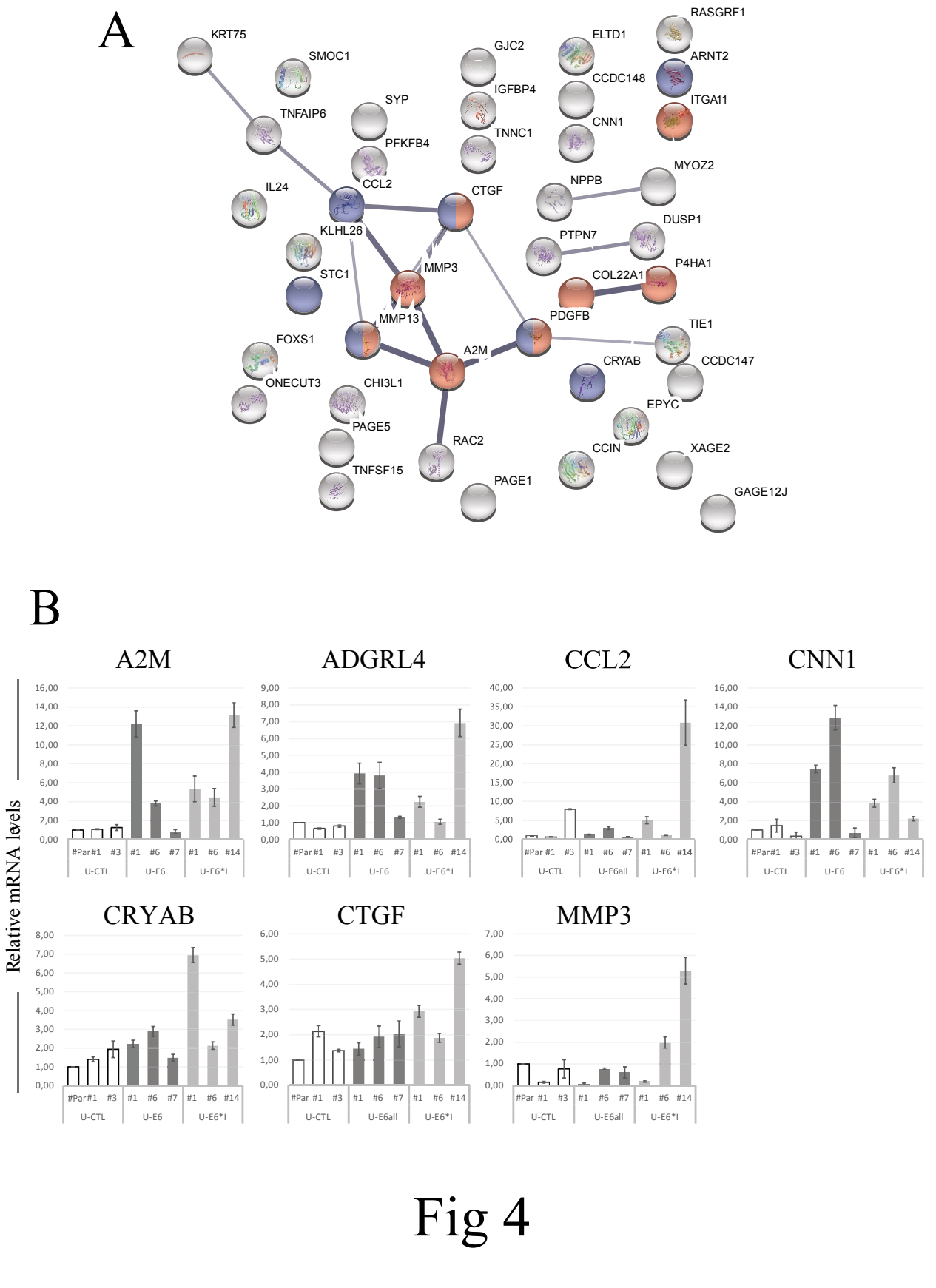

A
KRT75
SMOC1
TNFAIP6
SYP
PFKFB4
CCL2
IL24
KLHL26
STC1
MMP3
MMP13
FOXS1
ONECUT3
CHI3L1
PAGE5
TNFSF15
A2M
RAC2
PAGE1
GJC2
IGFBP4
TNNC1
CTGF
ELTD1
CCDC148
CNN1
NPPB
PTPN7
PDGFB
RASGRF1
ARNT2
ITGA11
MYOZ2
DUSP1
P4HA1
COL22A1
TIE1
CRYAB
CCDC147
EPYC
CCIN
XAGE2
GAGE12J
B
A2M
ADGRL4
CCL2
CNN1
CRYAB
CTGF
MMP3
Relative mRNA levels
U-CTL
U-E6
U-E6*I
U-E6all

Fig 4

(A) Trancripts deregulated in both U-E6all and U-E6*I compared to U-CTL		Log2 Fold-change (RNA-seq)	Pearson Correlation (r) RNA-seq/RT-qPCR
A2M	Alpha-2-macroglobulin	2.376	0.975
ADGLR4	Adhesion G Protein-Coupled Receptor L4	1.752	0.838
CCL2	C-C Motif Chemokine Ligand 2	1.695	0.995
KRT75	Keratin 75	3.064	0.980
PAGE-1	P Antigen Family, Member 1	6.168	0.982
SMOC1	Secreted Modular Calcium-Binding Protein 1	2.612	0.891
TNFAIP6	TNF Alpha Induced Protein 6	3.714	0.933
TNNC1	Troponin C Type 1	1.059	0.982
CH507-210P18.1	lincRNA	3.539	0.747
IGFBP4	Insulin Like Growth Factor Binding Protein 4	-1.272	0.908
PFKFB4	6-Phosphofructo-2-Kinase/Fructose-2,6-Biphosphatase 4	-1.224	0.928
(B) Trancripts deregulated in U-E6*I compared to U-CTL		**Log2 Fold-change (RNA-seq)**	**Pearson Correlation (r) RNA-seq/RT-qPCR**
CFAP58	Cilia and Flagella Associated Protein 58	1.525	0.815
CNN1	Calponin 1	1.974	0.841
CRYAB	CRYstallin Alpha B	1.887	0.874
CTGF	Connective Tissue Growth Factor	0.996	0.951
MMP3	Matrix Metallopeptidase 3	1.682	0.985
PICSAR	p38 inhibited cutaneous squamous cell carcinoma-associate	1.364	0.697
P4HA1	Prolyl 4-Hydroxylase subunit Alpha 1	-1.420	0.977
SYP	SYnaptoPhysin	1.307	0.980

Table 1

711

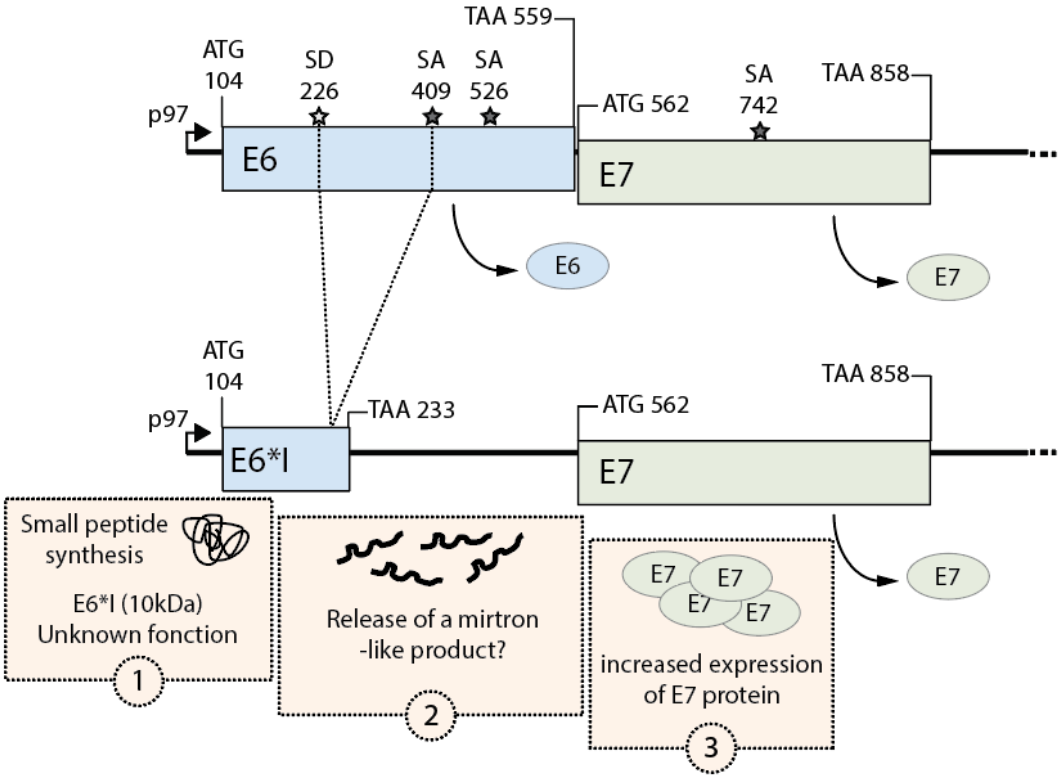
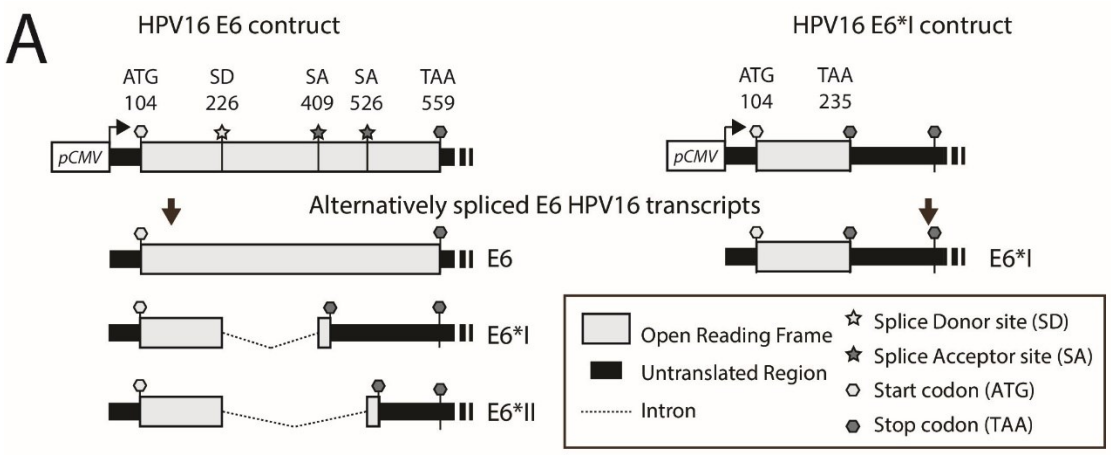

TAA 559
ATG
104
SD
226
SA
409
SA
526
SA
742
TAA 858
p97
ATG 562
E6
E7
E6
E7
ATG
104
TAA 858
p97
TAA 233
ATG 562
E6*I
E7
Small peptide
synthesis
E6*I (10kDa)
Unknown fonction
1
Release of a mirtron
-like product?
2
E7
E7
E7
E7
increased expression
of E7 protein
3
E7

Fig 5

712



B

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

          ACAACTATACATGATATAATATTAGAATGTGTGTACTGCAAGCAACAGTTACTGCGACGT
           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

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

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

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

          AAAAAGCAAAGATTCCATAATATAAGGGGTCGGTGGACCGGTCGATGTATGTCTTGTTGC
           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

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

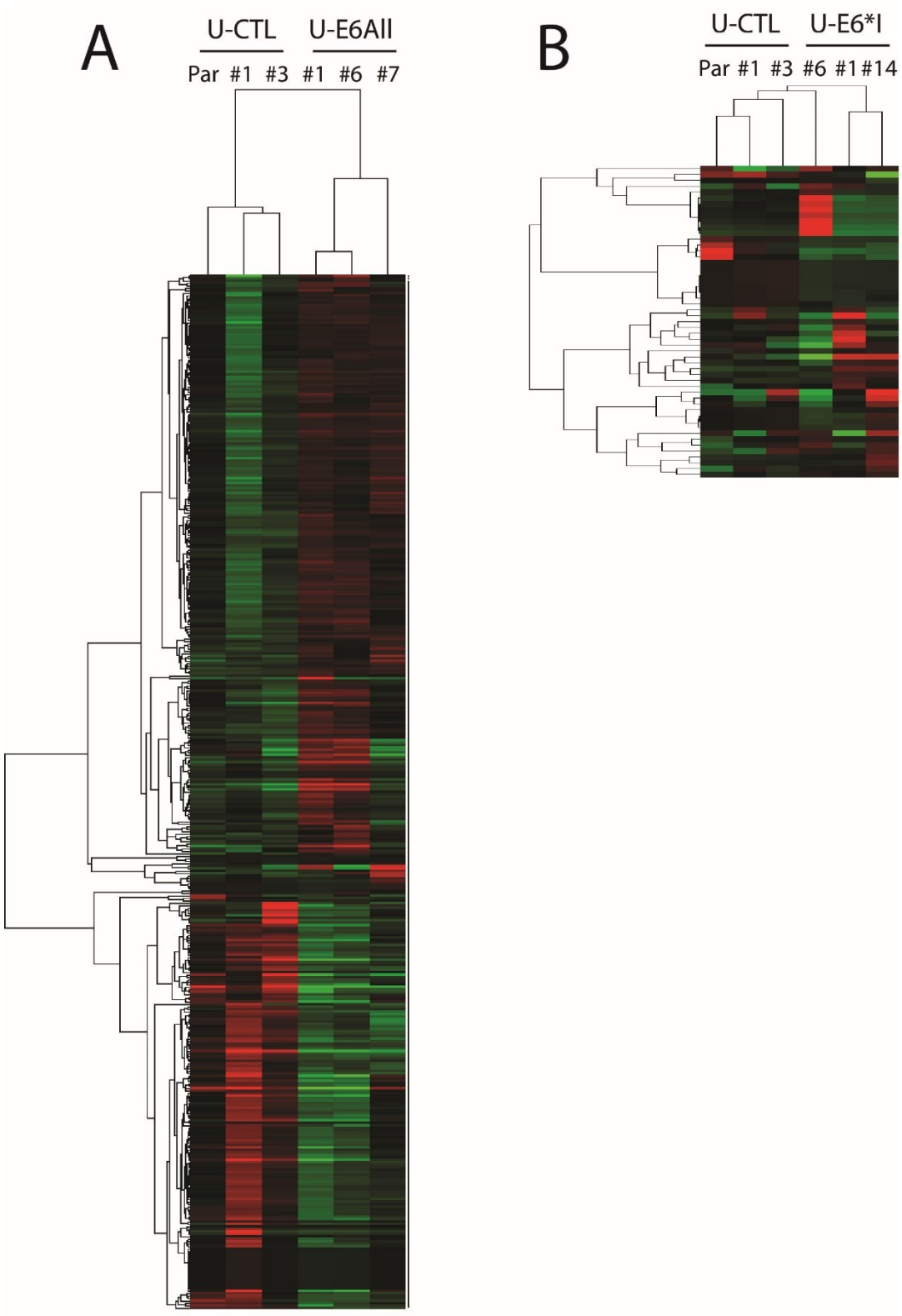
           GAGGTGTATTAActgtcaaaagccactgtgtcctgaagaaaagcaaagacatctggacaa
            E  V  Y  *  43

           aaagcaaagattccataatataaggggtcggtggaccggtcgatgtatgtcttgttgcag


           atcatcaagaacacgtagagaaacccagctgtaa
```

Supp Figure 1

713



Supp Figure 2

715 **Table S3. List of RT-PCR and RTqPCR primers**

Gene id	Target	Sequence (5' --> 3')	Use
ENSG00000175899	A2M_ex19_F	TGCCTTACTCTGTGATTCGTG	qPCR
	A2M_ex20_R	CGCTTGTTCCTTCTCCACTG	
ENSG00000162618	ADGRL4_ex10_F	CTTTGCTGTAGCCTATTTCTTGC	
	ADGRL4_ex11_R	GCCTTCAATGCACATCCATG	
ENSG00000279998	CH507_ex1/2_F	GGATAGAAAATGCAGTTGAAAGATATTTC	
	CH507_ex3_R	CTCTCGCTCCAGGTCCTAG	
ENSG00000141753	IGFBP4_ex2_F	GCACTTCGCCAAAATTCGAG	
	IGFBP4_ex3_R	CTCTGTGAAGCGGCCAG	
ENSG00000170454	KRT75_ex2_F	GCCCCTCTTTGATTCCTATACC	
	KRT75_ex3_R	TGCGCTTGTTAATTTCATCTTCG	
ENSG00000068985	PAGE1_ex4_F	AAGACTGGGTGTGAGCTTG	
	PAGE1_ex5_R	GTGAACCTGTTCCTGGCTATC	
ENSG00000114268	PFKFB4_ex8_F	GTGCTATGAGAACTCCTACGAG	
	PFKFB4_ex9_R	GAGGTAATATACGATGCGGCTC	
ENSG00000198732	SMOC1_ex9_F	TGATGCCCAGTTGTGAGAG	
	SMOC1_ex10_R	CAGTAGGCTGGTGATAAACTCC	
ENSG00000123610	TNFAIP6_ex2_F	AATACAAGCTCACCTACGCAG	
	TNFAIP6_ex3_R	GGTATCCAACTCTGCCCTTAG	
ENSG00000108691	CCL2_ex2_F	TGTCCCAAAGAAGCTGTGATC	
	CCL2_ex3_R	ATTCTTGGGTTGTGGAGTGAG	
ENSG00000114854	TNNC1_ex4_F	TGACCTCTTCCGCATGTTTG	
	TNNC1_ex5_R	TTGTTCTTGTCTCCGTCCTTC	
ENSG00000196074	SYCP2_ex34_F	GTAACAGTGGAGTTGGAGGTAC	
	SYCP2_ex35_R	CTCTTAGTTGGAGTATGTCTGGG	
ENSG00000100697	DICER1_F	GTCTCCCCAGCATACTTTATCG	
	DICER1_R	TTTCCACCCGAAGTCTAAGTTAG	
	E6_F	TGACTTTGCTTTTCGGGATTT	
	E6_R	ACAGCATATGGATTCCCATCTC	
	E6*I_F	CTGCGACGTGAGGTGTATTA	
	E6*I_R	TGTCCAGGTGTCTTTGCTTT	
	E6_orf_F	ATGCACCAAAAGAGAACTGC	PCR
	E6_orf_R	TTACAGCTGGGTTTCTCTACGTGT	
	RPLP0-F	GAAGGCTGTGGTGCTGATGG	
	RPLP0-R	CCGGATATGAGGCAGCAGTT	

716

717 Ex, Exon. F, forward. R, reverse

718

2. Rôle de la protéine E6 dans la liaison aux ARN

2.1. Introduction

Il a été décrit qu'E6 se lie aux acides nucléiques, ADN et ARN (Bodaghi *et al*., 2009; Mallon *et al*., 1987; Ristriani *et al*., 2000). Cette liaison serait liée à la présence de 2 domaines à doigt de zinc au sein de la protéine E6. De façon intéressante, seule la protéine E6 des HPV-HR mais pas celles des HPV-BR, a la capacité de se lier à ses propres ARN *in vitro* (Bodaghi *et al*., 2009). E6, qui régule l'expression de nombreux gènes cellulaires, pourrait intervenir au niveau post-transcriptionnel. D'ailleurs, E6 avec NFX1-123 et PABP4, permet de réguler la stabilité des ARNm de TERT (Katzenellenbogen *et al*., 2007, 2009, 2010). Nous avons voulu investiguer l'ampleur de la liaison d'E6 avec les ARNm au niveau cellulaire. Pour cela, nous avons effectué une RNA-IP (RNA-immunoprecipitation) anti-E6 et procédé à l'analyse des ARNm cellulaires et viraux co-précipités avec E6. En outre, une analyse des transcrits co-précipités avec E6, par RNA-seq, nous permettrait d'identifier des ARNm cibles d'E6 au niveau post-transcriptionnel.

2.2. Méthodologie

- Cellules et plasmides

Les cellules U2OS (HPV négatives) ont été cultivées dans un milieu DMEM (Dulbecco's Modified Eagle Medium) supplémenté de 10% de Sérum de veau fœtal, à 37°C dans une atmosphère humide contenant 5% de CO2. Les plasmides pXJ40-vide et pXJ40-E6 ont été généreusement offert par le Dr Murielle Masson (ESBS, Strasbourg). Le plasmide pXJ40-E6mut^{SD226}, qui contient une mutation du site d'épissage SD226 et exprimant uniquement l'isoforme E6Fl, a été généré par mutagénèse dirigée à partir du pXJ40-E6, en utilisant les amorces sens E6mutSD226-F (GCTATATGAC-TTTGCTTTTCGGG) et antisens E6mutSD226-R (TCGCGTCGCAGTAACTGTTGCTTG). Le plasmide pXJ40-E6mut^{SD226}-Flag a été généré à partir du pXJ40-E6mut^{SD226}en utilisant les amorces sens E6-Flag-F (GATGATGATAAATAAGCGGCCGCCCCGGGC) et antisens E6-Flag-F (ATCTTTATAA-TCCAGCYGGGTTTCTCTACGTGTTCTTGATGATCTGCAA-C). Les mutagenèses ont été réalisées à l'aide du kit Q5 Site-Directed Mutagenesis (New England Biolabs) selon les recommandations du fournisseur.

- Immunoprécipitation de l'ARN

Des cellules U2OS cultivées dans des boites 10cm, à 60-75% de confluence, ont été transfectées avec les plasmides pXJ40-vide, -E6mut^{SD226} ou -E6mut^{SD226}-Flag en utilisant le jetPEI (Polyplus Transfection) suivant les recommandations du fabricant. Brièvement, 30 µl de jetPEI ont été utilisés pour transfecter 15 µg de plasmides dans chaque boite de cellules. Après 24h de transfection, les cellules ont été lavées avec du PBS 1x froid, récoltées à l'aide d'un scraper avant d'être culottées par une centrifugation pendant 10 min à 1500 g à 4°C. La lyse cellulaire a été réalisée par l'ajout du tampon NET-2 (50 mM Tris-HCl pH7.4, 150 mM NaCl, 10 mM MgOAc, 0.05% (v/v) NP-40, 1 mM DTT, inhibiteurs de protéase (Roche Diagnostics), 80 U/mL RNAse OUT (Invitrogen)), une sonication 5 x 1sec, puis une centrifugation pendant 10 min à 16100 g à 4°C. Les concentrations protéiques ont été déterminées à l'aide du réactif Bio-Rad protein assay (Bio Rad).

Pour l'étape de *pre-clearing*, 1 mg/1 ml de protéines de chaque échantillon a été incubé en présence de 40 µl de billes de protéines A/G (Thermo Scientific) à 4°C pendant 30min sous agitation, puis centrifugé à 500 g pendant 5 min à 4°C. Les surnageants ont été divisés en fractions de 500 mg et incubées avec 0,4 µg d'anticorps pendant 2h à 4°C sous agitation. Des anticorps de lapin anti-Flag (Sigma) et de souris anti-E6 (1E-6F4) (Euromedex) ont été utilisés, en parallèle d'isotope IgG contrôle correspondant à chaque espèce (Santa Cruz), pour l'immunoprécipitation des protéines E6 et E6-Flag respectivement. Chaque mélange a été ensuite additionné de 40 µl de billes de protéines A/G et 1 mg/ml d'ARNt de levure (Roche Diagnostics), puis incubé à 4°C sous agitation pendant la nuit. Deux étapes de lavage ont été réalisées en ajoutant un tampon NET-2 contenant 300 mM de NaCl (1x), suivies d'une centrifugation à 500 g pendant 5 min 4°C, puis d'une incubation pendant 10 min à 4°C sous agitation avec un tampon NET-2 contenant 300 mM NaCl et 1 mg/ml d'héparine, suivie d'une centrifugation à 500 g pendant 5 min 4°C. Après 8 autres lavages réalisés par l'ajout du tampon NET-2 et une centrifugation à 500 g pendant 5 min à 4°C, les complexes anticorps-protéine ont été récupérés par l'ajout de 50 µl de tampon d'élution (0.1 M Tris-HCl pH 6.8, 20% (v/v) glycerol, 4% (m/v) SDS, 12% (v/v) β-mercaptoéthanol) suivie d'une centrifugation à 500 g pendant 5 min à 4°C. Chaque éluat a été divisé en 2 fractions : une fraction analysée par western blot et une autre utilisée pour l'extraction d'ARN et l'analyse par RT-PCR.

- Western blot

L'extraction des protéines totales a été réalisée en utilisant un tampon RIPA (50 mM Tris-HCl pH7.4, 150 mM NaCl, 1% NP40, 0.5% Na deoxycholate, 1mM EDTA) additionné de 30 µg/mL d'anti-protéases. Après sonication et centrifugation à 10 000 g pendant 10 min, les surnageants protéiques ont été dosés en utilisant le kit Bio-Rad Protein Assay (Bio-Rad) selon les recommandations du fournisseur. Les protéines ont été séparées par électrophorèse SDS-PAGE puis transférées sur une membrane PVDF Amersham^{TM}Hybond^{TM} ((GE Healthcare). Après la saturation par incubation avec du lait 5% pendant une nuit à 4°C, les membranes ont été incubées avec l'anticorps primaire puis avec l'anticorps secondaire conjugué à la HRP, pendant 2h et 1h à température ambiante respectivement. La chimiluminescence émise par les complexes immuns a été révélée à l'aide du substrat PierceECL 2 Western Blotting (Thermo Scientific).

Les anticorps primaires de souris anti-E6 (1E-6F4) et anti-Flag (FG4R) (Thermo Scientific), ont été utilisés pour révéler les protéines E6 et E6-Flag respectivement. Un anticorps secondaire de chèvre anti-souris (BD Pharmingen) a été utilisé pour révéler les deux protéines.

- Extraction d'ARN et RT-PCR

Les ARN totaux/immunprécipités ont été extraits en utilisant le réactif Ribozol (VWR chemicals) suivant les recommandations du fournisseur. La rétro-transcription des ARN a été réalisée en utilisant le kit Maxima first strand cDNA synthesis (ThermoScientific) suivant les recommandations du fabricant. Les cDNA obtenus ont été amplifiés par PCR semi-quantitative à l'aide du kit DreamTaq polymerase (ThermoScientific) suivant les recommandations du fabricant. Les amorces utilisées sont : E6-ORF-F (ATGCACCAAAAGAGAACTGC) et E6-ORF-R (TTACAGCTGGGTTTCTCTACGTGT). Les produits de PCR ont été analysés par électrophorèse sur gel d'agarose 2% (m/v) préparé dans un tampon Tris-Borate-EDTA. Pour E6ORF, le programme de PCR a consisté en : (i) une dénaturation initiale à 95°C pendant 2 min, (ii) 35 cycles d'amplification à 95°C pendant 30 sec, 50°C pendant 20 sec et 72°C pendant 30 sec suivis (iii) d'une élongation finale à 72°C pendant 5 min. Les produits de PCR ont été analysés par une électrophorèse sur un gel d'agarose 2% (m/v) dans un tampon TBE (Tris Borate EDTA). Les produits ont été visualisés aux ultraviolets à l'aide du système ChemiDoc XRS (BioRad).

2.3. Résultats

Afin de confirmer l'interaction entre la protéine E6 et ses propres ARNm, nous avons réalisé des expériences d'immunoprécipitation RNA-IP à partir d'extraits de cellules U2OS transfectées avec les plasmides pXJ40-Vide, -E6^{mutSD226} (plasmide codant la protéine E6Fl) ou -E6^{mutSD226}-Flag (plasmide codant la protéine E6Fl contenant une étiquette Flag). Comme illustré dans la figure 27, l'immunoprécipitation est spécifique des protéines E6 et E6-Flag avec les anticorps anti-E6 et anti-Flag respectivement. De façon intéressante les ARNm E6 ont co-précipité avec la protéine E6, dans les deux conditions d'IP, indiquant l'existence d'une interaction entre E6 et ses ARNm.

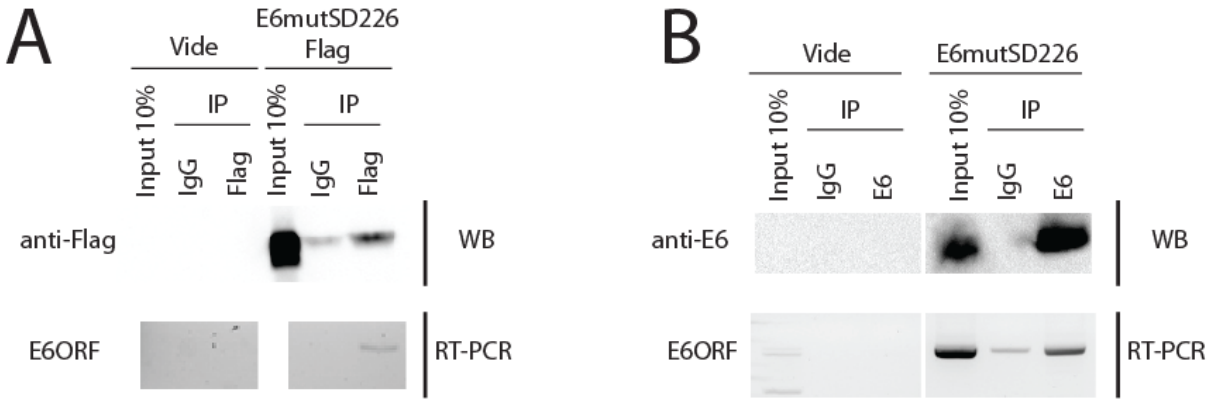


Figure 27. Interaction entre la protéine E6 avec ses ARNm. (A) Une immunoprécipitation de l'ARN (RIP) à partir d'extraits de cellules U2OS transfectées pendant 24h avec des plasmides pXJ40 vide ou pXJ40-E6^{mutSD226}-Flag a été réalisée avec des anticorps IgG ou anti.-Flag. L'IP d'E6 a été analysée par western blot et les ARN co-précipités révélés par RT-PCR et analysés sur d'agarose. pXJ40-E6^{mutSD226}-Flag, plasmide qui exprime exclusivement la protéine E6Fl-Flag. (B) RNA-IP réalisée à partir d'extrait de cellules U2OS transfectées pendant 24h avec des plasmides pXJ40 -vide ou -E6mutSD226, en utilisant des anticorps de souris anti-E6 ou contrôle (IgG non spécifiques). L'IP d'E6 a été analysée comme indique dans (A).

2.4. Conclusion

Nous avons pu confirmer la liaison d'E6 *in cellulo* avec ses propres ARNm. Toutefois, nous n'avons pas appréhendé la nature de cette liaison, notamment si celle-ci était directe ou dépendante d'un autre facteur cellulaire. Il serait intéressant d'étudier la liaison d'E6 avec les ARNm dans un contexte plus « physiologique », en effectuant une RNA-IP avec une protéine E6 endogène. Ces résultats s'avèrent prometteurs en vue de l'appréhension des activités d'E6 des HPV-HR. Dans ce sens, il serait aussi intéressant d'étudier la capacité des protéines E6 à se lier aux ARN viraux mais aussi cellulaires, chez d'autres HPV-HR, mais aussi des HPV-BR. Cela étant, des expériences menées *in vitro* ont montré qu'E6 d'HPV11 (HPV-BR) n'était pas capable de se lier aux ARN. Quel serait donc le rôle de cette propriété de RBP dans les fonctions et effets d'E6 au niveau cellulaire ? Quel serait son rôle dans le cycle viral et dans la carcinogenèse associée aux HPV-HR ?

3. Implication de l'EJC et du NMD dans la régulation de l'expression des gènes E6 et E7 d'HPV16

3.1. Résumé de la publication 2

La progression de la carcinogenèse induite par les HPV-HR est corrélée avec une augmentation de l'expression des protéines E6 et E7. Appréhender les mécanismes de régulation de l'expression d'E6 et E7 permettrait de mieux connaitre les processus sous-jacents de la carcinogenèse induite par les HPV-HR. L'expression des gènes viraux, réalisée par la machinerie de la cellule hôte, est finement régulée particulièrement au niveau posttranscriptionnel. A ce niveau, le devenir des ARNm est régulé par la fixation de protéines RBP, formant des mRNP. Les ARNm des HPV-HR sont soumis à un épissage alternatif permettant de générer une vingtaine de transcrits viraux, mais aussi de remodeler la composition des mRNP, en permettant le recrutemement de l'EJC. L'EJC est un complexe multiprotéique, composé d'eIF4A3, Y14, MAGOH et MLN51, qui est déposé sur les ARNm *via* l'épissage et qui y reste fixé jusqu'à leur traduction. De fait, l'EJC influence les nombreuses réactions auxquelles sont soumis les ARNm : l'export nucléocytoplasmique, la traduction, l'épissage et la dégradation par le NMD. En outre, le NMD est un mécanisme de contrôle de qualité des ARNm qui assure la dégradation des transcrits contenant un PTC. De façon intéressante, les codons stop retrouvés dans les ARNm polycistroniques, comme ceux d'HPV16, peuvent être assimilié à des PTC.

Afin de caractériser les mécanismes impliqués dans la régulation de l'expression des ARNm des HPV-HR, nous avons étudié l'impact de l'EJC et du NMD dans la régulation de l'expression des transcrits d'HPV16. Nous avons observé que l'inhibition des facteurs de l'EJC, eIF4A3 et Y14, dans des cellules HPV16+ (SiHa, Ca Ski) induit une augmentation de l'expression des transcrits et des protéines E6 et E7, tandis que l'inhibition de MLN51 a montré un effet inverse. En outre, alors que l'EJC n'affecte pas l'épissage des ARNm d'HPV16, eIF4A3 permet leur déstabilisation dans les cellules Ca Ski. De plus, les ARNm viraux interagissent avec eIF4A3, lequel permet la fixation de l'EJC sur les ARN. eIF4A3 pourrait être recruté sur les mRNP d'HPV16 d'une façon dépendante mais aussi indépendante de l'épissage.

Ensuite, nous avons étudié l'impact du NMD sur les transcrits d'HPV16. Pour cela, nous avons inhibé le NMD et étudié l'expression des gènes viraux. Nous avons observé que l'inhibition du NMD par le CHX (Cycloheximide) et le 5-azaC (5-azacytidine) induit une

augmentation des transcrits E6 et E7 d'HPV16, dans les cellules SiHa mais pas dans les Ca Ski. Le traitement par le 5-azaC augmente le niveau des protéines E7 et E6 à un moindre niveau. De plus, l'inhibition du NMD *via* le ciblage des facteurs SMG-1 par des siRNA a également augmenté l'expression d'E6 et E7. L'inhibition d'UPF1 induit une plus faible augmentation, mais non significative, des mêmes transcrits. Ces résultats suggèrent que les transcrits d'HPV16 soient ciblés par le NMD dans les cellules SiHa et qu'ils puissent y échapper dans les cellules Ca Ski. De façon intéressante, l'efficacité du NMD est réduite dans les cellules HPV16+ (SiHa et Ca Ski) par rapport aux cellules HPV- (C33A).Le NMD est inhibé plus fortement dans les cellules Ca Ski. Par conséquent, l'expression de transcrits cellulaires régulés par le NMD a été dérégulée. Nous avons montré que l'expression ectopique d'E6 réduit l'activité du NMD, suggérant que HPV16 inhibe le NMD. En somme, l'activité du NMD (lequel peut inhiber l'expression des ARNm viraux) est réduite par HPV16, permettant à celui-ci d'avoir un contrôle sur ses transcrits mais aussi sur les cibles cellulaires du NMD. Il serait intéressant de connaitre l'implication de cette inhibition mutuelle, entre le NMD et HPV16, dans le cycle viral et la carcinogenèse induite par les HPV-HR.

The Exon Junction Complex (EJC) and the NMD modulate HPV16 E6 and E7 oncoproteins expression

Koceila Meznad [1,2], Jérôme Perrard [1,2], Philippe Paget-Bailly[1,2], Anne Peigney[1,2], François Aubin [1,2,3], Christiane Mougin[1,2,3], Jean-Luc Prétet [1,2,3], Aurélie Baguet[1,2*]

[1] EA3181, LabEx LipSTIC ANR-11-LABX-0021, UFR SMP, 19 rue Ambroise Paré, Besançon.
[2] Université de Bourgogne Franche Comté, France.
[3] Centre Hospitalier Régional Universitaire, Bvd Alexandre Fleming, Besançon, France.

Running title: The EJC and the NMD modulates HPV16 E6 and E7 oncoproteins expression

[*]Corresponding author
Aurélie Baguet, UFR SMP, 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

Abstract: 228 words. Text: 4986 words.

Abstract

High-risk human papillomavirus (hrHPV) infection is the major etiologic factor for the development of ano-genital cancers and some head and neck squamous cell carcinoma (HNSCC). hrHPVs-induced carcinogenesis, including HPV16, is driven by an increased and continuous expression of viral oncoproteins, E6 and E7, whose expression is mainly regulated at post-transcriptional level. Post-transcriptionnal regulation involved RNA binding proteins that regulate mRNA fate. RNA splicing provides for the production of several polycistronic viral transcripts, but could also allow for Exon Junction Complex (EJC) deposition. The EJC is involved mRNA expression regulation by several process, among them the Nonsense-mediated mRNA decay (NMD). NMD is an mRNA quality control that ensures for degradation of aberrant mRNAs, but also for cellular and viral RNAs. To get insights into HPV-driven carcinogenesis, we investigated the implication of EJC and NMD in the regulation of HPV16's gene expression. We detected an interaction between eIF4A3, an EJC coupound, and HPV16 transcript. Then, we showed that EJC-depletion upregulated both E6 and E7 protein expression. In addition, NMD inhibition increased E6 and E7 expression, in SiHa cells not in Ca Ski cells. Interestingly, we observed an inhibition of NMD in HPV16+ cell lines, particularly Ca Ski cells, which altered cellular NMD-targets expression. Our results show for the first time that EJC and NMD can modulate HPV16 transcripts expression. Indeed, NMD is inhibited by HPV16, thus deregulating cellular NMD-naturel targets

Importance

Recently, it has been reported that Nonsense-mediated mRNA decay targets to degradation genomic and/or messenger viral RNAs from RNA viruses. Our study describe for the first time that mRNAs from DNA virus (HPV16) could be sensitive to NMD. We also show that HPV16, as it has been observe for others viruses, altered NMD pathway to protect their own RNAs but also to modulate host cell gene expression, thus questioning about the role of the NMD-inhibition on the cellular gene expression and on HPV-induced carcinogenesis.

Introduction

Human papilloma viruses (HPV) are small non-enveloped viruses with circular double-stranded DNA genome. More than 200 types of HPVs infect humans (Schiffman et al., 2016) and are classified in two categories: High-risk carcinogenic HPV (HR-HPV) (including HPV16 and HPV18), which are the etiological agents of cervical and anal carcinoma and frequently found in head and neck squamous cell carcinoma (HNSCC), and low-risk carcinogenic HPV (LR-HPV) (including HPV6 and HPV11), which are associated with begnin lesions such as genital warts (Guan et al., 2012). In most cases, HPV infections are efficiently cleared by the immune system within few months. But, persistence of HPV infection is the main cause to develop a High-grade squamous intra-epithelial lesions (HSILs) corresponding to an abortive HPV infection associated with deregulation of viral genes expression and a higher risk of progression to cancer (Shulzhenko et al., 2014). 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 early proteins whereas late proteins are produced from the late promoter p670. HR-HPV -driven carcinogenesis is most often related to the E2 ORF disruption as a consequence of the viral integration into the host genome (Shulzhenko et al., 2014). The loss of the E2 protein leads to increase p97-mediated transcription, which consequently produce E6 and E7 proteins. Both HPV16 E6 and E7 proteins have pleiotropic effect by interacting with multiple host cell proteins, such as, respectively, p53 and pRb tumor suppressor proteins (Roman and Munger, 2013; Vande Pol and Klingelhutz, 2013). Continuous expression of HPV16 E6 and E7 proteins are essential for the initiation and maintenance of cancer.

In addition to transcriptional regulation, HPV16 RNA processing, mediated by the interaction of RNA-binding proteins (RBPs), largely contribute to the regulation of viral protein expression (Graham and Faizo, 2017). For instance, HPV utilize alternative splicing mediated by cellular splicing factors to diversify viral proteins-encoding mRNAs. Indeed, p97-derived HPV16 transcripts can be submitted to alternative splicing to generate almost 20

polycistronics mRNAs encoding early proteins (Chen et al., 2014; Schwartz, 2013). Four cryptic splice sites have been identified in E6 and E7 ORFs including one splice donor site SD226 and three splice acceptor sites, SA409, SA526 and SA742, leading to the expression of truncated proteins, respectively named E6*I, E6*II and E6^E7.

The EJC is a multiprotein complex assembled onto pre-messenger RNA (pre-mRNA) during splicing. The EJC is loaded at 24 nucleotides upstream the exon-exon junctions, in a sequence independent manner. The core EJC is composed of four proteins, eIF4A3 (DDX48), Y14 (RBM8A), Magoh and MLN51 (CASC3), which act as a binding platform to recruit others peripheral EJC factors (Le Hir et al., 2016). Consequently, the composition of EJC evolves all along the nuclear and cytoplasmic mRNA lifecycle and ultimately, EJC factors are removed by the translation machinery during the first round of translation. The EJC greatly influence the mRNA outcome by affecting the RNA processing. The EJC i) influences the splicing process (Ashton-Beaucage et al., 2010; Fukumura et al., 2016; Michelle et al., 2012; Wang et al., 2014) and the mRNA transport (Giorgi and Moore, 2007; Martin and Ephrussi, 2009); ii) promotes the nuclear export (Le Hir et al., 2001) and the translation (Chazal et al., 2013; Ma et al., 2008; Nott et al., 2004); iii) have a crucial role in mRNA surveillance pathway (Palacios et al., 2004). The EJC allows the recognition of Premature Termination Codon (PTC)-harboring transcripts, and targets their degradation by the Nonsense mediated mRNA Decay (NMD) (Le Hir et al., 2016; Lejeune, 2017).

In this study, we addressed the role of the Exon Junction Complex (EJC) in the HPV16 E6 and E7 oncoproteins expression. We showed that EJC-depletion in Ca Ski and SiHa (HPV16+) cell lines modulate the expression of both E6 and E7 proteins. We observed that eIF4A3 depletion causes a stabilization of early HPV16 RNAs without modulating their alternative splicing. Surprisingly, MLN51 depletion decreased E6 and E7 expression. Also, we reported that eIF4A3 are in the same ribonucleoprotein particles that early HPV16 RNAs. Interestingly, NMD inhibition in SiHa but not Ca Ski cells increased E6 and E7 expression in RNA and protein levels. These data show for the first time that viral transcripts from DNA virus could be NMD substrates at least in SiHa Cells. In addition, we showed in an attractive

way that NMD is inhibited in HPV16+ cell lines, probably through E6 protein activity. Such activity could be used by the virus to control its own mRNA levels, but also the many NMD-naturel targets involved in cellular process.

Results

Depletion of EJC core factors alters the expression of HPV16 E6 and E7 oncoproteins

The expression of HPV16 E6/E7 early transcripts are regulated by alternative splicing generating various transcripts by recognition of 1 donor splice site (SD226) and 3 acceptor splice sites (SA409, SA526 and SA742). The recognition of SD226/SA409 or SD226/SA526 provides spliced transcripts encoding E7 protein and truncated E6 proteins, E6*I and E6*II, respectively. A fusion protein called E6^E7 is potentially encoded by the spliced transcript SD226/SA742. Finally, the oncogenic E6 proteins are exclusively encoded by unspliced transcripts. Interestingly, spliced transcripts potentially encoding E7 and E6*I proteins are mainly expressed in SiHa and Ca Ski cells whereas spliced transcripts encoding E6^E7 fusion protein are rarely observed (Fig. S1).

To investigate whether EJC core factors affect HPV16 gene expression, HPV16 RNA levels were quantified by RT-qPCR in HPV16-positive cell lines, SiHa and Ca Ski depleted for eIF4A3, Y14 and MLN51. Primer sets have been designed to monitor all early transcripts (E6All and E7), unspliced (E6Fl) or spliced transcripts (E6*I and E6^E7) (Fig. 1A). Consistent with protein and RNA levels, around 60-80% of eIF4A3, Y14 and MLN51 depletion was achieved in each cell line (Figs. 1B and 1D). Interestingly, siRNA-mediated depletion of eIF4A3 or Y14 induces a 1.5 at 2.5-fold increase of spliced and unspliced HPV16 early transcripts, E6All, E6Fl, E6*I, E7 and E6^E7 (Fig 1C). Consecutively, higher E6 and E7 protein levels are observed in eIF4A3 and Y14 depleted cells compared to control (Fig. 1D). In contrast, MLN51 depletion leads to 2-fold decrease of E6/E7 HPV16 transcripts and lower E6 and E7 protein levels (Figs. 1C and 1D). These data suggest that EJC core factors alter the

E6/E7 protein expression but, surprisingly, MLN51 has a distinct affact than Y14 and eIF4A3. So, we next investigated how EJC core factors modulate HPV16 gene expression.

EJC core factors do not regulate alternative splicing pattern of HPV16 early transcripts

Recently, eIF4A3 has been involved in the splicing regulation of some specific pre-mRNAs in both Drosophila (Ashton-Beaucage et al., 2010; Roignant et al., 2010) and mammalian cells (Wang et al., 2014). To access whether EJC core factors depletion can modulate the splicing pattern of E6/E7 transcripts, specific primers sets was developed for RT-PCR to simultaneously or individually detect HPV16 E6/E7 alternative transcripts (Fig. 1A). RT-PCR using primer targeting early HPV16 transcripts (F6-R7) are represented in Figure 2A and 2C. As expected, eIF4A3 depletion causes an increase in exon skipping of the constitutive exon 11 of KPNA1, used as positive control (Wang et al., 2014). Although more modest, Y14 depletion also induces exclusion of exon 11 in KPNA1 contrary to MLN51 depletion which have no effect on KPNA1 splicing pattern (Figs. 2A and 2C). However, the reduction of eIF4A3, Y14 or MLN51 does not alter the splicing profile of HPV16 early transcripts (Figs. 2A and 2C). Similar results using real-time PCR were obtained by calculating the ratio of E6Fl/E6All (Figs. 2B and 2D), ruling out the hypothesis that the increase of E6/E7 expression in EJC depleted cells provides from a modulation of alternative splicing pattern of HPV16 early transcripts.

The eIF4A3 depletion causes a stabilization of HPV16 E6/E7 transcripts

We examined whether eIF4A3 or Y14 depletion impairs RNA stability of HPV16 early transcripts. RT-qPCR showing RNA levels during a time course of actinomycin D treatment in Ca Ski and SiHa cells are represented in Figure 3. For all HPV16 transcripts tested (E6All, E6*I and E7) a significantly (from 1.5 to 2.5-fold) increase RNA half-live is observed in eIF4A3-depleted Ca Ski cells (Fig. 3A, right panel and Fig. 3B). However, neither eIF4A3 nor Y14 depletion in SiHa cells did impaired RNA half-live of HPV16 early transcripts (Figs. 3A, left panel and Fig. S2). Altogether, these data suggest the implication of eIF4A3 protein in the regulation of RNA stability of HPV16 early transcripts in Ca Ski but not in SiHa cells.

As Y14 or MLN51 have not displayed similar effect on E6 and E7 expression, we hypothesized that eIF4A3 may influence E6E7 mRNA stability through noncoding RNAs.

Recently, others and we showed that miR-375 targets HPV16 transcripts to decrease E6 protein level (Jung et al., 2014; Morel et al., 2017). To test whether eIF4A3 affects miR-375 expression, RT-qPCR was performed in EJC core factors depleted cells (Fig. S3). No difference of miR-375 levels was observed indicating that the stabilization of E6/E7 transcripts, which occurs in eIF4A3-depleted cells, is independent of miR-375.

EIF4A3 associates with ribonucleoparticles containing HPV16 early transcripts

To study the potential interaction between eIF4A3 protein and HPV16 transcripts, RNA-IPs were performed in both SiHa and Ca Ski cell lines. Total RNA extracts were incubated with eIF4A3 (4A3-IP) or control (IgG-IP) antibodies, and the immunoprecipitated complexes were analyzed by western blotting for proteins and semiquantitative RT-PCR for RNAs (Fig. 4). The eIF4A3 immunoprecipitation was monitored by western blotting and shows, as previously reported, that eIF4A3 interacts with MLN51 (Fig. 4A). Then, immunoprecipitated RNAs were analyzed by RT-PCR. As expected, SF3B5 intronless RNA (negative control) does not precipitates whereas KPNA1 RNA (positive control) binds eIF4A3 (Fig. 4B). Altogether, these data are relevant for the RNA-IP validation. Interestingly, in both cell lines, the E6*I spliced RNA is significantly enriched in 4A3-IP compared to IgG-IP. While the detection of unspliced E6 RNA is less obvious in Ca Ski cells, a significant enrichment is observed in SiHa cells. Nevertheless, using primers targeting E7 ORF, we confirmed that eIF4A3 proteins are in ribonucleoprotein particles containing E6/E7 transcripts. These data are in agreement with those of Martinez-Salazar who identified eIF4A3 among several proteins that bind E6 ORF (Martínez-Salazar et al., 2014). Moreover, inhibition of splicing in their experiment suggests that eIF4A3 bind E6 mRNA in a splicing-independent manner.

So, in order to identify which part of the E6 sequence is important for the association of HPV16 transcripts with eIF4A3 protein and also to determine whether this association could

be splicing-independent, RNA-IPs were realized in U2OS cells transfected with a vector encoding E6 ORF which can be spliced, an E6 mutated ORF and E6*I ORF which cannot be spliced (Fig. 5A). We noticed that no interaction between eIF4A3 and E6 proteins was detected by western blotting (Fig. 5B). Both spliced and unspliced E6/E6*I ORF were enriched in 4A3-IP bound fraction compared to IgG-IP whereas SF3B5 RNA negative control was not enriched (Fig. 5C). Therefore, eIF4A3 protein does bind to E6 ORF independently of splicing. Moreover, enrichment of E6*I ORF in 4A3-IP fraction indicated that nucleotides from 226 to 409 are not relevant for its association with eIF4A3 (Fig. 5C). So, these results indicate that eIF4A3 binds E6 sequence in a splicing-independent manner.

HPV16 E6 and E7 gene expression is regulated by NMD

The EJC have a key role in post-transcriptional regulation of mRNA including the elimination of abnormal transcripts by NMD pathway. HPV16 transcripts harbor numerous NMD-eliciting features such as numerous ORFs submitted to alternative splicing. In order to determine whether NMD pathway affects the expression of HPV16 transcripts, we explored the impact of drug-induced NMD inhibition on HPV16 E6/E7 RNA levels. Because NMD requires translation, SiHa and Ca Ski cells were treated with cycloheximide (CHX), a protein synthesis inhibitor. As expected, we observed in CHX-treated cells an increase of SC35 RNA level considered as physiological NMD substrate (Fig. 6A). In addition, CHX treatment induces an upregulation of E6 and E7 RNA levels in SiHa but not in Ca Ski cells (Fig. 6A). Using a RNA FISH assay *(RNAscope Technology)* designed to detect the HPV16 E6/E7 or c-jun RNAs (used as positive control, Sherman et al., 1990), we observed the presence of few cytoplasmic dots for corresponding E6/E7 RNAs (from 2 to 4 dots per cell) in DMSO-treated cells (Fig. 6B). However, in CHX-treated cells, the number of dots corresponding E6/E7 RNAs, was significantly increased (from 5 to 7 dots per cell) (Figs. 6B and 6C). The fluorescent signal corresponding to c-jun (positive control) increases in CHX-treated cells compared to DMSO-treated cells.

Recently, it has been reported that 5-azacytidine (5-azaC) treatment alters NMD activity (Bhuvanagiri et al., 2014). Endogenous E6 and E7 RNA and protein levels were monitored in

SiHa and Ca Ski cells treated with 5-azaC or 5-aza-2'-deoxycytidine as negative control (5-azadC). In SiHa cells, the amount of E6 and E7 RNA increases in a dose-dependent fashion with 5-azaC compared to 5-azadC (Fig. 7A). As positive controls, two physiologic NMD substrates, SC35 and asparagine synthetase (ASNS) were tested and show, respectively, from 1.4 to 20-fold induction in 5-azaC (Fig. 7A, upper panel) compared to 5-azadC-treated cells (Fig. 7A, lower panel). Moreover, NMD inhibition, through 5-azaC, upregulates HPV16 E6 and E7 protein levels to 1.3 and 2.3 fold respectively (Fig. 7B). Consequently, the level of p53, targeted by E6, was decreased (Fig. 7B). Otherwise, no difference was observed in cells treated with 5-azadC. In contrast to SiHa, 5-azaC treatment in Ca Ski cells downregulates E6 and E7 mRNAs in a dose dependent-manner (Fig. S3A). This was accompanied by a downregulation of E6 and E7 proteins and an upregulation of p53 (Fig. S3B).

We next performed siRNA-mediated knockdown (KD) of NMD factors, such as UPF1 and SMG1 in SiHa and Ca Ski cells. Protein depletion around 60-80% of UPF1 and SMG1 was achieved (Figs. 8A and 8C). Interestingly, we observed that the inhibition of SMG1 component causes an increase of E6 and E7 protein levels in SiHa (Fig. 8A) and Ca Ski cells (Fig. 8C). Surprisingly, UPF1 knock down decreases E6 and E7 protein level (Figs. 8A and 8C). Concomitantly, quantification by RT-qPCR of HPV16 early transcripts (E6All, E6*I and E7) shows an increase of HPV16 transcripts from 1.5 to 2-fold in SMG1-depleted SiHa cells (Fig. 8B) but not in Ca Ski cells (Fig. 8D). No significant increase was observed for HPV16 early transcripts (E6All, E6*I and E7) in UPF1-depleted for SiHa cells (Fig. 8B). For both cell lines, NMD substrates (SC35 and ASNS) were increased in UPF1 and SMG1-depleted cells (Figs. 8B and 8D), yet with a stronger impact from the latter. Altogether, these results suggest that HPV16 E6/E7 transcripts can be target by NMD dependently of the cellular context.

NMD is altered in HPV16-infected cells

We previously observed some discrepancies between SiHa and Ca Ski cell lines concerning HPV16 gene expression after NMD inhibition. Both Ca Ski and SiHa cells are infected by

HPV16 but differ by their integration status. Ca Ski cells contain approximately 500 integrated copies per cell; whereas SiHa cells have only 1-2- integrated copies. HPV16 integration into host cell causes the production of fusion transcripts poorly characterized. Moreover, these cell lines were subjected to selective pressure for E6/E7 gene expression. In view of recent evidence indicates that some viruses develop strategies to escape to NMD, we can hypothesize that NMD activity in HPV-infected cells could be modified and/or HPV evolved to counteract NMD degradation pathway. Supporting this, we compared NMD activity in Ca Ski, SiHa cells and C33A HPV-negative cells, using WT or PTC β-globin NMD-reporter genes (Fig. 9A). As shown in Figure 9B, NMD induces PTC-β-globin reporter degradation in C33A with high efficiency compared to SiHa and Ca Ski cells. Thus, NMD is 5-fold and 2-fold less efficient in Ca Ski and SiHa cells respectively compared to C33A (Fig. 9C). In addition, NMD inhibition in HPV16+ cell lines is correlated with upregulated of NRP1 and GABARAPL1 physiological NMD targets whereas intronless non NMD target, SF3B5 and Hist2H3C mRNAs were unaffected (Fig. 9D).

Next, to get insights on the mechanism of NMD inhibition in HPV16+ cells lines, we analyzed HPV16 E6 protein effect on NMD activity. To do so, we explored the expression of β-globin NMD-reporters in C33A cells, transfected with HPV16 E6 WT containing vector. As shown in Figure 7E, E6Fl and E6*I were expressed from the vector. HPV16 E6 expression upregulated PTC-containing β-globin mRNAs, thus inhibiting the NMD efficiency (Figs. 9E and 9F). E6 affect seems to be dose-unspecific at 24h of transfection, but specific at 48h. Altogether, these data suggest that HPV16 could inhibit the NMD through E6 protein.

Discussion

Viral gene expression, performed by host cell machinery, is tightly regulated by both viral and cell factors. HPV16 gene expression is particularly regulated at posttranscriptional level (Graham, 2017; Graham and Faizo, 2017). Its polycistronic transcripts undergo alternative splicing and polyadenylation events that contribute to the regulation of expression but also

generates variability in viral transcriptome and proteome. In this study, we investigated the role of splicing mark EJC on HPV16 mRNAs. We observed distinct impacts of EJC components on HPV16 E6 and E7 expression. eIF4A3 or Y14 KD led to the upregulation of E6 and E7, whereas MLN51 KD led to its downregulation. Such discrepancies could be attributed to: (i) the composition of EJC onto HPV16 transcripts that may differ from cellular transcripts, and/or (ii) EJC-independent functions of eIF4A3, Y14 or MLN51. Indeed, eIF4A3, Y14 along with MAGOH are part of pre-EJC, a stable complex without MLN51, which exist on mRNA before exon ligation (Gehring et al., 2009). Recently, in addition to canonical EJCs, which are located around 20 nucleotides upstream the exon-exon junction, EJCs deposition were detected on the exons at noncanonical sites (ncEJC). To date, the role of noncanonical EJC deposition on an mRNA is unknown. Recent studies suggest that MLN51 may not be present in all EJCs, particularly in ncEJC (Hauer et al., 2016; Saulière et al., 2010, 2012; Singh et al., 2012). In addition, MLN51 has EJC-independent functions, involving in mRNA translation, P-bodies disassembly and stress granule assembly (Baguet et al., 2007; Chazal et al., 2013; Cougot et al., 2014). The pre-EJC components eIF4A3, Y14 and MAGOH affects strongly mRNA splicing whereas MLN51 has a milder affect (Ashton-Beaucage et al., 2010; Fukumura et al., 2016; Hayashi et al., 2014; Liu et al., 2016; Malone et al., 2014; Michelle et al., 2012; Roignant and Treisman, 2010; Wang et al., 2014). However, EJC core components KD did not affect HPV16 E6 and E7 mRNA splicing. Instead, eIF4A3 KD induced stabilization of E6 and E7 transcripts, in the Ca Ski cells, but not in SiHa cells. These results in SiHa cells may suggest that eIF4A3 and Y14 have an indirect effect on E6 and E7 expression. It could also due to technical bias.

Recently, eIF4A3 has been identified by mass spectrometry in RNPs containing HPV16 intron 1 indicating that eIF4A3 could be involved in HPV16 RNA processing (Martinez-Salazar M et al., 2014). Our results confirmed the eIF4A3 binding to HPV16 E6 and E7 mRNA, but also showed the existence of a splicing-independent eIF4A3 recruitment to the E6 ORF. eIF4A3 binding to E6 may thus be sequence-dependent although the intron 1 sequence was dispensable for this recruitment. The eIF4A3 protein is a DEAD box helicase, forming an

ATP-dependent clamp onto mRNA (Andersen et al., 2006; Ballut et al., 2005). However, a direct binding of eIF4A3 to E6 transcripts was not investigated in our study. The binding of eIF4A3 to HPV16 E6 and E7 mRNA suggests a direct effect on the regulation of their expression, instead of what we suggested above.

So, the EJC core components could affect viral mRNA stability through their role in NMD. Indeed, HPV16's, as many viral transcripts possess several NMD-triggering features such as multiple stop codons. When we inhibited NMD through drug targeting, we observed an upregulation of alternative E6 and E7 but not E6FL mRNAs in SiHa cells. Indeed, E6Fl levels remained unchanged under UPF1 or SMG1 KD. These results may suggest that HPV16 transcript were targeted by NMD, whereas E6FL mRNAs may be immunized. However, in Ca Ski cells we did not observe any upregulation of viral mRNAs suggesting that they were not targeted by NMD. Interestingly, although E6 and E7 mRNAs remained unchanged upon SMG1 KD, their protein levels were upregulated. This could be due to enhanced of E6 and E7 translation efficiency and/or increased of protein stability induced by SMG1 KD. Indeed, SMG1 is a kinase may phosphorylate E6 and E7 proteins, inducing a destabilization by preventing interaction with cellular partners.

HPV16 transcript may be immune to NMD in a way or another. Indeed, many RNA viruses whose RNAs were targeted by the NMD developed different strategies for NMD escaping (Balistreri et al., 2014; Garcia et al., 2014; LeBlanc and Beemon, 2004; Mocquet et al., 2012; Nakano et al., 2013). Rous Sarcoma Virus RNAs possess *cis*-acting element protecting them from NMD whereas Human T-cell Leukemia Virus 1 expresses Tax and Rex which inhibit the NMD (LeBlanc and Beemon, 2004; Mocquet et al., 2012; Nakano et al., 2013). Thus, HPV16 could inhibit NMD to protect his own mRNAs. When we monitored NMD activity in cervical carcinoma derived cell lines, we found that NMD was altered in HPV16+ SiHa and Ca Ski cells. In addition, HPV16 E6 ectopic expression in HPV- cells was sufficient to decrease NMD activity. Otherwise, HPV16 may inhibit NMD by targeting one or more its factors. Indeed, SMG1, a kinase involved in NMD, is expressed at lower levels in HPV-positive than in HPV-negative HNSCC, suggesting that it could one of the HPV16's targets

(Gubanova et al., 2012). However, neither SMG1 nor other UPF and EJC factor expression was impaired in HPV16+ cervical cancer cells (data not shown). The finding that HPV16 could inhibit the NMD suggested that he could exploit NMD to regulate viral transcripts expression. As NMD is known to regulate a subset of cellular mRNAs involved different pathways, HPV16 could also modulate host cell transcriptome. Interestingly, we showed an upregulation of NMD target SC35 1.6 and GABARAPL1 in HPV16+ cell lines. Therefore, it will be interesting to identify the mechanism whereby E6 inhibit NMD and the potential implication of NMD inhibition in HPV16 life cycle and induced carcinogenesis.

Materiels and Methods

Cell culture

The human cervical cancer derived cell lines, C33A (HPV-), SiHa (HPV16+) and Ca Ski (HPV16+) were cultured in EMEM, DMEM and RPMI 1640 media (Lonza, Verviers, Belgium) respectively. Osteosarcoma derived cells U2OS (HPV-) were cultured in DMEM. All media were supplemented with of 10% of Fetal Calf Serum (Eurobio, Courtaboeuf, France). Cells were incubated at 37°C in a humidified atmosphere with 5% of CO2.

siRNA, plasmid and chemical reagents

Negative control (SR-CL000-005), eIF4A3, Y14, MLN51, UPF1 and SMG1 targeting siRNA duplex were purchased from Eurogentec (Seraing, Belgium). siRNA targeting sequences are listed in Supplementary Table 1.

Intronless pXJ39 (-E6, E6*I and -E6^{mutSD226}) vectors were generated by a deletion of the rabbit β-globin intron, from corresponding pX40 plasmids, using Q5Site-Directed Mutagenesis Kit (New England Biolabs) following supplier's recommendations. pXJ40-empty and pXJ40-E6 vector (containing the wild type HPV16 E6 coding sequence) have been generously provided by Murielle Masson (ESBS, Illkirch, France). pXJ40-E6^{mutSD226} and -E6*I were generated from the pXJ40-E6, by the mutation of the splice site SD226 and the

deletion of E6 intron 1 (nucleotides from 227 to 408) respectively. Primers used for mutagenesis are listed in Supplementary Table 2.

Cycloheximides (Sigma), Actinomycin D (Euromedex), were used at a final concentration of 100 µg/ml and 5µg/ml respectively. The nucleotide analogues 5-Aza-2'-deoxycytidine and 5-Azacytidine were purchased from Epigentek and StressMarq Biosciences respectively.

Transfection

For all transfections, cells were plated in 6 well plate or 10cm dish (RNA-IP experiments) one day before transfection, so that they were at 60-75% confluent the day of transfection. siRNA transfections were performed using Lipofectamine 2000 reagent (Invitrogen, Carlsbad, USA) following supplier's recommendations. Cells were cultured at 37°C for 48h before cell collection.

Plasmid transfections were performed using JetPei reagent (Polyplus-transfection, Illkirch, France) according to manufacturer's instructions. Cells were incubated 24 h at 37°C before collection.

RNA extraction, RT-PCR and RT-qPCR

Total RNA was extracted using Ribozol reagent (VWR Chemicals) following manufacturer's instructions. For mRNA quantification, total RNAs were retro transcribed into cDNA using Maxima first strand cDNA synthesis kit (Thermo Scientific) according to supplier's recommendations. Quantitative PCR was performed with Power SYBR Green PCR mix (Life Technologies), using 7500 Real-Time PCR system (Applied Biosystems) according to supplier's instructions. Thermal Cycling of all qPCR consisted of 5 min at 95°C, followed by 40 cycles of 30 sec at 95°C and 1 min at 60°C. mRNA levels were normalized to rRNA18S using the $2^{-\Delta\Delta Ct}$ method. The specific primer sequences used for PCR and qPCR are listed in Supplementary Table 3. For semi-quantitative PCR, cDNAs were amplified using DreamTaq polymerase (Thermo Scientific) according to manufacturer's instructions. PCR products were analyzed in 2 % (w/v) agarose gel using Tris-Borate-EDTA buffer. The thermal conditions of PCR are described in Supplemental material and methods.

For miR-375 quantification, total RNAs were retro transcribed using the TaqMan™ MicroRNA Reverse Transcription Kit (Applied Biosystems) following supplier's recommendations. Quantitative PCR conditions used for miR-375 were the same as for

mRNA quantifications (see above). miR-375 levels were normalized to snRNA U6. The custom primers used for RT and for qPCR are listed in Supplementary Table 4.

Protein Extraction and Western blotting

Cell lysis was performed using Radio-immunoprecipitation buffer (50 mM Tris-HCl pH7.4, 150 mM NaCl, 1% NP40, 0.5% Na deoxycholate, 1mM EDTA) supplemented with 30 µg/mL anti-protease mixture (Roche Diagnostics, Mannheim, Germany) followed by sonication. After centrifugation at 10 000g for 10 min at 4°C, protein concentrations were determined using Bio-Rad Protein Assay (Bio-Rad, Munich, Germany) according to manufacturer's instructions.

Proteins were then separated by SDS-PAGE and transferred onto Amersham^{TM}Hybond^{TM}-PVDF membranes (GE Healthcare, Munich, Germany). Membranes were blocked overnight at 4°C with 5% nonfat milk (or 5% BSA for UPF1 and SMG1 antibodies), probed with primary antibodies (see below) then with HRP-conjugated secondary antibodies. Chemiluminescent signals were detected using PierceECL 2 Western Blotting Substrate (Thermo Scientific).

Antibodies

Mouse antibodies anti-E7, anti-E6 (2E-3F8 and 1E-6F4 clones), anti-p53 and anti-βactin (AC-15) were purchased from Santa Cruz, Euromedex and Sigma respectively. Rabbit antibodies anti-UPF1 and anti-SMG1 were purchased from Cell Signal. Mouse anti-eIF4A3 (2E5), rabbit anti-eIF4A3, rabbit anti-MLN51 (1608) were generously offered by Catherine Tommasetto (IGBMC, Ilkirch, Strasbourg). HRP-conjugated goat anti-mouse and anti-rabbit secondary antibodies were purchased from BD Pharmingen.

RNA-Immunoprecipitation

Cells were washed 3 times with ice cold PBS 1x, scrapped in 1ml PBS 1x per 10cm dish then subjected to centrifugation at 1500 g for 10 min at 4°C. 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 x 1 second and centrifuged at 16100 g for 10 min at 4°C.

For pre-clearing, 1 mg/ml of protein were added to 40 µl of protein A/G beads (Thermo Scientific), incubated under agitation at 4°C for 1h and centrifuged at 500 g for 5 min at 4°C. Supernatants have been divided on two fractions of 500 mg, which were incubated with 0.4

µg of either rabbit anti-eIF4A3 or IgG antibodies, under agitation for 2h at 4°C. Then, 40 µl of protein A/G beads and 1 mg/ml of yeast tRNA (Roche Diagnostics, Mannheim, Germany) were added to each fraction, and incubated overnight under agitation at 4°C.

Next, washing step were performed with NET-2 buffer 300 mM NaCl followed by centrifugation at 500 g for 5 min at 4°C. NET-2 buffer with 300 mM NaCl and 1 mg/ml heparin was added to tubes and incubated under agitation at 4°C for 10 min, followed by centrifugation at 500 g for 5 min at 4°C. After 8 washing steps with NET-2 buffer and centrifugation at 500 g for 5 min at 4°C, protein and antibodies complexes were eluted with 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 centrifugation at 500g for 5 min at 4°C. Eluate was splitted into 2 fractions: one fraction was used to perform a western blotting and the other was subjected to RNA extraction followed by RT-PCR analysis.

Immunofluorescence and RNA-FISH

For Immunofluorescence and RNA-FISH experiment, cells were cultured in coverslips on 24 well plates. siRNA transfections were performed as described above. Twenty-four hour after transfection, cells were fixed in 4% paraformaldehyde for 30 min at RT°. RNA-FISH were performed using RNAscope® Fluorescent Multiplex Assay (Advanced Cell Diagnostics Bio, Milan, Italy) following supplier's recommendations. Briefly, cells were dehydrated by successive incubations, for 2 min at RT°, with increasing concentrations of EtOH (50%, 70% then 100%). After changing EtOH 100%, cells were incubated 10 min at RT° before being rehydrated with incubation for 2 min at RT° with decreasing concentrations of EtOH (70% and 100%). Cells were incubated with PBS 1x for 10 min at RT° then subjected to permeabilization in PBS 1x with 0.1% Tween20 for 10 min at RT°. After washing in PBS 1x, the cells were incubated with Protease III for 10 min at RT°. Finally, cells were washed with PBS 1x and subjected to RNAscope probe hybridization. Coverslips were mounted on slides using fluorescent mounting medium and subjected to confocal microscopy revelation using Olympus FluoView1000 (Olympus).

Acknowledgements

The authors thanks to Murielle Masson for pXJ40-E6 vector, Véronique Dalstein and Christine Clavel for protein extracts from HNSCC, Catherine Tomasetto for NMD reporters, eIF4A3 and MLN51 antibodies. Julie Frejaville and Anne Peigney for the technical assistance.

K.M., J.P. and A.M. had fellowships from French Ministère de la Recherche et de la Technologie and Conseil Régional de Franche-Comté. Funds were from La Ligue Contre le Cancer (Comité du Doubs and CCIR-GE), Cancéropôle du Grand-Est (CGE), Université de Franche-Comté (UFC), Région Franche-Comté. The authors thank Fiona Ecarnot (EA3920, University Hospital Besancon, France) for editorial assistance.

References

Andersen, C.B.F., Ballut, L., Johansen, J.S., Chamieh, H., Nielsen, K.H., Oliveira, C.L.P., Pedersen, J.S., Séraphin, B., Le Hir, H., and Andersen, G.R. (2006). Structure of the exon junction core complex with a trapped DEAD-box ATPase bound to RNA. Science *313*, 1968–1972.

Ashton-Beaucage, D., Udell, C.M., Lavoie, H., Baril, C., Lefrançois, M., Chagnon, P., Gendron, P., Caron-Lizotte, O., Bonneil, E., Thibault, P., et al. (2010). The exon junction complex controls the splicing of MAPK and other long intron-containing transcripts in Drosophila. Cell *143*, 251–262.

Baguet, A., Degot, S., Cougot, N., Bertrand, E., Chenard, M.-P., Wendling, C., Kessler, P., Le Hir, H., Rio, M.-C., and Tomasetto, C. (2007). The exon-junction-complex-component metastatic lymph node 51 functions in stress-granule assembly. J. Cell. Sci. *120*, 2774–2784.

Balistreri, G., Horvath, P., Schweingruber, C., Zünd, D., McInerney, G., Merits, A., Mühlemann, O., Azzalin, C., and Helenius, A. (2014). The host nonsense-mediated mRNA decay pathway restricts Mammalian RNA virus replication. Cell Host Microbe *16*, 403–411.

Ballut, L., Marchadier, B., Baguet, A., Tomasetto, C., Séraphin, B., and Le Hir, H. (2005). The exon junction core complex is locked onto RNA by inhibition of eIF4AIII ATPase activity. Nat. Struct. Mol. Biol. *12*, 861–869.

Bhuvanagiri, M., Lewis, J., Putzker, K., Becker, J.P., Leicht, S., Krijgsveld, J., Batra, R., Turnwald, B., Jovanovic, B., Hauer, C., et al. (2014). 5-azacytidine inhibits nonsense-mediated decay in a MYC-dependent fashion. EMBO Mol Med *6*, 1593–1609.

Chazal, P.-E., Daguenet, E., Wendling, C., Ulryck, N., Tomasetto, C., Sargueil, B., and Le Hir, H. (2013). EJC core component MLN51 interacts with eIF3 and activates translation. Proc. Natl. Acad. Sci. U.S.A. *110*, 5903–5908.

493 Chen, J., Xue, Y., Poidinger, M., Lim, T., Chew, S.H., Pang, C.L., Abastado, J.-P., and
494 Thierry, F. (2014). Mapping of HPV transcripts in four human cervical lesions using RNAseq
495 suggests quantitative rearrangements during carcinogenic progression. Virology *462–463*, 14–
496 24.

497 Cougot, N., Daguenet, E., Baguet, A., Cavalier, A., Thomas, D., Bellaud, P., Fautrel, A.,
498 Godey, F., Bertrand, E., Tomasetto, C., et al. (2014). Overexpression of MLN51 triggers P-
499 body disassembly and formation of a new type of RNA granules. J. Cell. Sci. *127*, 4692–
500 4701.

501 Fukumura, K., Wakabayashi, S., Kataoka, N., Sakamoto, H., Suzuki, Y., Nakai, K., Mayeda,
502 A., and Inoue, K. (2016). The Exon Junction Complex Controls the Efficient and Faithful
503 Splicing of a Subset of Transcripts Involved in Mitotic Cell-Cycle Progression. Int J Mol Sci
504 *17*.

505 Garcia, D., Garcia, S., and Voinnet, O. (2014). Nonsense-mediated decay serves as a general
506 viral restriction mechanism in plants. Cell Host Microbe *16*, 391–402.

507 Gehring, N.H., Lamprinaki, S., Hentze, M.W., and Kulozik, A.E. (2009). The hierarchy of
508 exon-junction complex assembly by the spliceosome explains key features of mammalian
509 nonsense-mediated mRNA decay. PLoS Biol. *7*, e1000120.

510 Giorgi, C., and Moore, M.J. (2007). The nuclear nurture and cytoplasmic nature of localized
511 mRNPs. Semin. Cell Dev. Biol. *18*, 186–193.

512 Graham, S.V. (2017). Keratinocyte Differentiation-Dependent Human Papillomavirus Gene
513 Regulation. Viruses *9*.

514 Graham, S.V., and Faizo, A.A.A. (2017). Control of human papillomavirus gene expression
515 by alternative splicing. Virus Res. *231*, 83–95.

516 Guan, P., Howell-Jones, R., Li, N., Bruni, L., de Sanjosé, S., Franceschi, S., and Clifford,
517 G.M. (2012). Human papillomavirus types in 115,789 HPV-positive women: a meta-analysis
518 from cervical infection to cancer. Int. J. Cancer *131*, 2349–2359.

519 Gubanova, E., Brown, B., Ivanov, S.V., Helleday, T., Mills, G.B., Yarbrough, W.G., and
520 Issaeva, N. (2012). Downregulation of SMG-1 in HPV-positive head and neck squamous cell
521 carcinoma due to promoter hypermethylation correlates with improved survival. Clin. Cancer
522 Res. *18*, 1257–1267.

523 Hauer, C., Sieber, J., Schwarzl, T., Hollerer, I., Curk, T., Alleaume, A.-M., Hentze, M.W.,
524 and Kulozik, A.E. (2016). Exon Junction Complexes Show a Distributional Bias toward
525 Alternatively Spliced mRNAs and against mRNAs Coding for Ribosomal Proteins. Cell Rep
526 *16*, 1588–1603.

527 Hayashi, R., Handler, D., Ish-Horowicz, D., and Brennecke, J. (2014). The exon junction
528 complex is required for definition and excision of neighboring introns in Drosophila. Genes
529 Dev. *28*, 1772–1785.

530 Jung, H.M., Phillips, B.L., and Chan, E.K. (2014). miR-375 activates p21 and suppresses
531 telomerase activity by coordinately regulating HPV E6/E7, E6AP, CIP2A, and 14-3-3ζ. Mol.
532 Cancer *13*, 80.

Le Hir, H., Gatfield, D., Izaurralde, E., and Moore, M.J. (2001). 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.

Le Hir, H., Saulière, J., and Wang, Z. (2016). The exon junction complex as a node of post-transcriptional networks. Nat. Rev. Mol. Cell Biol. *17*, 41–54.

LeBlanc, J.J., and Beemon, K.L. (2004). Unspliced Rous sarcoma virus genomic RNAs are translated and subjected to nonsense-mediated mRNA decay before packaging. J. Virol. *78*, 5139–5146.

Lejeune, F. (2017). Nonsense-mediated mRNA decay at the crossroads of many cellular pathways. BMB Rep *50*, 175–185.

Liu, M., Li, Y., Liu, A., Li, R., Su, Y., Du, J., Li, C., and Zhu, A.J. (2016). The exon junction complex regulates the splicing of cell polarity gene dlg1 to control Wingless signaling in development. Elife *5*.

Ma, X.M., Yoon, S.-O., Richardson, C.J., Jülich, K., and Blenis, J. (2008). SKAR links pre-mRNA splicing to mTOR/S6K1-mediated enhanced translation efficiency of spliced mRNAs. Cell *133*, 303–313.

Malone, C.D., Mestdagh, C., Akhtar, J., Kreim, N., Deinhard, P., Sachidanandam, R., Treisman, J., and Roignant, J.-Y. (2014). The exon junction complex controls transposable element activity by ensuring faithful splicing of the piwi transcript. Genes Dev. *28*, 1786–1799.

Martin, K.C., and Ephrussi, A. (2009). mRNA localization: gene expression in the spatial dimension. Cell *136*, 719–730.

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., Martines-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.

Michelle, L., Cloutier, A., Toutant, J., Shkreta, L., Thibault, P., Durand, M., Garneau, D., Gendron, D., Lapointe, E., Couture, S., et al. (2012). Proteins associated with the exon junction complex also control the alternative splicing of apoptotic regulators. Mol. Cell. Biol. *32*, 954–967.

Mocquet, V., Neusiedler, J., Rende, F., Cluet, D., Robin, J.-P., Terme, J.-M., Duc Dodon, M., Wittmann, J., Morris, C., Le Hir, H., et al. (2012). 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.

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.

Nakano, K., Ando, T., Yamagishi, M., Yokoyama, K., Ishida, T., Ohsugi, T., Tanaka, Y., Brighty, D.W., and Watanabe, T. (2013). Viral interference with host mRNA surveillance, the nonsense-mediated mRNA decay (NMD) pathway, through a new function of HTLV-1 Rex: implications for retroviral replication. Microbes Infect. *15*, 491–505.

Nott, A., Le Hir, H., and Moore, M.J. (2004). Splicing enhances translation in mammalian cells: an additional function of the exon junction complex. Genes Dev. *18*, 210–222.

Palacios, I.M., Gatfield, D., St Johnston, D., and Izaurralde, E. (2004). An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay. Nature *427*, 753–757.

Roignant, J.-Y., and Treisman, J.E. (2010). Exon junction complex subunits are required to splice Drosophila MAP kinase, a large heterochromatic gene. Cell *143*, 238–250.

Roman, A., and Munger, K. (2013). The papillomavirus E7 proteins. Virology *445*, 138–168.

Saulière, J., Haque, N., Harms, S., Barbosa, I., Blanchette, M., and Le Hir, H. (2010). The exon junction complex differentially marks spliced junctions. Nat. Struct. Mol. Biol. *17*, 1269–1271.

Saulière, J., Murigneux, V., Wang, Z., Marquenet, E., Barbosa, I., Le Tonquèze, O., Audic, Y., Paillard, L., Roest Crollius, H., and Le Hir, H. (2012). CLIP-seq of eIF4AIII reveals transcriptome-wide mapping of the human exon junction complex. Nat. Struct. Mol. Biol. *19*, 1124–1131.

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 Primers *2*, 16086.

Schwartz, S. (2013). Papillomavirus transcripts and posttranscriptional regulation. Virology *445*, 187–196.

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, G., Kucukural, A., Cenik, C., Leszyk, J.D., Shaffer, S.A., Weng, Z., and Moore, M.J. (2012). The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. Cell *151*, 750–764.

Vande Pol, S.B., and Klingelhutz, A.J. (2013). Papillomavirus E6 oncoproteins. Virology *445*, 115–137.

Wang, Z., Murigneux, V., and Le Hir, H. (2014). Transcriptome-wide modulation of splicing by the exon junction complex. Genome Biol. *15*, 551.

Figures Legends

Figure 1: Y14 and eIF4A3 depletion causes an increase of HPV16 E6/E7 expression and MLN51 depletion induced a decrease of HPV16 E6/E7 expression. (A) Representative scheme of unspliced and alternatively spliced HPV16 early transcripts. Real time PCR primer sets are in dotted line (E6All, E6Fl and E7) and RT-PCR primer sets are represented by arrows (F6, R6 and R7). (B) RNAs relative expression levels of EJC core factors in eIF4A3, Y14 or MLN51-depleted SiHa and Ca Ski cells. Dashed line represents mRNA levels in control siRNA-treated cells. Error bars = s.e.m., n=3. (C) RNAs relative expression levels of HPV16 E6/E7 early transcripts in eIF4A3, Y14 or MLN51-depleted cells. Dashed line represents mRNA levels in control siRNA-treated cells. Error bars = s.e.m., n=3. *, $p < 0.05$ (student's t-test using siRNA CTL treated cells as a reference (D) Western blot showing endogenous E6 and E7 protein levels in RNAi knockdown (KD) of eIF4A3, Y14 or MLN51 in SiHa (left panel) Ca Ski (right panel) cells.

Figure 2: EJC core factors depletion does not alter the splicing pattern of HPV16 early transcripts. (A and C) RT-PCR showing alternative splicing patterns of HPV16 E6/E7 transcripts in eIF4A3, Y14 or MLN51-depleted SiHa (A) and Ca Ski (C). KPNA1 is used as positive control; B2M mRNAs and 18S rRNA are both used as loading controls. (B and D) E6Fl/E6All ratio of RNA levels quantified by real-time PCR in EJC-depleted SiHa (B) and Ca Ski (D) cells. Error bars = s.e.m., n=3.

Figure 3: Control of HPV16 early transcripts stability by eIF4A3 in Ca Ski cells. (A). RT-qPCR of HPV16 alternatively spliced transcripts in SiHa and Ca Ski cells depleted (gray) or not (black) for eIF4A3, and treated with 5 µg/ml of actinomycin D for 0, 30 min, 1h, 1h30min, 3h, 5h, and 8h to determine mRNAs half-lives. 18S rRNA is used for normalization. Error bars = s.e.m., n=3. (B) Fold change of mRNA half-lives relative to control conditions. Dashed line represents mRNA half-lives in control siRNA-treated cells. Error bars = s.e.m., n=3. *, $p < 0.05$ (student's t-test).

Figure 4: eIF4A3 associates with HPV16 early transcripts. (A) Western blots showing the immunoprecipitation of eIF4A3 and MLN51 protein in Ca Ski (left panel) and SiHa (right panel) cells. (B) RIP analysis of RNAs from Ca Ski (left panel) and SiHa (right panel) cells

associated with either control (rabbit IgG) or eIF4A3 antibodies. The KPNA1 mRNAs are used as positive control and the SF3B5 intronless mRNAs are used as negative control.

Figure 5: The intron 1 part of E6 ORF is dispensable for the recruitment of eIF4A3. (A) Representative scheme of unspliced and/or alternatively spliced HPV16 E6 transcripts expressed from -E6WT, -E6^{mutSD226} and -E6*I vector. RT-PCR primer sets are represented by arrows (F6 and R6). (B) blots showing the immunoprecipitation of eIF4A3 protein in U2OS cells transfected with a vector containing the various sequence of E6 ORF represented in (A). The HPV16 E6 WT (1), HPV16 E6^{mutSD226} (2) and E6*I (3) containing vectors were transfected. (C) RIP of mRNAs associated with either control (rabbit IgG) or eIF4A3 antibodies in U2OS cells transfected as in (B).

Figure 6: Inhibition of NMD through CHX upregulates E6 and E7 mRNA levels in SiHa but not in Ca Ski cells. (A) RT-qPCR analysis of SC35, ASNS and HPV16 E6all, E6*I, E6Fl and E7 mRNA expression in SiHa (upper panel) and Ca Ski (lower panel) cells treated for 6h with DMSO or CHX. rRNA 18S was used for normalization. Error bars = s.e.m., n=2. *, $p < 0.05$ (student's t-test). (B) HPV16 E6/E7 RNA expression in SiHa cells treated with DMSO or CHX analyzed by RNA FISH. Cells were treated for 6h with DMSO or CHX then subjected to RNA-FISH as described in Material and methods. (C) The number of HPV16 RNA dots per cell was counted in 50 and 53 cells for DMSO and CHX treated cells respectively.

Figure 7: Inhibition of NMD through 5-azaC upregulates E6 and E7 gene expression in SiHa cells. (A) RT-qPCR analysis of SC35, ASNS and HPV16 E6all, E6*I, E6Fl and E7 expression in SiHa cells treated for 18h with increasing amount (0, 0.5, 1, 3, 5 and 10 μM) of 5-azaC (upper panel) or 5-azadC (lower panel). rRNA 18S was used for normalization. Error bars = s.e.m., n=3. *, $p < 0.05$ (student's t-test using non-treated cells as a reference). (B) Western blot analysis of p53, HPV16 E6 and E7 protein expression in SiHa cells treated for 18h with 0, 5 and 10 μM of 5-azaC or 5-azadC. βactin was used as a loading control. Relative quantification of E6 and E7 expression was shown under each blot.

Figure 8: Inhibition of NMD factors alters HPV16 protein expression. (A and C) Relative expression levels of SC35, ASNS and HPV16 E6All, E6*I, E6Fl and E7 mRNAs in SiHa (A) and Ca Ski (C) cells depleted for UPF1 and SMG1. rRNA 18S was used for normalization.

Error bars = s.e.m., n=3. *, $p < 0.05$ (student's t-test using siRNA CTL treated cells as a reference). (B and D) Western blot showing endogenous UPF1, SMG1, E6 and E7 protein levels in RNAi knockdown of UPF1 and SMG1 in SiHa (B) and Ca Ski (D) cells. βactin was used as a loading control.

Figure 9: HPV16 E6 protein inhibit NMD activity in C33A cells. (A) Schematic representation of the β-globin wild type (WT) and PTC-containing (PTC) reporters. (B) RT-PCR showing WT and PTC b-globin RNAs levels in C33A, SiHa and Ca Ski cells. Quantifications of WT and PTC b-globin mRNA expression were represented as percentage of the WT under each lane. The luciferase mRNAs level was used as for normalization. (C) NMD efficiency is represented as the fold change of PTC-containing β-globin mRNA relative to C33A cells. (D) Relative mRNAs levels of endogenous NMD substrates and NMD-insensitive mRNAs (intronless mRNAs) in C33A, SiHa and Ca Ski cells. (E) RT-PCR showing WT and PTC β-globin RNAs levels in C33A cells transfected with increasing amount (1 µg or 2 µg) of either control or E6-expressing vector, for 24h or 48h. The luciferase mRNAs level was used as internal control. (F) NMD efficiency calculated by the quantification of the relative PTC- β-globin mRNA from (E). NMD efficiency is represented as the fold change of PTC-containing β-globin mRNA relative to C33A cells

Supplementary Figure 1: Splicing pattern of HPV16 early transcripts. RT-PCR showing alternative splicing patterns of HPV16 E6/E7 transcripts in SiHa and Ca Ski cells; C33A cell line are HPV-negative cervical carcinoma. B2M mRNAs and 18S rRNA are both used as loading controls.

Supplementary Figure 2: HPV16 early transcripts stability in Y14-depleted SiHa cells. (A) RT-qPCR of HPV16 alternatively spliced transcripts in SiHa cells depleted (gray) or not (black) for Y14, and treated with 5 µg/ml of actinomycin D for 0, 30 min, 1h, 1h30min, 2h, 3h, and 5h to determine mRNAs half-lives. 18S rRNA is used for normalization. Error bars = s.e.m., n=3. (B) Fold change of mRNA half-lives relative to control conditions. Dashed line represents mRNA half-lives in control siRNA-treated cells. Error bars = s.e.m., n=3.

Supplementary Figure 3: Expression levels of miR-375 under depletion of EJC core factors. RT-qPCR analysis of miR-375 expression in eIF4A3, Y14 or MLN51 depleted SiHa

and Ca Ski cells. snRNA U6 RNA levels were used for normalization. Error bars = s.e.m., n=2 for SiHa cells.

Supplementary Figure 4: Inhibition of NMD through CHX upregulates HPV16 early gene expression in SiHa but not in Ca Ski cells. (A) RT-qPCR analysis of SC35, ASNS and HPV16 E6all, E6*I, E6Fl and E7 expression in Ca Ski cells treated for 18h with increasing amount (0, 0.5, 1, 3, 5 and 10 μM) of 5-azaC (upper panel) or 5-azadC (lower panel). rRNA 18S was used for normalization. Error bars = s.e.m., n=3. *, $p < 0.05$ (student's t-test using non-treated cells as a reference); **, $p < 0.01$. (B) Western blot analysis of p53, HPV16 E6 and E7 protein expression in SiHa cells treated for 18h with 0, 5 and 10 μM of 5-azaC or 5-azadC. βactin was used as a loading control.

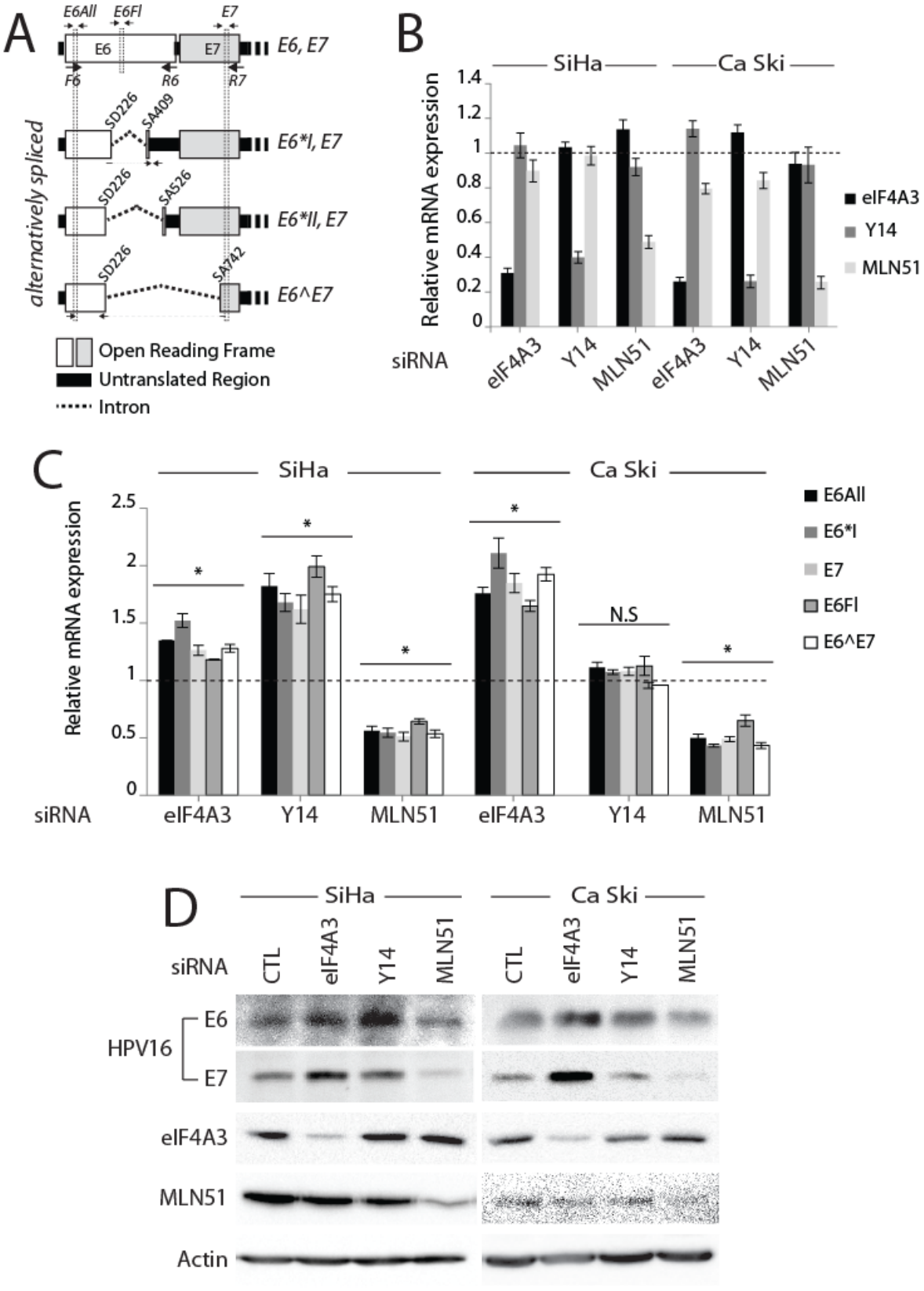

A
E6All
E6Fl
E7
E6
E7
E6, E7
F6
R6
R7
alternatively spliced
SD226
SA409
E6*I, E7
SD226
SA526
E6*II, E7
SD226
SA742
E6^E7
Open Reading Frame
Untranslated Region
Intron
B
SiHa
Ca Ski
Relative mRNA expression
1.4
1.2
1
0.8
0.6
0.4
0.2
0
siRNA
eIF4A3
Y14
MLN51
eIF4A3
Y14
MLN51
eIF4A3
Y14
MLN51
C
SiHa
Ca Ski
Relative mRNA expression
2.5
2
1.5
1
0.5
0
*
*
*
*
N.S
*
siRNA
eIF4A3
Y14
MLN51
eIF4A3
Y14
MLN51
E6All
E6*I
E7
E6Fl
E6^E7
D
SiHa
Ca Ski
siRNA
CTL
eIF4A3
Y14
MLN51
CTL
eIF4A3
Y14
MLN51
HPV16
E6
E7
eIF4A3
MLN51
Actin

Fig 1

719

720

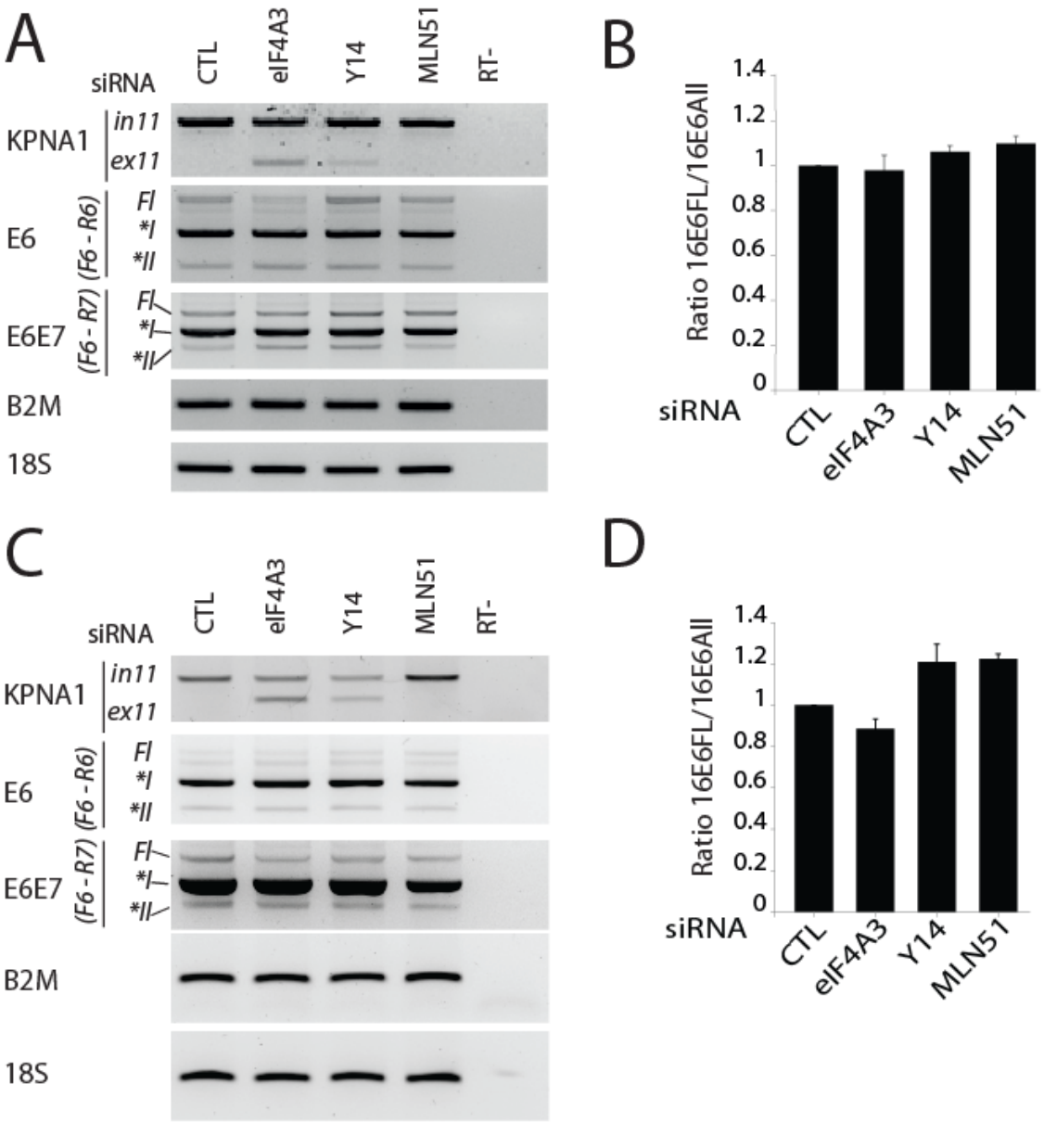

A
siRNA
CTL
eIF4A3
Y14
MLN51
RT-
KPNA1
in11
ex11
E6
(F6 - R6)
Fl
*I
*II
E6E7
(F6 - R7)
B2M
18S
B
Ratio 16E6FL/16E6All
1.4
1.2
1
0.8
0.6
0.4
0.2
0
siRNA
CTL
eIF4A3
Y14
MLN51
C
D

Fig 2

721

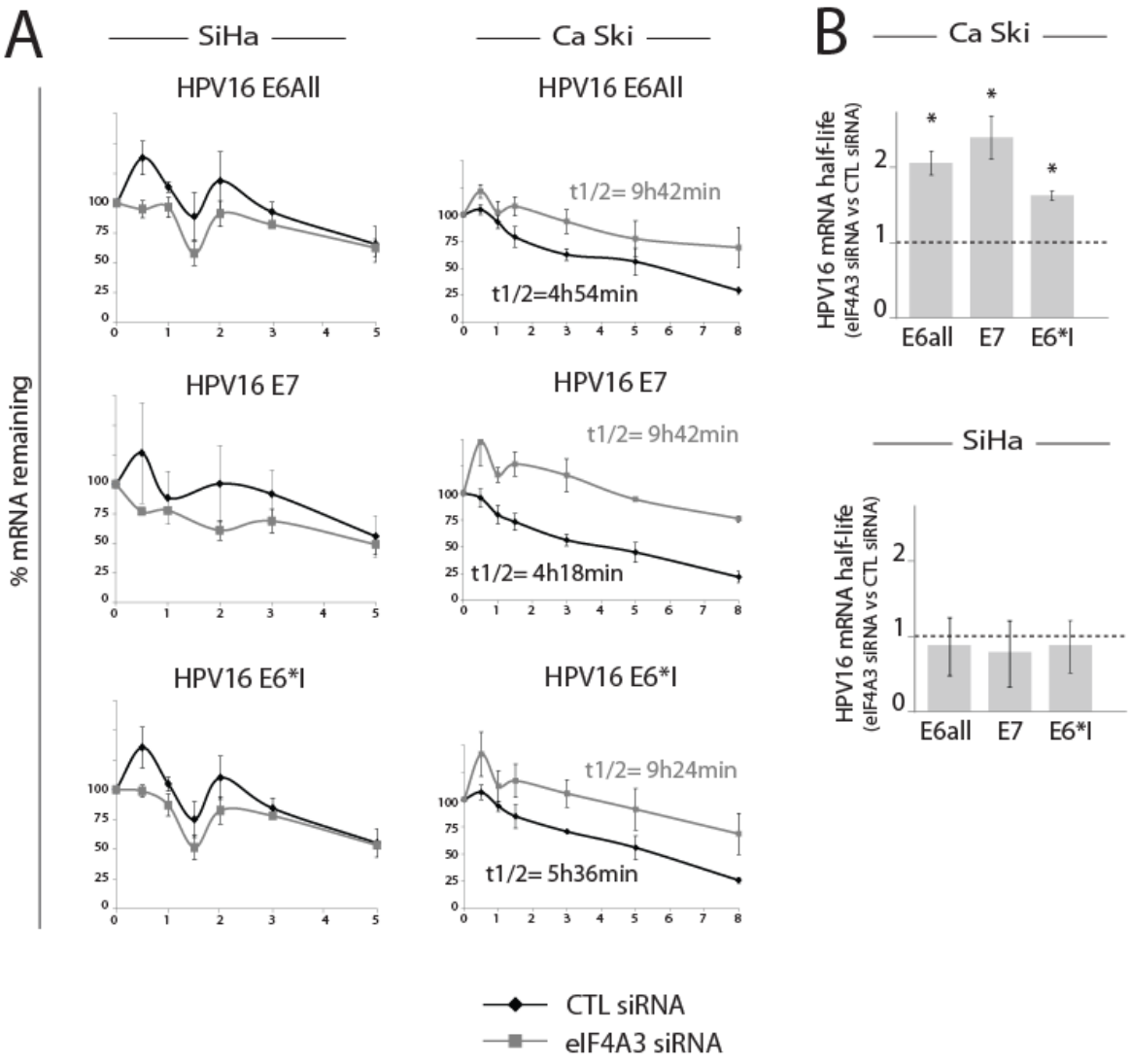

A
SiHa
Ca Ski
HPV16 E6All
HPV16 E7
HPV16 E6*I
% mRNA remaining
t1/2= 9h42min
t1/2=4h54min
t1/2= 9h42min
t1/2= 4h18min
t1/2= 9h24min
t1/2= 5h36min
CTL siRNA
eIF4A3 siRNA
B
Ca Ski
HPV16 mRNA half-life
(eIF4A3 siRNA vs CTL siRNA)
E6all
E7
E6*I
SiHa

Fig 3

722

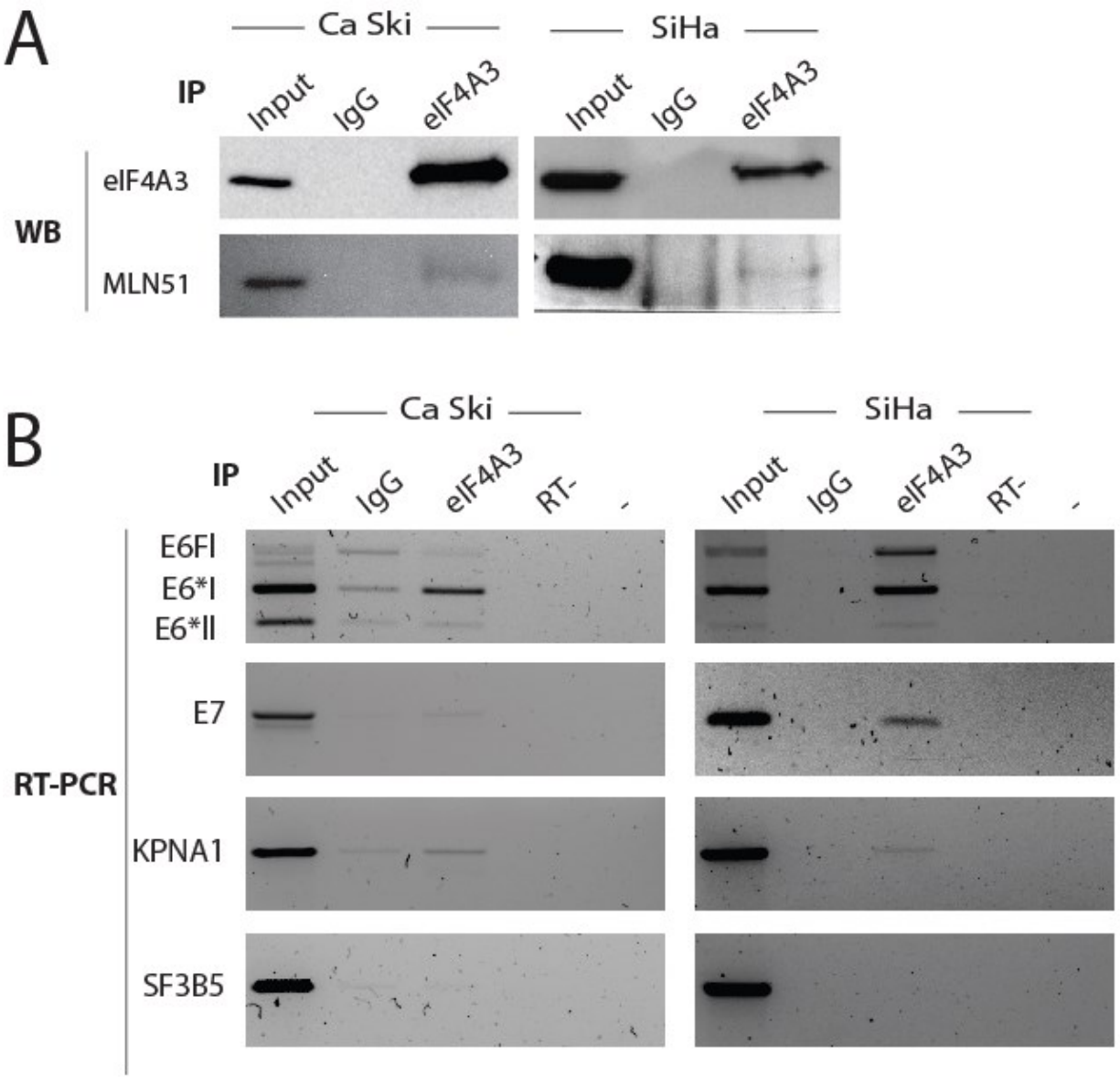

A
Ca Ski
SiHa
IP
Input
IgG
eIF4A3
Input
IgG
eIF4A3
WB
eIF4A3
MLN51
B
Ca Ski
SiHa
IP
Input
IgG
eIF4A3
RT-
-
Input
IgG
eIF4A3
RT-
-
RT-PCR
E6FI
E6*I
E6*II
E7
KPNA1
SF3B5

Fig 4

723

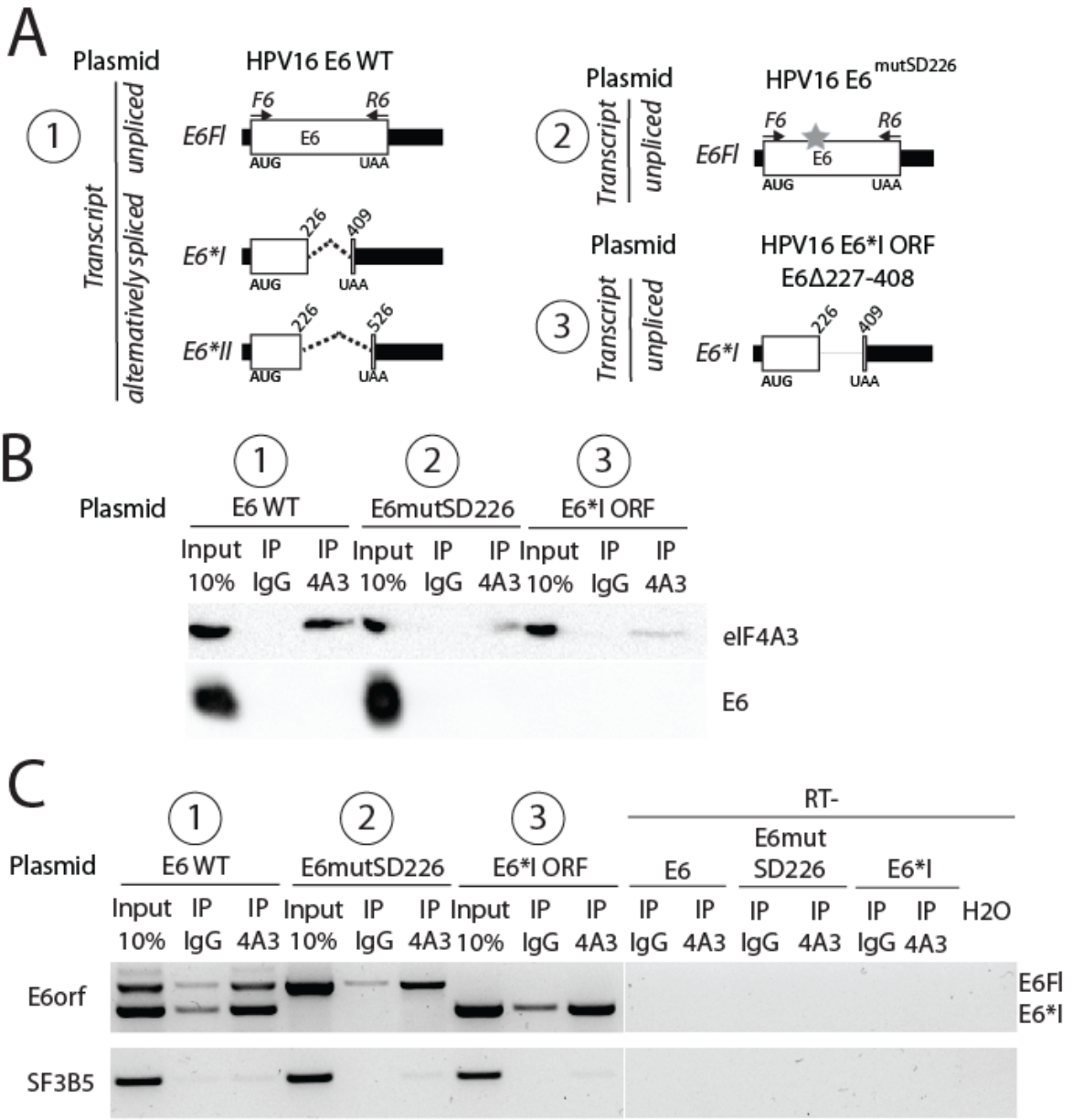

A
Plasmid
HPV16 E6 WT
1
Transcript
unpliced
alternatively spliced
E6Fl
F6
R6
E6
AUG
UAA
E6*I
226
409
E6*II
526
Plasmid
HPV16 E6 mutSD226
2
Transcript
unpliced
HPV16 E6*I ORF
E6Δ227-408
3
B
Plasmid
E6 WT
E6mutSD226
E6*I ORF
Input 10%
IP IgG
IP 4A3
eIF4A3
E6
C
RT-
E6mut SD226
E6
E6*I
H2O
E6orf
E6Fl
E6*I
SF3B5

Fig 5

724

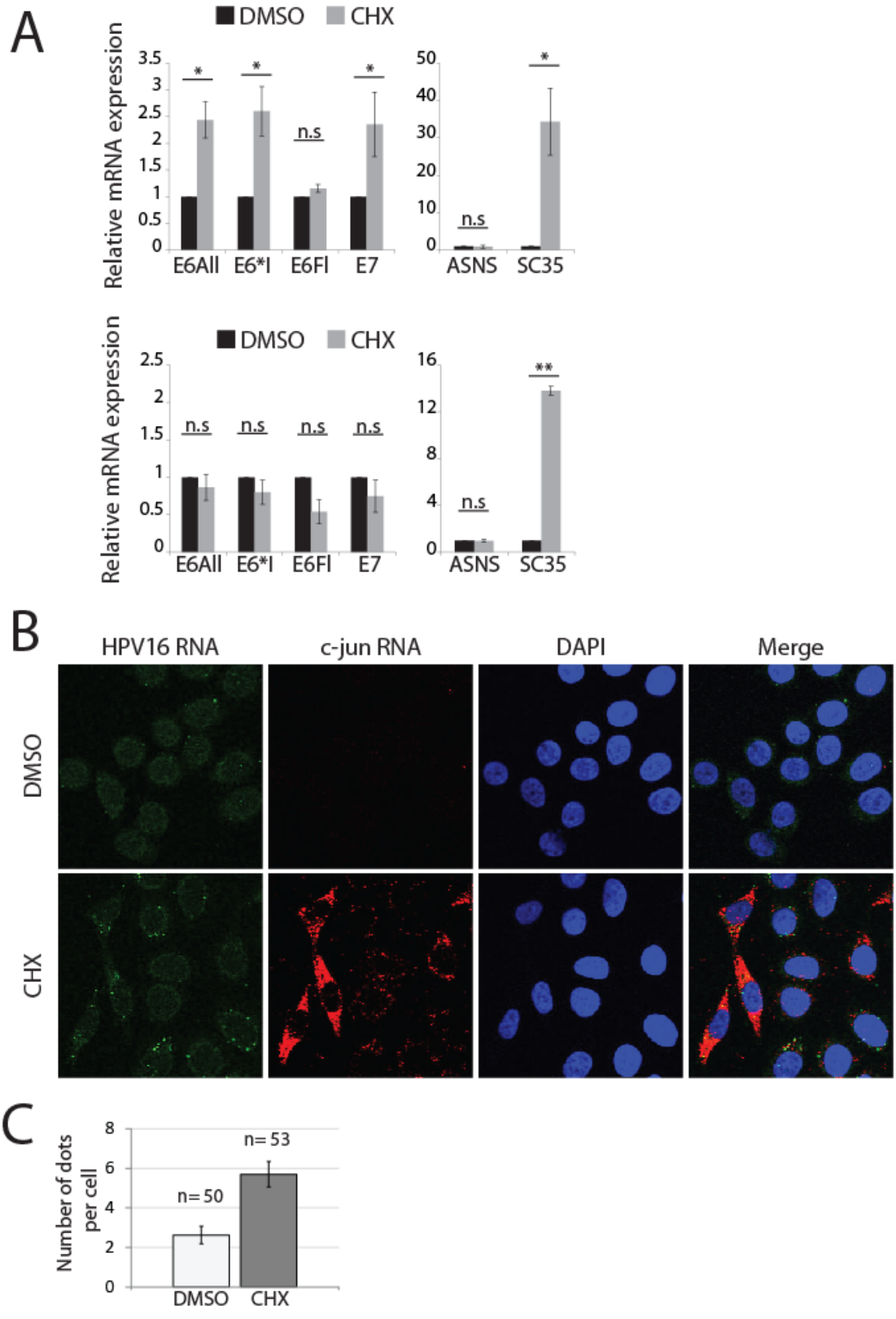

A
DMSO
CHX
Relative mRNA expression
3.5
3
2.5
2
1.5
1
0.5
0
*
*
n.s
*
E6AII
E6*I
E6FI
E7
50
40
30
20
10
0
n.s
*
ASNS
SC35
Relative mRNA expression
2.5
2
1.5
1
0.5
0
n.s
n.s
n.s
n.s
16
12
8
4
0
n.s
**
B
HPV16 RNA
c-jun RNA
DAPI
Merge
DMSO
CHX
C
Number of dots per cell
8
6
4
2
0
n= 50
n= 53
DMSO
CHX

Fig 6

725

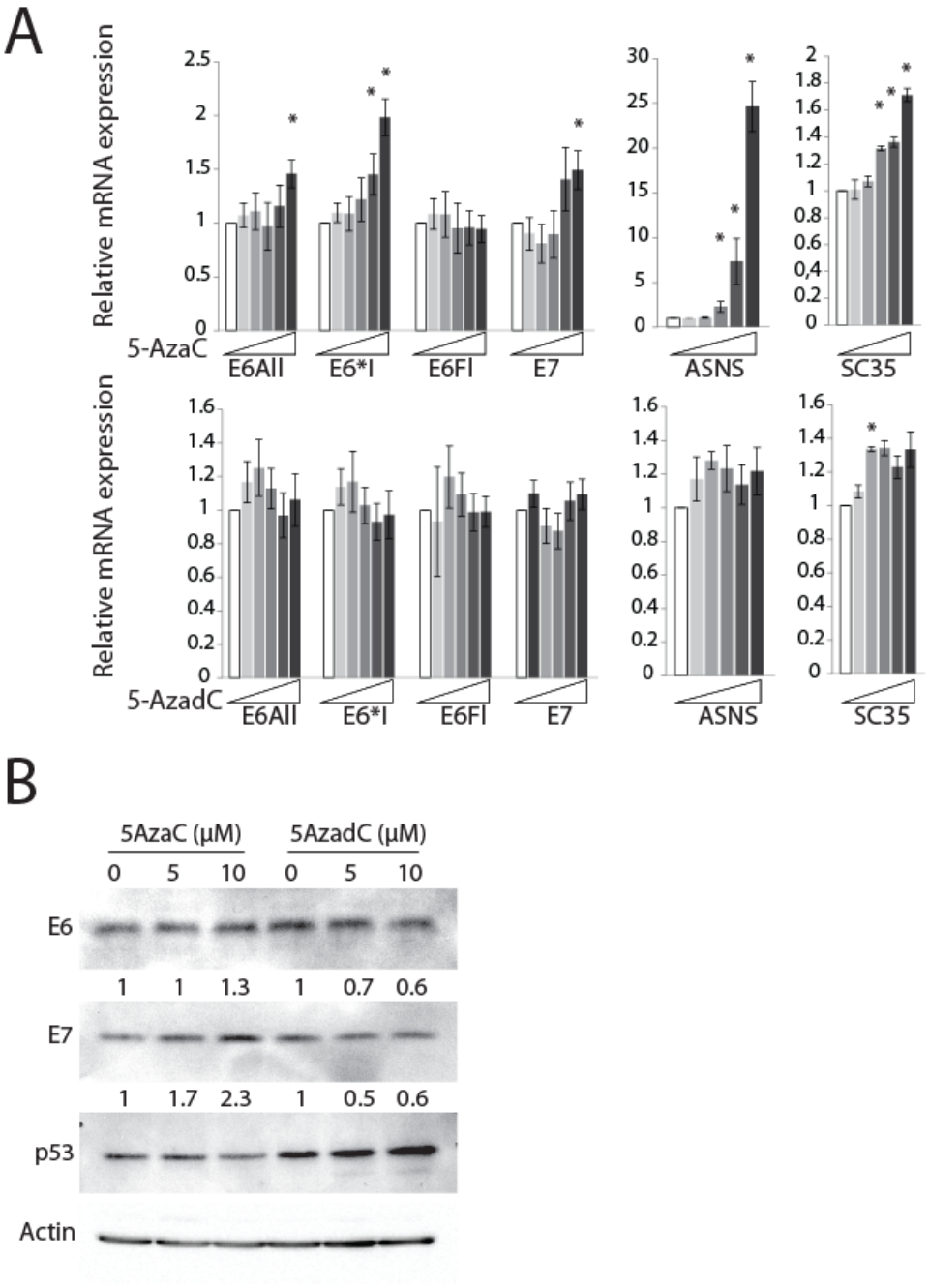

A
Relative mRNA expression
5-AzaC
E6AII
E6*I
E6FI
E7
ASNS
SC35
Relative mRNA expression
5-AzadC
E6AII
E6*I
E6FI
E7
ASNS
SC35
B
5AzaC (µM)
5AzadC (µM)
0 5 10 0 5 10
E6
1 1 1.3 1 0.7 0.6
E7
1 1.7 2.3 1 0.5 0.6
p53
Actin

Fig 7

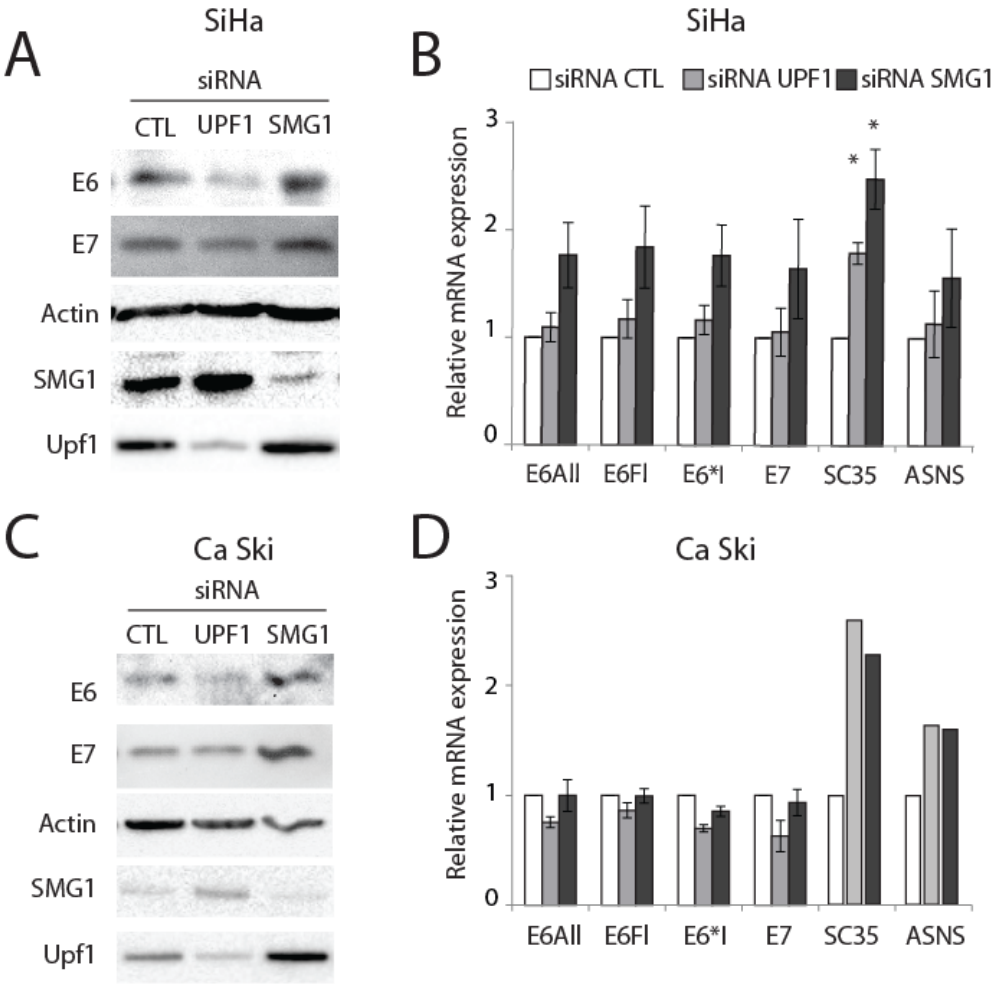

A
SiHa
siRNA
CTL
UPF1
SMG1
E6
E7
Actin
SMG1
Upf1
B
SiHa
siRNA CTL
siRNA UPF1
siRNA SMG1
Relative mRNA expression
0
1
2
3
*
*
E6AII
E6FI
E6*I
E7
SC35
ASNS
C
Ca Ski
siRNA
CTL
UPF1
SMG1
E6
E7
Actin
SMG1
Upf1
D
Ca Ski
Relative mRNA expression
0
1
2
3
E6AII
E6FI
E6*I
E7
SC35
ASNS

Fig 8

728

729

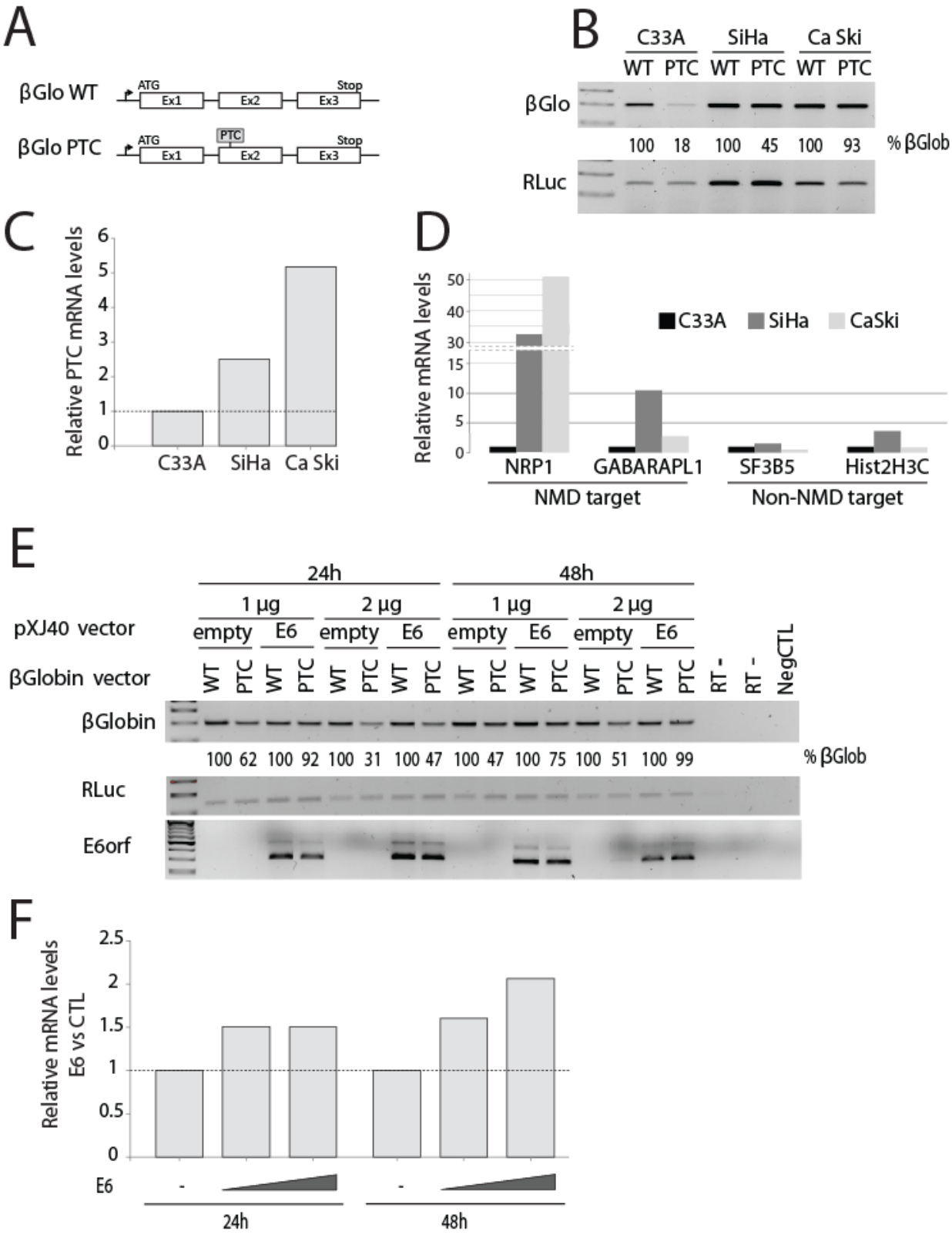

A
βGlo WT
ATG
Ex1
Ex2
Ex3
Stop
βGlo PTC
PTC
B
C33A
SiHa
Ca Ski
WT PTC WT PTC WT PTC
βGlo
100 18 100 45 100 93
% βGlob
RLuc
C
Relative PTC mRNA levels
C33A
SiHa
Ca Ski
D
Relative mRNA levels
C33A
SiHa
CaSki
NRP1
GABARAPL1
SF3B5
Hist2H3C
NMD target
Non-NMD target
E
24h
48h
1 µg
2 µg
pXJ40 vector
empty
E6
βGlobin vector
WT
PTC
RT -
NegCTL
βGlobin
100 62 100 92 100 31 100 47 100 47 100 75 100 51 100 99
% βGlob
RLuc
E6orf
F
Relative mRNA levels
E6 vs CTL
E6
24h
48h

Fig 9

730

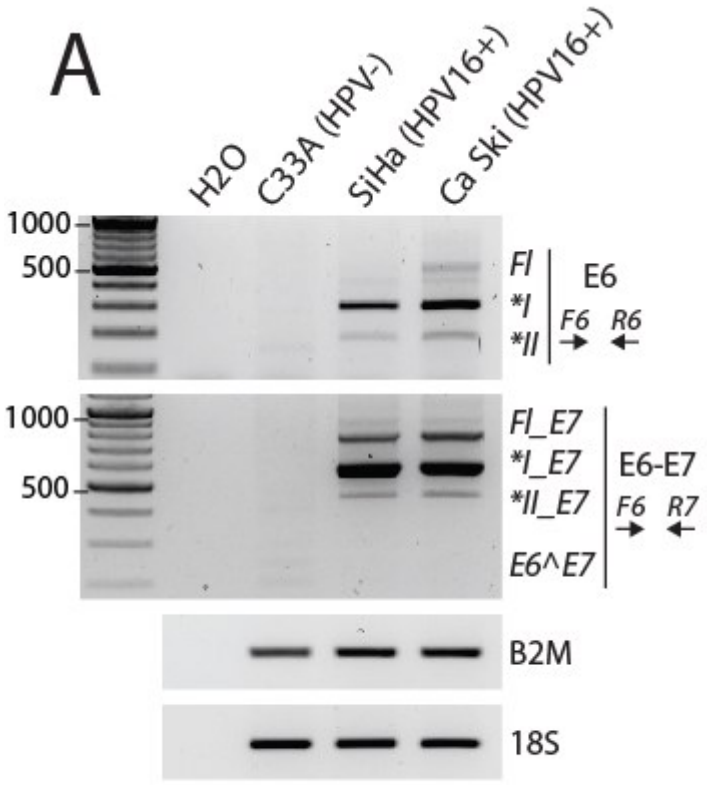

A
H2O
C33A (HPV-)
SiHa (HPV16+)
Ca Ski (HPV16+)
1000
500
Fl
*I
*II
E6
F6
R6
1000
500
Fl_E7
*I_E7
*II_E7
E6^E7
E6-E7
F6
R7
B2M
18S

Fig S1

731

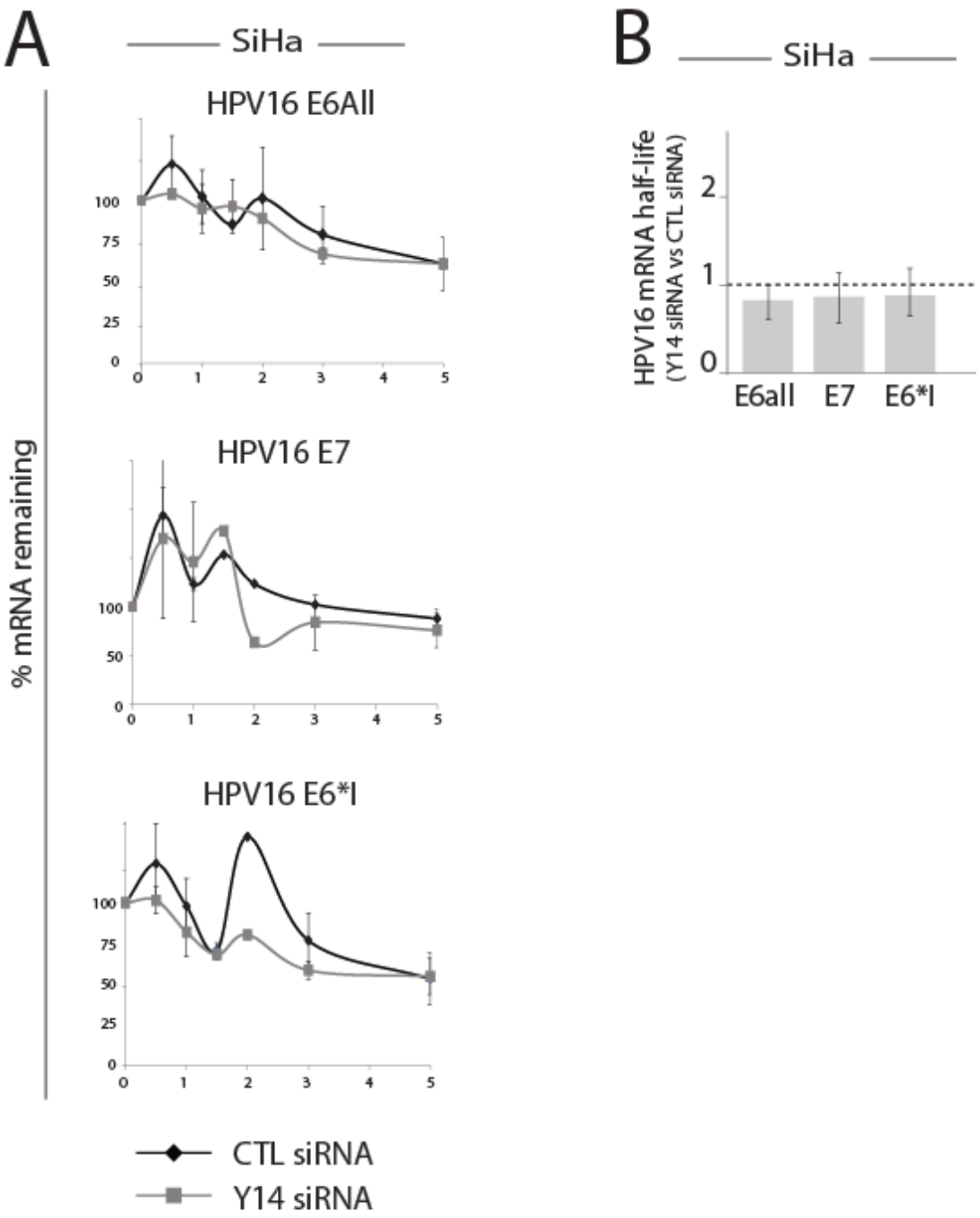

A
SiHa
HPV16 E6All
HPV16 E7
HPV16 E6*I
% mRNA remaining
CTL siRNA
Y14 siRNA
B
SiHa
HPV16 mRNA half-life
(Y14 siRNA vs CTL siRNA)
E6all
E7
E6*I

Fig S2

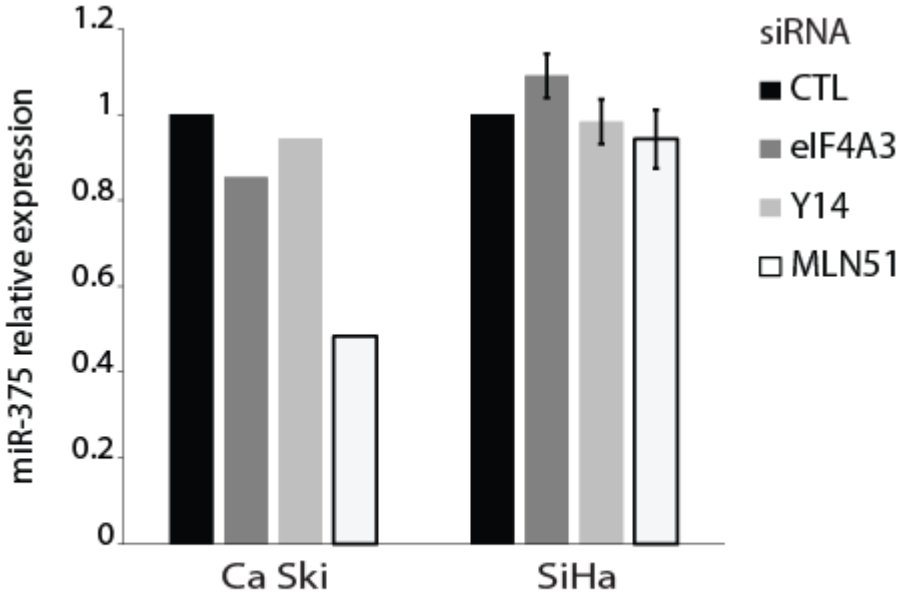

1.2
1
0.8
0.6
0.4
0.2
0
miR-375 relative expression
Ca Ski
SiHa
siRNA
CTL
eIF4A3
Y14
MLN51

Fig S3

737

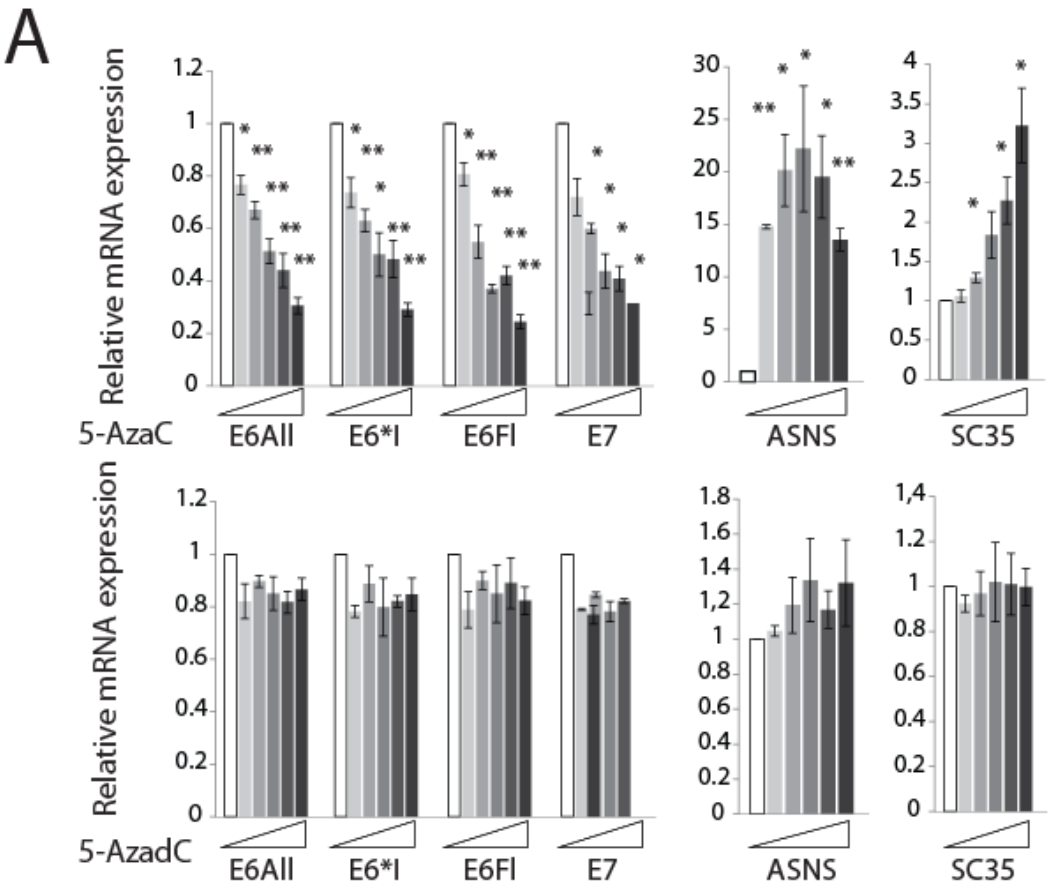

A
Relative mRNA expression
5-AzaC
E6AII
E6*I
E6FI
E7
ASNS
SC35
5-AzadC
B
5AzaC (µM)
5AzadC (µM)
0 5 10 0 5 10
E6
E7
p53
Actin

Fig S4

Supplementary Table 1.

siRNA	Target sequence
eIF4A3-63-	AGACATGACTAAAGTGGAA
Y14	CGCUCUGUUGAAGGCUGGA
MLN51	GAUCGGAAGAAUCCAGCAU
UPF1	GAUGCAGUUCCGCUCCAUU
SMG1	GUGUAUGUGCGCCAAAGUA

Supplementary Table 2.

Primer	Sequence	Use
Q5MutHPV16E6SD226_F	GCTATATGACTTTGCTTTTCGGG	Mutation of SD226 from E6 sequence
Q5MutHPV16E6SA226_R	TCGCGTCGCAGTAACTGTTGCTTG	
Q5MutHPV16E6*I_F2	GTGTATTAACTGTCAAAAGCC	Deletion of nucleotide 227-408 from pXJ39-16E6= Production of E6*I vector
Q5MutHPV16E6*I_R2	CTCACGTCGCAGTAACTG	
Q5dbgloiII_R	CCTGAAGTTCTCAGGATCG	Deletion of rabbit βglobin intron from pXJ40 vector = Production of pXJ39
Q5dbgloiII_F	CTCCTGGGCAACGTG	

Supplementary Table 3.

Primer	Forward sequence (5'-3')	Reverse sequence (5'-3')	Use
16E6all	GCAATGTTTCAGGACCCACA	TTGTTTGCAGCTCTGTGCAT	qPCR
16E6Fl	TGACTTTGCTTTTCGGGATTT	ACAGCATATGGATTCCCATCTC	qPCR
16E6*I	CTGCGACGTGAGGTGTATTA	TGTCCAGGTGTCTTTGCTTT	qPCR
16E7	ACAAGCAGAACCGGACAGAG	GCCCATTAACAGGTCTTCCA	qPCR and PCR
16E6^E7	CACAGGAGCGACCCAGAAA	AGTCACACTCACGTCGCAGTAA	qPCR
rRNA18S	GCAATTATTCCCCATGAACG	GGCCTCACTAAACCATCCAA	qPCR and PCR
eIF4A3	ACGAGCAATCAAGCAGATCA	AGGTGGCTGTTTTTCCTGTG	qPCR
RBM8A(Y14)	AAGATTTCGCCATGGATGAG	TTCCGTTTCTTCGCTTTTTC	qPCR
Magoh	TTCAGGCTCGGTTGTCTTTT	TCCGGTCGAAACTCAAACTC	qPCR
MLN51	AGCCTGACACCAAAAGCACT	TAGGGCCCTTTTTACCCACT	qPCR
SC35_1.6kb	CGGTGTCCTCTTAAGAAAATGATGTA	CTGCTACACAACTGCGCCTTTT	qPCR
ASNS	GGAAGACAGCCCCGATTTACT	AGCACGAACTGTTGTAATGTCA	qPCR
SMG1	CAAGCGATGTCAGCAGATGT	TGCTGACAAAAGCCATTCAG	qPCR
UPF1	CTGCACACCAAGCTCTACCA	GCACACAGGACAGGATGATG	qPCR
NRP1-F	CTCAGAATGGAGCTGCTGGG	GTTCCCGTTGGGAGTGGTC	qPCR
GABARAPL1	CCCTCCCTTGGTTATCATCCA	ACTCCCACCCCACAAAATCC	qPCR
SF3B5	ACCGCTACACCATCCATAGC	AGGCTGAAGCATCTTTTCCA	qPCR and PCR
Hist2H3C-F	TTCGAAGACACGAACCTGTG	ACTTCTTAAGCCCGCTCTCC	qPCR
B2M	GATGAGTATGCCTGCCGTGTG	CAATCCAAATGCGGCATCT	PCR
KPNA1	AGTGGTTTCTCCTGCTTTGC	GCTGATCCTCCAGAAGTTGC	PCR
E6ORF	ATGCACCAAAAGAGAACTGC (F6)	TTACAGCTGGGTTTCTCTACGTGT (R6)	PCR
E6E7ORF	ATGCACCAAAAGAGAACTGC (F6)	GCCCATTAACAGGTCTTCCA (R7)	PCR
βglobine	GAAGGCTCATGGCAAGAAAG	CACTGGTGGGGTGAATTCTT	PCR
Renil-Luciferase	ACAAGTACCTCACCGCTTGG	GACACTCTCAGCATGGACGA	PCR

Supplementary Table 4.

RT Primers	Reverse sequence (5'-3')	
miR-375	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTCACGC	
snRNA U6	AACGCTTCACGAATTTGCGT	
qPCR Primers	**Forward sequence (5'-3')**	**Reverse sequence (5'-3')**
miR-375	AGCCGTTTGTTCGTTCGGCT	GTGCAGGGTCCGAGGT
snRNA U6	CTCGCTTCGGCAGCACA	AACGCTTCACGAATTTGCGT

Supplemental material and methods

The thermal conditions for semi-quantitative PCR are described below. For E6ORF, an initial denaturation at 95°C for 2 min, followed by 35 amplification cycles (95°C for 30 sec, 50°C for 20 sec, 72°C for 30 sec) and a final elongation at 72°C for 5 min. For E7 and E6E7ORF, an initial denaturation at 95°C for 2 min, followed by 35 amplification cycles (95°C for 30 sec, 50°C for 20 sec, 72°C for 1 min) and a final elongation at 72°C for 5 min.

For rRNA18S, B2M and KPNA1, an initial denaturation at 95°C for 5 min, followed by 25 amplification cycles (95°C for 30 sec, 60°C for 20 sec, 72°C for 30 sec) and a final elongation at 60°C for 10 min. Excepting for the cycle number (35 cycles), the same conditions as for rRNA18S were used to amplify SF3B5.

3.2. Résultats supplémentaires : expression des facteurs de l'EJC et du NMD dans les lignées cancéreuses HPV16 positives

3.2.1. Introduction

Dans l'article, nous avons observé la diminution de l'activité du NMD dans les cellules HPV16+ (SiHa et Ca Ski). L'activité du NMD étant entre autre dépendante de l'expression de ses facteurs effecteurs, nous avons donc voulu appréhender si l'inhibition du NMD dans les cellules HPV16+ était liée à une modification de leur expression. Nous avons donc étudié l'expression des transcrits codant les facteurs de l'EJC (eIF4A3, Y14, MAGOH, MLN51, PYM1) et du NMD (UPF1, UPF2, UPF3a, UPF3b, SMG1), dans les cellules HPV16 + *versus* HPV16-.

3.2.2. Méthodologie

- Culture cellulaire

Les cellules C33A, SiHa et Ca Ski ont été cultivées dans un milieu EMEM, DMEM et RPMI (Lonza) respectivement. Tous les milieux ont été supplémentés avec 10% de sérum de veau fœtal (Eurobio). Les cellules ont été incubées à 37°C, dans un atmosphère humide contenant 5% de CO2.

- Extraction d'ARN et RT-PCR

Les ARN totaux ont été extraits en utilisant le réactif Ribozol (VWR chemicals) suivant les recommandations du fournisseur. La rétro-transcription des ARN a été réalisée en utilisant le kit Maxima first strand cDNA synthesis (Thermo Scientific) suivant les recommandations du fabricant. Les cDNA obtenus ont été amplifiés par PCR quantitative à l'aide du kit Power SYBR Green PCR mix (Life Technologies) et avec l'appareil 7500 Real-Time PCR system (Applied Biosystems) suivant les recommandations du fabricant. Le programme de qPCR a consisté en une dénaturation initiale à 95°C pendant 5 min at 95°C, suivie de 40 cycles d'amplification à 95°C pendant 30 sec et 60°C pendant 1 min. Le niveau de transcrit a été normalisé à celui de l'ARN ribosomale 18S, et quantifié suivant la méthode du $2^{-\Delta\Delta Ct}$. Les amorces utilisées sont listées dans le Tableau IV.

Tableau IV. Liste des amorces utilisées pour quantifier l'expression des facteurs de l'EJC et du NMD.

Primer	Forward sequence (5'-3')	Reverse sequence (5'-3')
ARN18S	GCAATTATTCCCCATGAACG	GGCCTCACTAAACCATCCAA
eIF4A3	ACGAGCAATCAAGCAGATCA	AGGTGGCTGTTTTTCCTGTG
RBM8A-Y14	AAGATTTCGCCATGGATGAG	TTCCGTTTCTTCGCTTTTTC
MAGOH	TTCAGGCTCGGTTGTCTTTT	TCCGGTCGAAACTCAAACTC
MLN51	AGCCTGACACCAAAAGCACT	TAGGGCCCTTTTTACCCACT
SMG1	CAAGCGATGTCAGCAGATGT	TGCTGACAAAAGCCATTCAG
UPF1	CTGCACACCAAGCTCTACCA	GCACACAGGACAGGATGATG
UPF2	GAAGATGAGGAAGCCAGCAC	CCCGTACCAACTTCTTCCTG
UPF3b	AAGGGGGAAGATAAGCAGGA	GTTCCTGAAGCTGCTCCTTG
UPF3a	GAAAAGAGCGACTGGCAAAC	TGCCTTGCAGATCCTCCTG
PYM1	GGCAAGTATATCGCGTCAACA	TTCAGGTTACGTTTGGCTGTC

3.2.3. Résultats

Comme illustré dans la Figure 28, aucune baisse de l'expression des facteurs impliqués dans le NMD et pouvant expliquer l'inhibition de l'activité du NMD, n'a été observée dans les cellules HPV16+ (SiHa, Ca Ski) par rapport aux cellules HPV- (C33A). Une augmentation de l'expression de certains facteurs (eIF4A3, Y14, MAGOH, UPF3a, PYM1, UPF2) a même été observée. Cette expérience n'ayant été réalisée qu'une seule fois, il faudrait confirmer ces résultats.

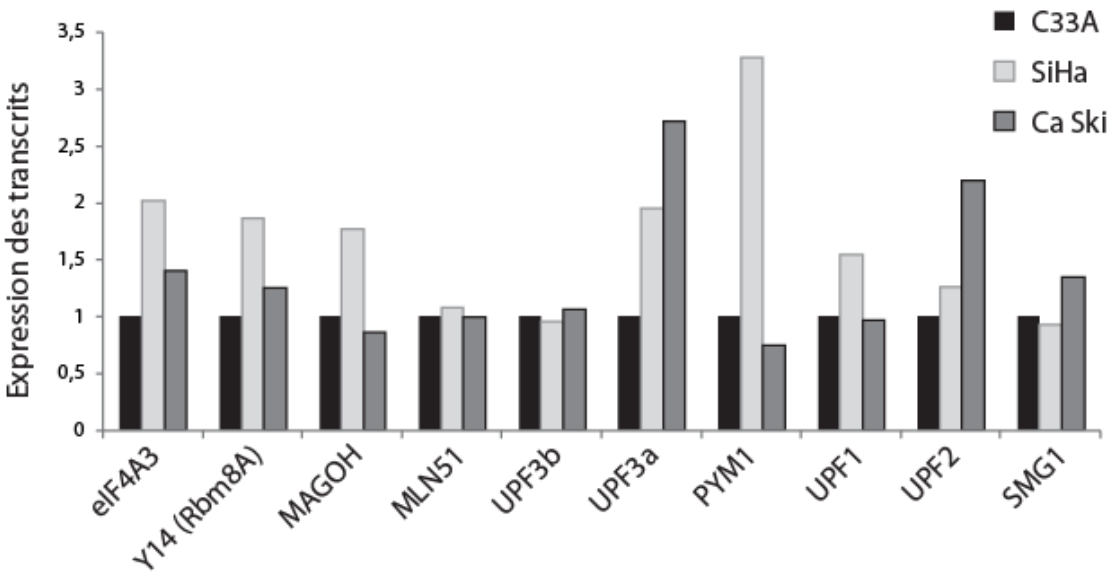


Figure 28. Expression des facteurs de l'EJC et du NMD dans les cellules HPV positives. (A) RT-qPCR montrant l'expression des transcrit codant pour les facteurs de l'EJC et du NMD dans les cellules C33A, SiHa et Ca Ski. L'ARNr ribosomale 18S a été utilisé pour la normalisation. L'expression des transcrits a été analysée suivant la méthode du $2^{-\Delta\Delta Ct}$, en utilisant l'échantillon C33A comme référence.

3.2.4. Conclusions

L'analyse de l'expression des facteurs de l'EJC et du NMD, au niveau transcriptomique, n'a montré aucun changement de leur expression dans les cellules HPV16+, pouvant indiquer le mécanisme de l'inhibition du NMD dans ces cellules. En outre, nous n'avons pas trouvé d'altération de l'expression de SMG1, dont la transcription était diminuée dans les cancers des VADS HPV+ par rapport au cancers HPV- (Gubanova *et al*., 2012). L'inhibition du NMD dans les cellules SiHa et Ca Ski pourrait être liée à une altération de l'expression de facteurs non analysés, comme SMG5, SMG6 et SMG7. L'inhibition du NMD dans ces cellules pourrait aussi résulter de l'altération de l'expression/activité des protéines du NMD. Cela ne serait pas surprenant au regard des propriétés d'E6 à altérer l'expression et l'activité de nombreuses protéines. Une analyse de l'expression de ces facteurs effecteurs, au niveau protéique, pourrait apporter les informations sur les mécanismes d'inhibition du NMD par HPV16.

4. Régulation de l'épissage d'E6 : Identifier des mutations au niveau des ORF E6 et E7 ayant un impact sur l'épissage

4.1. Introduction

L'épissage au sein des ORF E6 et E7 caractéristique des HPV-HR, permettrait entre autre l'expression de l'isoforme E6*I qui présente plusieurs activités biologiques, ainsi que le contrôle de l'expression des protéines E6 et E7. L'activité biologique d'E6 et E6*I pouvant dépendre de leur concentration cellulaire, l'épissage pourrait affecter ces fonctions, voire le phénotype des tumeurs. L'épissage est particulièrement régulé par la présence d'éléments *cis*-régulateurs au sein des ARNm, mais aussi par les niveaux d'expression des facteurs *trans*-régulateurs qui y sont recrutés. Nous avons voulu étudier les profils d'épissage des ORF E6 et E7 dans des cancers des VADS et les comparer avec le profil observé dans les cancers du col de l'utérus, afin d'identifier des changements de l'épissage pouvant avoir un rôle dans le cancer. Nous avons aussi prévu d'appréhender les mécanismes à l'origine de changements d'épissage observés et ce, en tentant d'identifier des mutations affectant l'épissage des ARNm E6E7. Les mutations altérant l'épissage d'E6Fl et E6*I permettraient d'appréhender les mécanismes de régulation de l'épissage de l'ORF E6, notamment les éléments ESE et ESS potentiellement impliqués.

4.2. Méthodologie

Les ARN totaux et l'ADN extraits de biopsies de cancers des VADS ont été fournis par l'équipe du Dr. Alain Jung (Centre Paul Strauss, Strasbourg). Les ARN totaux ont été extraits en utilisant le kit RNeasy (Qiagen) alors que l'ADN a été extrait à l'aide du kit Manual-MagNa Pure LC DNA II kit (Roche) selon les recommandations des fournisseurs respectifs. L'ADN et l'ARN de la lignée cellulaire Ca Ski dérivant d'un cancer du col de l'utérus ont été extraits avec le kit Nucleospin DNA (Macherey-Nagel) et le Ribozol (VWR chemicals) selon les recommandations de chaque fournisseur respectivement.

Les cDNA ont été synthétisés à partir de 100 ng d'ARN totaux en utilisant le kit Maxima First Strand synthesis (Thermo) selon les recommandations du fournisseur. Les cDNA ont été ensuite amplifiés par PCR, en utilisant les amorces E6orf-F (ATGCACCAAAAGAGAACT-GC) et E6orf-R (TTACAGCTGGGTTTCTCTACGTGT), à l'aide du kit DreamTaq DNA pol (Thermo) suivant les recommandations du fabricant. Le programme de PCR a consisté en : (i) une dénaturation initiale à 95°C pendant 5 min, (ii) 35 cycles d'amplification à 95°C pendant 30 sec, 50°C pendant 20 sec et 72°C pendant 30 sec suivis (iii) d'une élongation finale à 72°C

pendant 5 min. Les produits de PCR ont été analysés par une électrophorèse sur un gel d'agarose 2% (m/v) dans un tampon TBE. Les produits ont été visualisés aux ultraviolets à l'aide du système ChemiDoc XRS (BioRad) et l'intensité des bandes quantifiée à l'aide du logiciel Image Lab (Biorad).

Afin de séquencer les ORF E6 et E7 des échantillons sélectionnés, l'ADN viral a été amplifié en utilisant les amorces LCRseq-F (AAACTGCACATGGGTGTGTG) et E7-R (GCCCATT-AACAGGTCTTCCA), à l'aide du kit kit Phusion High-Fidelity PCR Kit (New England Biolabs) suivant les recommandations du fabricant. Le programme de PCR a consisté en : (i) une dénaturation initiale à 98°C pendant 30sec, (ii) 30 cycles d'amplification à 98°C pendant 10 sec, 64°C pendant 20 sec et 72°C pendant 20 sec suivis (iii) d'une élongation finale à 72°C pendant 5 min. Après la séparation des produits de PCR par une électrophorèse sur un gel d'agarose 2%, les bandes d'intérêt ont été découpées et extraites du gel en utilisant le kit NucleoSpin Gel and PCR Clean-up (Macherey-Nagel) suivant le recommandations du fournisseur. Les bandes ont été séquencées par la méthode de Sanger et les séquences ont été comparées au génome de référence GenBank: K02718.1 afin d'identifier des mutations potentielles.

Les prédictions in silico de l'effet des mutations sur l'épissage ont été réalisées par SpliceAid2 (http://193.206.120.249/splicing_tissue.html) et Human Splicing Finder 3.1 (http://www.umd.be/HSF3/HSF.shtml).

4.3. Résultats

Afin d'étudier les profils d'épissage de l'ORF E6 dans des cancers VADS, nous avons amplifié les ARNm E6 par une RT-PCR semi-quantitative à partir d'ARN totaux extraits de tumeurs des VADS HPV16 positives et de cellules CaSki.

Comme illustré dans la Figure 29, l'épissage alternatif au sein de l'ORF E6 génère des transcrits non épissés (E6Fl) et épissés (E6*I et E6*II). Comme attendu, E6*I représente les transcrits majoritaires dans les échantillons de patients et dans les cellules Ca Ski (Fig. 29). Par contre, les profils d'épissage varient considérablement entre les échantillons, présentant de fait différentes proportions d'E6Fl, E6*I et E6*II (Fig. 29). En particulier, les taux d'E6Fl augmentent dans les échantillons 225, 855, 938, 1359, 1366, 1426, 1427 et 2301 en comparaison avec ceux observés dans les cellules Ca Ski (Fig. 29). Dans les échantillons 635, 825, 985, 1461, les taux d'E6Fl ne varient pas en comparaison à ceux observés dans les cellules Ca Ski, alors qu'ils baissent dans les échantillons, 801, 882, 1237, 1319, 1405, 1630,

1635, 1682, 1869, 2311. Enfin, aucune amplification des ARNm E6 d'HPV16 n'a été observée dans les échantillons 1436 et 1542, issus de cancer HPV16 négatifs.

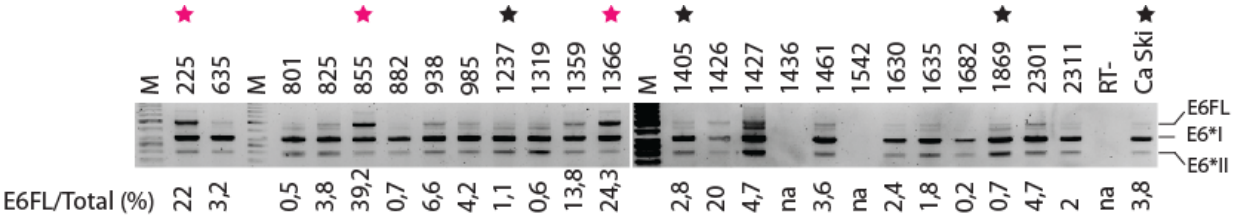


Figure 29. Profils d'épissage des ARNm E6 dans les échantillons de cancers des VADS.

Afin de déterminer si le changement des profils d'épissage était lié à l'altération d'ESE et/ou d'ESS, nous avons sélectionné les échantillons 225, 855, 1366 et Ca Ski, 1869 1237, 1405, (22%, 39,2%, 24,3% d'E6Fl) et procédé au séquençage de leurs séquences E6 et E7.

Après séquençage et comparaison avec la séquence de référence d'HPV16 du lignage européen A1 (GenBank: K02718), plusieurs variations nucléotidiques ont été identifiées dans les échantillons analysés (Tableau V). De façon intéressante, des mutations (173C>T et 233G>A) ont été retrouvées dans les échantillons ayant présenté un profil différentiel de l'épissage de l'ORF E6 (225 et 855 respectivement) (Tableau V). Ces mutations pourraient avoir induit un changement d'épissage dans l'ORF E6, en altérant potentiellement des ESE et/ou des ESS. Les mutations en position 131A>G et 350T>G, identifiées dans les échantillons 225, 855 et Ca Ski, sont caractéristiques du lignage européen A2 (GenBank: AF536179). La présence de ces mutations dans les cellules Ca Ski suggère qu'elles ne soient pas à l'origine de la variation des profils d'épissage observé dans les échantillons 225 et 855.

L'échantillon 1869, dont l'épissage induit une baisse des transcrits E6Fl, présente 2 mutations : 350T>G observée dans de nombreux lignages et 109T>C qui pourrait être impliquée dans le changement du profil d'épissage observé chez ce patient. Enfin, la séquence des échantillons 1366 et 1405 est identique à celle de la séquence de référence HPV16 de lignage européen A1. Le profil d'épissage particulier observé dans l'échantillon 1366 pourrait être lié à d'autres facteurs tels le niveau d'expression des facteurs *trans*-régulateurs d'épissage.

Tableau V. Variations nucléotidiques au niveau de l'ORF E6 d'HPV16 dans les échantillons de cancers des VADS.

	Nucléotide						
sample	**109**	**131**	**173**	**233**	**315**	**350**	
HPV16 A1	T	A	C	G	C	T	
1366	-	-	-	-	-	-	Lignage A
225	-	G	T	-	-	G	Lignage A2
855	-	G	-	A	-	G	Lignage A2
CaSki	-	G	-	-	-	G	Lignage A2
1869	C	-	-	-	-	G	Lignage A
1237	-	-	-	-	G	-	Lignage A
1405	-	-	-	-	-	-	Lignage A

A l'instar de l'échantillon 1366, l'épissage des autres échantillons pourrait résulter non pas des mutations altérant les ESE et ESS, mais plutôt d'une variation de l'expression des facteurs *trans*-régulateurs d'épissage et/ou de la vitesse de transcription. Afin de confirmer l'effet potentiel de chaque mutation sur l'épissage, nous avons effectué des prédictions *in silico* en utilisant HSF3.1 (Human Splicing Finder 3.1) et SpliceAid2 (Voir matériels et méthodes). SpliceAid2 regroupe les 2339 motifs de liaisons de 71 protéines impliquées dans l'épissage (Piva *et al*., 2012). Ces motifs ont été identifiés de façon expérimentale. HSF3.1 est un outil de prédiction de l'effet des mutations sur l'épissage (Desmet *et al*., 2009). Il regroupe les nombreuses matrices permettant d'identifier des sites canoniques d'épissage, des ESE/ESS/ISE/ISS ainsi que la liaison par des facteurs d'épissage.

Comme attendu, la mutation 350T>G ne semble pas avoir d'effet sur l'épissage d'après HSF3.1, même si SpliceAid2 suggère la destruction d'un site de liaison de TIA-1 (TIA1 cytotoxic granule associated RNA binding protein) et/ou TIAL-1 (TIA1 cytotoxic granule associated RNA binding protein like 1) et son remplacement par un site ELAVL2 (ELAV like RNA binding protein 2) (Tableau VI). En revanche, les prédictions de HSF3.1 suggèrent que la mutation 315C>G pourrait induire la création d'un ESS, tandis que d'après SpliceAid2, elle induit la destruction de sites reconnus par SLM-2 (KH RNA binding domain containing, signal transduction associated 3) et Sam68 et la création d'un site de liaison d'ELAVL2 (Tableau VI).

La mutation 233G>A, pourrait induire des modifications d'épissage en altérant un ESE, notamment un site pour le facteur 9G8 (SRSF7) (Tableau VI). Selon SpliceAid2, elle crée des sites de liaison de SLM-2 et Sam68 pouvant inhiber le site SD226. Du fait de sa proximité avec le site SD226, la liaison de RBP pourrait inhiber l'épissage en induisant un

encombrement stérique, empêchant ainsi la fixation de snRNP U1. La mutation 173C>T altère aussi un ESE contenant des sites de liaison de 9G8 et crée un ESS permettant la liaison de hnRNP A1, pouvant inhiber l'épissage et favoriser la production de la forme E6Fl.

Tableau VI. Prédictions *in silico* de l'impact des mutations identifiées au sein de l'ORF E6, sur l'épissage.

	109T>C	131A>G	173C>T	233G>A	315C>G	350T>G
SpliceAid 2	Destruction de sites ELAVL2, TIA-1, TIAL-2 et Nova-1.	Création de site pour SF2/ASF et SC35.	Pas de changement.	Création de site pour SLM-2 et Sam68.	Destruction de sites SLM-2 et Sam68. Création de site pour ETR-3, Nova-1 et CUG-BP1.	Destruction de sites TIA-1 et TIAL-1. Création de site pour ELAVL2.
Human Splicing Finder 3.1	Création d'un ESS.	Altération d'un ESE (9G8).	Création d'un ESS (hnRNP A1). Altéraction d'un ESE.	Altéraction d'un ESE (9G8).	Création d'un ESS.	Pas d'effet.

Par ailleurs, bien qu'elle ne soit pas associée à un changement d'épissage, le mutation 131A>G est prédite comme pouvant altérer un ESE et des sites de liaison de 9G8, SF2/ASF et SC35 (Tableau VI). Quant à la mutation 109T>C, elle pourrait créer un ESS et /ou détruire des sites de liaison d'ELAVL2, TIA-1, TIAL-2 et Nova-1 (Neuro-oncological ventral antigen 1).

4.4. Conclusions

L'étude des profils d'épissage de l'ORF E6 d'HPV16 a montré que ceux-ci étaient globalement similaires entre les différents échantillons de cancers VADS et de lignées dérivées de cancer du col de l'utérus. En effet, le produit majoritaire dans les différents échantillons est le plus souvent E6*I, soulignant au passage l'importance de ce produit. Nous avons aussi observé des modifications du profil d'épissage dans certains échantillons, notamment une augmentation des ARNm E6Fl. Ces modifications de l'épissage pouvent être liées à un changement d'éléments *cis*-régulateurs et/ou de l'expression des facteurs *trans*-régulateurs qui y sont recrutés. La recherche de variations d'éléments *cis*-régulateurs, nous a permis d'identifier plusieurs mutations, plus particulièrement les mutations 173C>T et 233G>A, identifiés dans les échantillons 225 et 855, qui présentent un profil d'épissage particulier. Les prédictions *in silico* ont montré que ces deux mutations altèrent des éléments *cis*-régulateurs de l'épissage, et pourraient donc être à l'origine de changement d'épissage. Les données de HSF3.1 ont été obtenues suite à des études sur l'épissage d'exon cassettes, alors que l'épissage du premier intron d'E6 pourrait être réglé différemment. Il faut donc prendre prudemment ces prédictions en ce qui concerne l'effet potentiel sur l'épissage d'E6. Par conséquent, afin de déterminer l'effet biologique de ces mutations sur l'épissage, il est nécessaire d'effectuer des expériences de transfection *in cellulo*.

Discussion et perspectives

L'infection par les HPV conduit à des défauts de prolifération cellulaire, pouvant parfois induire l'apparition de lésions cancéreuses. En particulier, l'infection par les HPV-HR est responsable de l'apparition de nombreux cancers, tandis que l'infection par les HPV-BR est uniquement liée à l'apparition de lésions bénignes. La carcinogenèse induite par les HPV-HR, notamment au niveau du col de l'utérus est un processus lent, qui dure plusieurs années. Au cours de cet intervalle, différentes lésions précancéreuses ont été identifiées, certaines lésions étant éliminées et d'autres progressant vers le stade cancer. Le dépistage et la prise en charge des lésions précancéreuses nécessitent une bonne connaissance de la carcinogenèse induite par les HPV-HR. Celle-ci est liée à l'expression accrue des oncoprotéines E6 et E7, lesquelles dérégulent de nombreux processus cellulaires.

Afin de mieux appréhender les mécanismes de la carcinogenèse induite par les HPV-HR, nous nous sommes intéressés d'une part, à l'étude des fonctions des oncogènes d'HPV16, notamment du gène E6 et d'autre part, aux mécanismes impliqués dans la régulation de l'expression d'E6 et E7.

1. Fonctions spécifiques du gène E6 des HPV-HR

Dans notre projet, nous nous sommes intéressés aux différences fonctionnelles des gènes E6 et E7, entre les HPV-HR et les HPV-BR. Plusieurs différences ont été décrites pour les activités des protéines E6 et E7, entre les HPV-HR et HPV-BR (Roman et Munger, 2013; Vande Pol et Klingelhutz, 2013; Wallace et Galloway, 2015). Une des différences les plus intrigante, est l'existence de sites d'épissage dans les ORF E6 et E7 chez les HPV-HR, mais pas chez les HPV-BR (De la Cruz-Hernández *et al*., 2005; Mesplède *et al*., 2012; Smotkin *et al*., 1989). Ces sites permettent l'expression d'ARNm E6*I, largement majoritaires par rapport à E6Fl (qui codent la protéine E6). La fonction de cet épissage a été largement débattue. D'une part, il générerait des ARNm E6*I-E7 qui permettrait la synthèse de la protéine E7 (Ajiro *et al*., 2012; Tan *et al*., 1994b; Tang *et al*., 2006; Zheng *et al*., 2004). D'autre part, les transcrits E6*I comporteraient des fonctions spécifiques, permettant entre autre l'expression d'une protéine E6*I (Pim *et al*., 1997; Schneider-Gädicke *et al*., 1988). Afin d'appréhender le rôle spécifique des ARN/protéines E6*I, nous avons étudié l'effet d'une expression ectopique d'E6*I dans des cellules U2OS (HPV-).

Nous avons généré des clones U2OS exprimant stablement les ARNm et les protéines E6 et/ou E6*I, générant ainsi des clones de cellules U2OS : U-E6All, exprimant E6Fl, E6*I et E6*II ; U-E6*I, exprimant uniquement E6*I ; et U-CTL, exprimant le vecteur vide. Puis, nous

avons étudié leurs transcriptomes respectifs à l'aide du RNA-seq, afin d'identifier des cibles dérégulées par l'expression ectopique d'E6 et/ou d'E6*I d'HPV16. Bien que les cellules U2OS soient mésenchymateuses, elles permettent l'expression des gènes viraux et l'amplification des génomes des HPV-HR et –BR (Geimanen *et al*., 2011; Isok-Paas *et al*., 2015; Toots *et al*., 2014). L'intérêt de ce modèle réside aussi dans sa possession de gènes p53 et pRb sauvages, faisant de ces cellules un modèle idéal pour l'étude du rôle des protéines virales E6 et E7. Nous avons ainsi pu confirmer l'expression de protéines E6 fonctionnelles dans les clones U-E6All, étant donné la baisse de l'expression de protéine p53. En revanche, bien que des ARNm E6*I soient exprimés dans les clones U-E6*I, l'expression de la protéine E6*I n'a pu être détectée. La protéine E6*I a une demi-vie très courte, étant uniquement détectée lors de la transfection transitoire, en présence d'inhibiteur du protéasome (MG132). Cela est en accord avec des études ultérieures ayant détecté E6*I en transfection transitoire mais pas lors de l'expression stable (Filippova *et al*., 2009). En outre, lors de ces études, c'est une protéine de fusion E6*I qui a été utilisée, tandis que dans notre étude nous l'avons exprimée dans sa forme native. Il faut tout de même méditer sur le rôle connu des ARN non codants sur l'expression génique, et donc du rôle potentiel des ARN E6*I (en lieu et place de la protéine) dans les effets observés dans les clones U-E6*I. Cela serait d'ailleurs une information très intéressante pour l'appréhension des mécanismes impliqués dans le cycle viral et dans la carcinogenèse induite par les HPV. En revanche, en dépit de notre tentative de générer des cellules exprimant uniquement les ARNm E6Fl seules, à partir de plasmide E6 dont le site SD226 a été muté, aucun clone n'a pu être produit. Ces observations suggèrent que l'expression excessive de la protéine E6 peut être cytotoxique pour la cellule. D'ailleurs, le seul clone ayant survécu dans cette condition, exprime une protéine E6 incapable de dégrader la protéine p53, étant probablement non fonctionnelle (Données non montrées). La cytotoxicité d'E6 pourrait être liée à sa capacité à former des agrégats lorsqu'elle est présente à de fortes concentrations (Zanier *et al*., 2007). Toutefois, dans certaines études il a été possible de produire de tels clones, en transfectant des cellules épithéliales. Une toxicité cellule-dépendante demeure donc une possibilité non négligeable.

Nous avons ensuite caractérisé les propriétés biologiques des clones transfectées stablement à l'aide de plusieurs tests fonctionnels. Comme décrit dans des études précédentes, nous n'avons observé aucun effet de l'expression d'E6 et E6*I ou E6*I seule sur la prolifération cellulaire (Pim *et al*., 1997; Sedman *et al*., 1991). En revanche, l'expression d'E6 réduit la capacité clonogénique des cellules transfectées comme décrit par Ristriani et collaborateurs

(Ristriani *et al*., 2009). Nos résultats montrent qu'E6*I inhibe plus fortement la clonogénicité des cellules U2OS. Ces résultats sont en adéquation avec ceux de Filippova et collaborateurs, qui montrent un rôle anti-oncogénique d'E6*I, inhibant la croissance tumorale dans un modèle de greffe de souris nude (Filippova *et al*., 2014).

De plus, contrairement aux cellules U-E6*I et U-CTL, les cellules U-E6All exprimant les isoformes d'E6, présentent une augmentation de leur radiosensibilité. Ces résultats sont en adéquation avec ceux montrant l'augmentation de radiosensibilité des cellules cancéreuses HPV-positives, exprimant la protéine E6 (Arenz *et al*., 2014; Mirghani *et al*., 2015; Pang *et al*., 2011; Rieckmann *et al*., 2013). Ces résultats seraient liés à l'inhibition, par la protéine E6, des systèmes de réparation de l'ADN médiés par p53 (Kessis *et al*., 1993). En effet, en inhibant plusieurs facteurs impliqués dans la réparation de l'ADN, comme XRCC1 et RAD51, E6 mais pas E6*I, empêche la réparation des dommages de l'ADN (Iftner *et al*., 2002; Nickson *et al*., 2017; Wallace *et al*., 2017). L'accumulation des dégâts sur l'ADN contribue à induire la mort des cellules irradiées. Contrairement à Pang et collaborateurs, qui ont observé une augmentation de la radiosensibilité des cellules exprimant E6*I, nous n'avons observé aucun effet d'E6*I sur la radiosensibilité des cellules U2OS (Pang *et al*., 2011). Cette différence pourrait être liée à la nature du modèle cellulaire utilisé, notamment au statut du gène TP53. En effet, alors que nous avons utilisé des cellules d'ostéosarcome (U2OS ; TP53wt), Pang et collaborateurs ont utilisé des cellules de cancers VADS (TP53 muté) (Pang *et al*., 2011).

Grâce au RNA-seq, nous avons identifié 419 transcrits dérégulés par l'expression d'E6 et E6*I dans les clones U-E6All. La plupart de ces transcrits ont été décrits comme dérégulés dans des modèles cellulaires d'HPV16 ou dans des échantillons cliniques de cancers HPV16 positifs. De façon intéressante, des transcrits codant des protéines ribosomales dont l'expression est inhibée dans des cancers du col de l'utérus, ont aussi été dérégulés dans les clones U-E6All (den Boon *et al*., 2015). On retrouve aussi des transcrits de gènes impliqués dans des voies relatives à l'expression des gènes viraux, la carcinogenèse virale, la transduction du signal et la traduction. Nous avons par ailleurs confirmé l'augmentation de l'expression de ces transcrits, comme SYCP2 (Synaptonemal complex protein 2), Nup210 (Nucleoporin 210) et DICER1 (Dicer 1, double-stranded RNA-specific endoribonuclease), dont l'expression a été aussi augmentée dans des cancers du col de l'utérus et des VADS, mais aussi dans des cellules exprimant E6 et E7 de façon ectopique (den Boon *et al*., 2015; Harden et Munger, 2017; Pyeon *et al*., 2007; Slebos *et al*., 2006). Nos résultats suggèrent

donc que c'est l'expression d'E6 (et/ou E6*I) qui serait à l'origine de la dérégulation de leur expression. En fait, cette dérégulation de l'expression génique serait majoritairement due à l'expression d'E6. L'effet d'E6 sur l'expression génique, notamment au niveau transcriptionnel, est d'ailleurs assez bien décrit (Durzynska *et al*., 2017; Wallace et Galloway, 2015). E6 altère le transcriptome de façon indirecte via la modulation de nombreuses protéines impliquées dans la signalisation cellulaire, mais aussi directement en modulant les facteurs de transcription et de remodelage de la chromatine (Durzynska *et al*., 2017; Johannsen et Lambert, 2013). A noter que l'effet d'E6 sur l'expression génique pourrait être aussi lié à son interaction avec l'ADN et les ARN, mais aussi à son impact sur le Nonsense-mediated mRNA decay comme nous l'avons montré (Voir discussion ci-dessous).

Les analyses transcriptomiques par RNA-seq ont aussi montré que l'expression d'E6*I seule affectait l'expression génique, quoique son effet soit moins fort que celui de la protéine E6. En effet, l'expression ectopique d'E6*I module l'expression de 53 transcrits, dont uniquement 11 sont communément dérégulés par E6 et E6*I. La confirmation de ces résultats par RT-qPCR a d'ailleurs montré l'existence d'une forte corrélation des variations de ces transcrits entre le RNA-seq et la RT-qPCR. Les gènes dont les transcrits sont dérégulés par E6*I sont impliqués dans des voies du cancer tels que : l'organisation de la matrice extracellulaire, l'adhérence cellulaire, la liaison et la réponse aux facteurs de croissance et des cytokines, ainsi que dans des voies de signalisation cellulaire (MAPK). Ces gènes sont aussi impliqués dans de métabolisme cellulaire, la réponse inflammatoire et la signalisation immunitaire, des voies dérégulées par HPV16 (Klymenko *et al*., 2017). Ces résultats indiquent qu'E6*I peut agir au niveau cellulaire, indépendamment d'E6, ayant donc potentiellement un rôle dans les effets d'HPV16. Il faut noter que la plupart des études sur l'effet d'E6 au niveau cellulaire a lieu avec l'expression commune d'E6 et E6*I, ne distinguant donc pas l'effet d'E6*I de celui d'E6. Le peu de transcrits dérégulés en commun dans U-E6All et U-E6*I pourrait être lié à l'expression d'E6 qui contrecarrerait l'effet d'E6*I. En effet, E6*I interagit avec E6 et inhibe sa capacité à dégrader p53 (Pim et Banks, 1999; Pim *et al*., 1997), et l'inverse serait aussi probable. De plus, il a été observé que l'effet d'E6 et E6*I, sur la réponse à la mort cellulaire induite par le TNF-α, pourrait varier selon le ratio entre leurs expressions respectives (Filippova *et al*., 2009). Enfin, la capacité d'E6*I à induire les ROS (Voir ci-dessous), dépend aussi du ratio E6/E6*I (Williams *et al*., 2014). La modulation de ce ratio en inhibant E6Fl ou en augmentant l'expression d'E6*I, induit les ROS dans les cellules HPV16+ (Williams *et al*.,

2014). Le ratio E6/E6*I est donc important pour réguler le rôle de chacune des isofomes d'E6, et il serait intéressant d'appréhender plus profondément son fonctionnement.

De façon intéressante, plusieurs gènes dérégulés par l'expression d'E6*I sont impliqués dans la réponse à l'hypoxie, laquelle est affectée entre autres par les ROS. En effet, les ROS sont impliqués dans la régulation de la prolifération cellulaire, l'apoptose, la communication et la signalisation médiée par les cytokines, chimiokines et facteurs de croissance (Reuter *et al*., 2010). De fait, le niveau de ROS est régulé notamment par un système anti-oxydatif, permettant de contrôler le rôle des ROS mais aussi de limiter son pouvoir pathogène, lié au stress oxydatif qu'il engendrerait. Ce système de défense anti-oxydatif participe donc au maintien de l'homéostasie cellulaire. Les virus qui dérégulent plusieurs voies cellulaires, modulent le niveau de ROS et induisent un stress oxydatif, permettant d'optimiser leur amplifications virales (Paracha *et al*., 2013; Reshi *et al*., 2014). D'ailleurs, des études sur HPV16 ont montré un lien entre l'expression d'E6*I et une augmentation des ROS (Evans *et al*., 2016; Mouret *et al*., 2005; Williams *et al*., 2014). Une expression élevée d'E6*I inhibe l'expression des protéines du système de défense anti-oxydatif, comme SOD et GPX1 (Evans *et al*., 2016; Mouret *et al*., 2005; Williams *et al*., 2014). Dans notre étude, l'expression des transcrits SOD et GPX1 n'a pas été modifié par l'expression d'E6*I, suggérant plutôt un effet de cette dernière au niveau protéique. Toutefois, plusieurs gènes dérégulés par E6*I sont impliqués dans le métabolisme des ROS : A2M (Alpha-2-macroglobulin) (Chen *et al*., 2014b), MMP3 (Matrix metallopeptidase 3) (Choi *et al*., 2014; Radisky *et al*., 2005), CRYAB f(Crystallin alpha B) (van de Schootbrugge *et al*., 2014), CCL2 (Li *et al*., 2014), CTGF (Connective tissue growth factor) (Murphy-Marshman *et al*., 2017), PDGFB (Platelet derived growth factor subunit B) (Adachi *et al*., 2005; Chen *et al*., 2007) et RAC2 (Rac family small GTPase 2) (Diebold et Bokoch, 2001; Nieborowska-Skorska *et al*., 2012). En outre, ces gènes pourraient favoriser l'invasion cellulaire, étant donné leur implication dans la progression tumorale, l'inflammation et l'organisation de matrice extracellulaire. La dérégulation de ces gènes dans des cancers HPV+ (Bhosale *et al*., 2017; Grozdanov *et al*., 2016), suggère une implication d'E6*I dans la carcinogenèse induite par les HPV-HR.

Le stress oxydatif lié à l'expression d'E6*I pourrait être impliqué dans la carcinogenèse induite par les HPV-HR, non seulement en dérégulant l'expression de nombreux gènes, mais aussi en provoquant une instabilité génétique. En effet, l'induction de certains cancers par la surexpression d'oncogènes cellulaires, tel MYC et KRAS, est réalisée *via* l'augmentation des ROS (Gao *et al*., 2007; Schieber et Chandel, 2014; Weinberg *et al*., 2010). Les ROS

permettent d'activer les voies de signalisation, entre autres PI3K et MAPK, assurant la promotion de la tumorigenèse. De plus, la prévention de la carcinogenèse par p53 serait aussi liée à sa capacité à baisser les ROS (Sablina *et al*., 2005). L'expression de gènes antioxydants est d'ailleurs régulée par p53 (Budanov, 2014; Budanov *et al*., 2004). Le stress oxydatif induit par l'expression d'E6*I pourrait aussi, *via* l'instabilité génétique engendrée, favoriser l'intégration du génome viral à celui de la cellule hôte (Camini *et al*., 2017; Williams *et al*., 2011). D'ailleurs, l'expression d'E6*I augmente la proportion de l'intégration de génome étranger au génome cellulaire ((Chen Wongworawat *et al*., 2016) ; Nos travaux, observations non montrées). Cela est en adéquation avec l'intégration du génome viral observée chez les HPV-HR, qui expriment E6*I, mais pas chez les HPV-BR, qui n'en expriment pas.

L'épissage alternatif dans les ORF E6 et E7, permet aux HPV-HR de réguler l'expression des protéines E6 et E7 (Ajiro *et al*., 2012; Tan *et al*., 1994b; Tang *et al*., 2006; Zheng *et al*., 2004) et/ou d'exprimer des ARN et des protéines tronquées fonctionnelles (Czech et Hannon, 2011; Pim et Banks, 1999; Schneider-Gädicke *et al*., 1988). De tels ARN/protéines, comme E6*I, permettent de contrôler/limiter l'expression de la protéine E6 mais aussi son activité (Pim *et al*., 1997 ; Pim *et al*., 1999 ; Filipovva *et al*., 2009) (Fig. 30). Des propriétés pro-oncogéniques ont été aussi identifiées pour E6*I, suggérant qu'elle pourrait participer à la carcinogenèse induite par les HPV-HR. En particulier, E6*I engendre les ROS, qui contribuent à l'intégration du génome viral des HPV-HR et à la carcinogenèse.

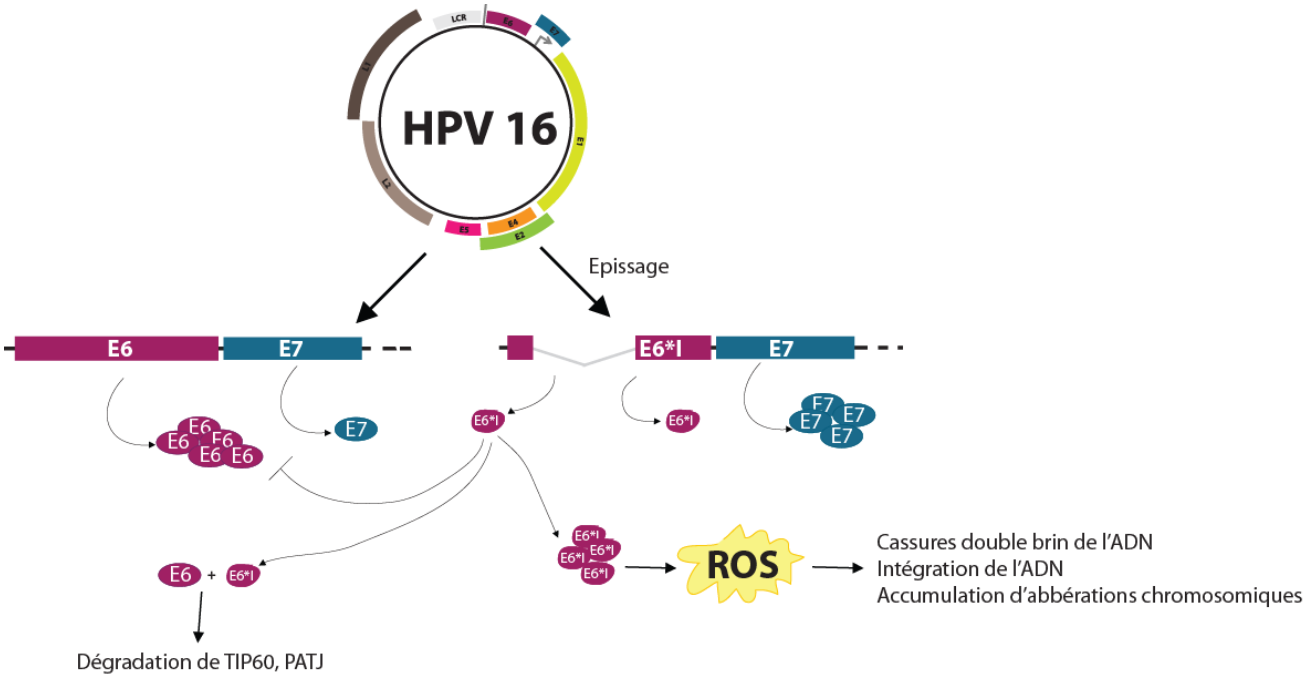


Figure 30. Le rôle de l'épissage dans l'ORF E6 des HPV-HR.

La régulation de l'expression d'E6 et E7, de la balance E6/E6*I, pourrait jouer un rôle dans le cycle viral, et une dérégulation pour être impliquée dans l'induction de la carcinogenèse. Le mécanisme régissant l'interaction entre E6 et E6*I, ainsi que le fonctionnement de la balance reste inconnu. Appréhender ces mécanismes serait fascinant du point de vue biologique, et pourrait apporter des pistes thérapeutiques potentielles ciblant les cancers/pré cancers associés aux HPV-HR. Cela souligne aussi l'importance de la régulation de l'épissage de l'ORF E6 sur la balance E6/E6*I.

2. La protéine E6, régulateur de l'expression des gènes au niveau post-transcriptionnel ?

Dans notre projet, nous avons aussi étudié les propriétés de liaison aux ARN d'E6. En utilisant la méthode de RNA-IP, nous avons confirmé la présence d'E6 dans des mRNP contenant ses propres ARNm. Ces résultats suggèrent qu'E6 se lie à ses ARNm, mais ne permettent pas d'affirmer qu'il s'agit d'une liaison directe. Les travaux de Bodaghi et collaborateurs qui ont montré une liaison directe d'E6 avec ses ARNm *in vitro*, suggèrent plutôt une interaction directe (Bodaghi *et al*., 2009). Bodaghi et collaborateurs ont aussi montré que seule les protéines E6 des HPV-HR avaient cette propriété de RBP, laquelle n'a pas été détectée pour les HPV-BR. La capacité de liaison d'E6 aux ARN pourrait avoir un rôle dans la carcinogenèse associée aux HPV-HR.

Du reste, appréhender l'importance que revêt cette protéine de liaison dans l'activité d'E6 et dans le cycle viral est une perspective fascinante. En particulier, rechercher des ARNm cellulaires interagissant avec la protéine E6, dont l'expression peut être modulée par cette dernière peut être intéressant. Une telle éventualité ne serait pas surprenante, vu qu'E6 affecte les ARNm de TERT (Katzenellenbogen *et al*., 2007, 2009). Une approche par RIP-sequencing serait intéressante pour identifier les mRNP contenant la protéine E6. Une première tentative préliminaire permettant de mettre au point les conditions de l'expérience a déjà été réalisée (Résultats non montrés). En complément, des études avec du CLIP-seq (Cross-linking immunoprecipitation-sequencing), permettant d'identifier les ARNm cellulaires liés à E6 de façon directe serait d'un apport considérable. Le CLIP-seq permettrait d'identifier, le cas échéant, les séquences nucléotidiques impliquées dans la liaison d'E6.

La liaison d'E6 avec les ARNm est assez palpitante, dans le sens où elle incite à porter un nouveau regard sur ses fonctions. E6 pourrait être vu comme un régulateur majeur de

l'expression génique et ce, à plusieurs échelons : ADN, ARN et protéines (Fig. 31). En effet, E6 est connu pour réguler l'expression et les fonctions de plusieurs protéines cellulaires. De plus, elle régule la transcription de plusieurs gènes, directement ou indirectement *via* la régulation de protéines impliquées dans la régulation transcriptionnelle et épigénétique (Durzynska *et al*., 2017; Wallace et Galloway, 2015). Enfin, l'interaction d'E6 avec les ARNm suggère son implication dans la régulation des ARNm. E6 affecte non seulement la transcription de TERT, mais aussi la stabilité de ses ARNm ainsi que la fonction de la protéine (Katzenellenbogen *et al*., 2007, 2009; Liu *et al*., 2009; Oh *et al*., 2001; Veldman *et al*., 2001). Il serait donc intéressant d'étudier l'impact de cette propriété RBP dans la fonction d'E6, notamment si cela permet d'affiner la régulation exercée par E6 sur l'expression génique. Comme E6, la protéine E7 est aussi impliquée dans l'altération de l'expression génique induite par les HPV-HR. De façon intéressante, le cancer est caractérisé par une dérégulation de l'expression génique qui découle d'altérations génétiques et épigénétiques (Chibon, 2013; Hanahan et Weinberg, 2011). E6 et E7 pourrait donc induire la carcinogenèse non seulement en inhibant directement des gènes suppresseurs de tumeur mais aussi *via* la dérégulation de l'expression des gènes.

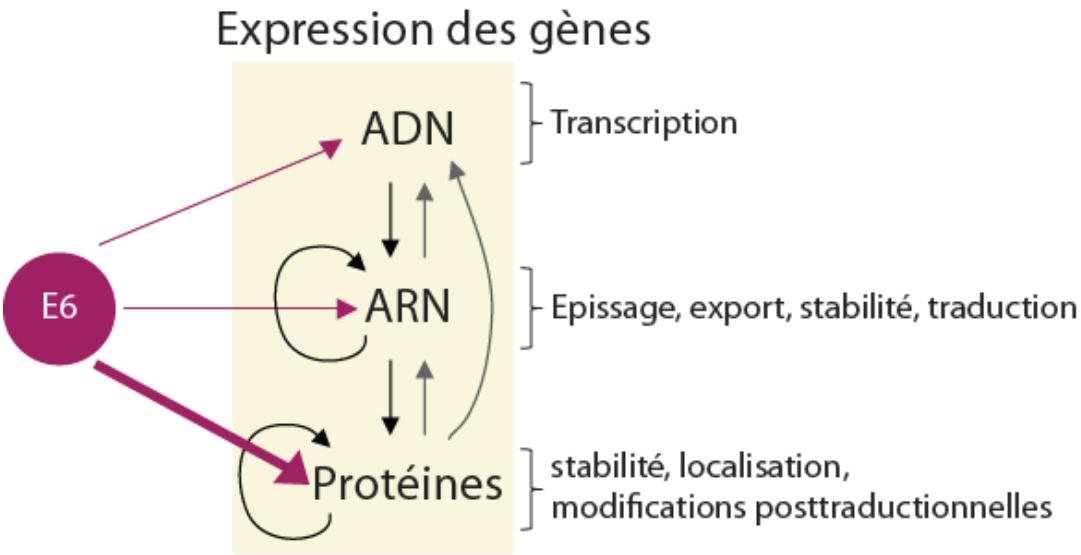


Figure 31. E6, un régulateur majeur de l'expression des gènes.

3. Régulation de l'expression des transcrits E6 et E7 par l'EJC et le NMD

L'autre versant de notre projet est l'étude des mécanismes de régulation de l'expression des protéines E6 et E7. L'expression des gènes viraux est régulée particulièrement au niveau post-transcriptionnel. A ce niveau, le devenir des ARNm est influencé par le recrutement de plusieurs protéines formant une mRNP. En outre, les transcrits polycistroniques d'HPV16 subissent un épissage alternatif, permettant ainsi de diversifier le transcriptome viral, modifier la composition du mRNP et réguler l'expression des gènes viraux. L'épissage permet entre autre le dépôt de l'EJC. L'EJC est considéré comme une plateforme essentielle à la régulation post-transcriptionnelle de l'expression des gènes, régulant l'épissage, l'export nucléo-cytoplasmique, la traduction et la stabilité des ARN. Nous avons donc voulu étudier son rôle dans la régulation de l'expression des transcrits E6 et E7.

Suite à l'inhibition de l'expression des protéines du cœur de l'EJC, nous avons observé des effets distincts des protéines de l'EJC sur l'expression d'E6 et E7. En effet, alors que l'inhibition d'eIF4A3 ou Y14 augmente l'expression de gènes viraux, l'inhibition de MLN51 induit une baisse de leur expression.

La différence de l'effet d'eIF4A3 et Y14 avec celui de MLN51 pourrait être liée soit : (i) à une composition différentielle de l'EJC et/ou (ii) à des fonctions indépendantes de l'EJC et spécifiques de chacune de ces protéines, notamment MLN51. Des études récentes ont montré que MLN51 n'était pas présente dans tous les EJC (Hauer *et al*., 2016; Saulière *et al*., 2012; Singh *et al*., 2012). En effet, eIF4A3, Y14 et MAGOH constituent un complexe stable sans MLN51, le pré-EJC auquel s'ajoute MLN51 lors de la ligature des exons lors de l'épissage (Gehring *et al*., 2009a). Ces études ont particulièrement montré qu'en plus des EJC présents aux positions canoniques c'est-à-dire à environ 20-25 nucléotides en amont des jonctions exon-exon, il y avait aussi des ncEJC. Bien que la fonction de ces ncEJC soit encore inconnue, il a été observé une absence de MLN51 dans ces jonctions. En outre, les protéines du pré-EJC, eIF4A3, Y14 et MAGOH peuvent affecter l'épissage de nombreux ARN alors que MLN51 n'a eu que peu/pas d'effet (Ashton-Beaucage *et al*., 2010; Fukumura *et al*., 2016; Hayashi *et al*., 2014; Liu *et al*., 2016; Malone *et al*., 2014; Michelle *et al*., 2012; Roignant et Treisman, 2010; Wang *et al*., 2014). De plus, MLN51 possède des fonctions indépendantes de l'EJC qui serait à l'origine de son effet sur les ARNm E6 et E7 d'HPV16. MLN51 est ainsi impliquée dans la traduction des ARNm, l'assemblage des granules de stress et le

désassemblage des P-bodies (Baguet *et al*., 2007; Chazal *et al*., 2013; Cougot *et al*., 2014). L'effet contraire observé par MLN51 pourrait être lié à un effet non spécifique du siRNA α-MLN51 utilisé. D'ailleurs, un second siRNA n'a montré que peu/voire aucun effet de l'inhibition de MLN51 sur l'expression d'E6 et E7 dans les cellules SiHa. Bien évidemment, il faudrait donc confirmer l'effet observé lors de l'inhibition de MLN51. Dans tous les cas, l'inhibition de MLN51 a un effet différent de l'inhibition d'eIF4A3 et Y14 dans les cellules SiHa.

Des nombreux effets connus de l'EJC sur les ARNm, nous avons étudié son impact sur l'épissage et la stabilité des transcrits viraux. En outre, nous avons montré que l'inhibition d'eIF4A3, Y14 ou MLN51 n'altérait pas l'épissage des transcrits E6 et E7 d'HPV16. En revanche, la stabilité des ARNm viraux est augmenté suite à l'inhibition d'eIF4A3 dans les cellules Ca Ski, alors que ni l'inhibition d'eIF4A3 ni de Y14 ne modifie la demi-vie des transcrits d'HPV16. Le mécanisme par lequel les composants de l'EJC affectent l'expression des ARNm viraux, serait différent entre les cellules SiHa et Ca Ski. Il pourrait être lié aux contextes cellulaires, variant entre ces cellules, notamment au statut du génome viral. Les cellules Ca Ski possèdent environ 200-500 copies de génomes viraux intégrés en cancatémères, tandis que les cellules SiHa possèdent 2 copies d'HPV16. L'altération de l'expression d'E6 et E7 observée dans les cellules Ca Ski et propre à eIF4A3, suggèrent un effet indépendant de l'EJC.

En outre, ces résultats suggèrent que la stabilisation induite par l'inhibition d'eIF4A3 dans les cellules pourrait être liée à une dégradation par des ARN non codants. Bien que nous n'ayons montré aucun changement du miR-375 lors de l'inhibition des composants de l'EJC, une implication d'eIF4A3 *via* des longs ARN non codants n'est pas à écarter. Comme décrit dans le cas de maturation des ARN ribosomaux (Alexandrov *et al*., 2011), eIF4A3 pourrait jouer un rôle dans la maturation et la fonction d'ARN non codants. En effet, il a été montré qu'eIF4A3 pouvait interagir avec des ARN non codants (Han *et al*., 2016; Zhang *et al*., 2018) avec lesquels il pourrait affecter l'expression de ARNm viraux. L'absence d'effet de l'EJC sur l'épissage et la stabilité des ARNm viraux dans les cellules SiHa pourrait suggérer un effet indirect de l'EJC sur l'expression d'E6 et E7. L'inhibition de l'EJC pourrait altérer l'expression d'un facteur de transcription impliqué dans la régulation de p97. L'absence d'effet sur la stabilité pourrait être liée à l'impossibilité de bien détecter celle-ci à cause de biais techniques.

Par ailleurs, en utilisant la spectrométrie de masse, Martinez-Salazar et collaborateurs ont récemment identifié eIF4A3 et les composants de l'EJC dans les protéines faisant partie du mRNP E6E7 (Martínez-Salazar *et al*., 2014). Ces résultats suggèrent que les composants de l'EJC affectent directement le métabolisme des ARNm d'HPV16. En utilisant des expériences de RNA-IP, nous avons montré qu'eIF4A3 faisait partie des mRNP E6 et E7 d'HPV16 dans les cellules SiHa et Ca Ski. En outre, eIF4A3 se lie aux ARNm E6 non seulement d'une façon dépendante mais aussi indépendante de l'épissage. Ce mode de liaison pourrait être dépendant de la séquence de l'ARN E6. Nos résultats montrent que la séquence de l'intron 1 d'E6, qui est excisé lors de la production d'E6*I, n'est pas nécessaire à l'interaction avec eIF4A3. Toutefois, d'autres expériences sont nécessaires pour confirmer l'existence d'une liaison dépendante de la séquence, mais aussi pour évaluer le mode de cette interaction, qu'elle soit directe et/ou indirecte. A la lumière de cette interaction entre eIF4A3, et probablement de l'EJC avec les ARN d'HPV16, nous avons décidé de considérer l'existence d'un effet direct de l'EJC sur l'expression des ARNm d'HPV16.

En particulier, l'EJC pourrait affecter l'expression des ARNm E6 et E7 *via* son rôle dans la dégradation des ARNm par le NMD. Même si l'EJC n'affecte pas la stabilité des ARNm viraux dans les cellules SiHa (Voir discussion ci-dessus), suggérant donc un effet indépendant du NMD, ce mécanisme pourrait affecter leur expression indépendamment de l'EJC. En effet, les transcrits polycistroniques d'HPV16, comme ceux de nombreux virus, possèdent des caractéristiques induisant leur reconnaissance et dégradation par le NMD. D'ailleurs, plusieurs ARN messagers et/ou génomiques de virus à ARN ont été identifiés comme des cibles du NMD (Balistreri *et al*., 2014; Garcia *et al*., 2014; LeBlanc et Beemon, 2004; Mocquet *et al*., 2012; Nakano *et al*., 2013). Chez les mammifères, le NMD est principalement dépendant de l'EJC bien que celui puissent être dispensable quant à la reconnaissance et la dégradation de certains transcrits. L'étude de l'effet de l'inhibition du NMD sur les transcrits d'HPV16, par de nombreuses approches, a montré des résultats contrastés, particulièrement entre les cellules SiHa et Ca Ski. Nos résultats ont montré que l'inhibition du NMD par CHX et le 5-azaC induit l'augmentation des transcrits E6 et E7, mais pas d'E6Fl, d'HPV16 dans les cellules SiHa. De plus, l'inhibition d'UPF1 et SMG1 a induit l'augmentation des transcrits d'HPV16, mais pas E6Fl. Ces résultats suggèrent que les ARNm E6 et E7 sont dégradés par le NMD, au moins dans les cellules SiHa, alors que l'isoforme E6Fl semble y être immunisé. Ces observations pourraient être liées à la courte demi-vie des ARNm E6Fl. En effet, l'efficacité de dégradation des ARNm par le NMD semble diminuer lorsque ceux-ci

possèdent une faible stabilité (Lindeboom *et al*., 2016). A l'inverse, dans les cellules Ca Ski, l'inhibition du NMD par des drogues ou *via* UPF1 et SMG1, n'a pas induit d'augmentation des ARNm E6 et E7 d'HPV16. Bien que les niveaux d'ARNm viraux soient inchangés lors de l'inhibition de SMG1 dans les cellules Ca Ski, nous avons observé une augmentation de leur expression au niveau protéique. Cette augmentation pourrait résulter d'une augmentation de la traduction des ARNm E6 et E7 et/ou de la stabilité des protéines produites suite à l'inhibition de SMG1. D'ailleurs, SMG1 est une kinase impliquée dans la phosphorylation de nombreuses protéines (De Cicco *et al*., 2015; Gewandter *et al*., 2011), dont certaines pourraient provoquer la déstabilisation d'E6 et E7.

La différence de résultats obtenus dans les cellules Ca Ski et SiHa pourrait être liée à une différence de l'activité du NMD dans ces cellules. D'ailleurs, l'augmentation des niveaux de SC35, un ARN ciblé par le NMD, est plus forte dans les SiHa que dans les Ca Ski. L'activité du NMD est-elle différente entre les cellules SiHa et Ca Ski ? HPV16 affecte-t-il l'activité du NMD ? En tout cas, l'expression des transcrits d'HPV16 suggère que ceux-ci puissent échapper au NMD d'une façon ou d'une autre. D'ailleurs, plusieurs virus à ARN ont développé différents mécanismes permettant de protéger leur ARN d'une dégradation par le NMD (Balistreri *et al*., 2014; Garcia *et al*., 2014; LeBlanc et Beemon, 2004; Mocquet *et al*., 2012; Nakano *et al*., 2013; Ramage *et al*., 2015). Les ARN du RSV possèdent un élément *cis*-protecteur leur permettant d'échapper au NMD, tandis que HTLV-1 inhibe l'activité du NMD via l'expression des protéines Tax et Rex (LeBlanc et Beemon, 2004; Mocquet *et al*., 2012; Nakano *et al*., 2013). Les ARNm E6Fl, insensibles au NMD, pourrait posséder des séquences/caractéristiques leur permettant d'échapper aux NMD. Celle-ci pourrait être liée à la séquence de l'intron 1 qui est spécifique à E6Fl, ou bien à l'absence de l'épissage dans tout ces transcrits. Notre stratégie et notre méthode d'analyse ne permettent pas de distinguer ces deux cas. De façon intéressante, nous avons observé une baisse de l'efficacité du NMD dans les cellules HPV16+ (SiHa et Ca Ski) par rapport aux cellules HPV- (C33A). L'inhibition du NMD est plus prononcée dans les cellules Ca Ski, fournissant un élément de réponse quant aux résultats observés dans cette lignée. De plus, l'expression ectopique d'E6 et E6*I dans les cellules C33A a été suffisante pour inhiber le NMD, confirmant un peu plus l'effet d'HPV16 sur l'activité du NMD. Le mécanisme par lequel E6 inhibe le NMD n'a pas été élucidé. D'une part, E6 pourrait altérer l'expression et la fonction d'un ou plusieurs facteurs du NMD (Fig. 32). D'ailleurs, il a été montré que l'expression de SMG1 est diminuée dans des cancers VADS HPV+ en comparaison avec les cancers HPV- (Gubanova *et al*., 2012). Toutefois, ni

l'expression de SMG1 ni celle des autres facteurs de l'EJC (eIF4A3, Y14, MAGOH, Y14, UPF3) et du NMD (UPF1, UPF2) que nous avons analysés ici n'a été dérégulée dans les cancers HPV+. L'inhibition du NMD dans les cellules du HPV+ résulte de la dérégulation d'autres facteurs (SMG5, SMG6 et SMG7). Enfin, il est toujours possible que la dérégulation du NMD soit due à une altération de l'expression et/ou de l'activité de ces facteurs au niveau protéique. D'autre part, l'inhibition du NMD pourrait résulter d'une élévation des ROS liée à l'expression d'E6 et E6*I (Fig. 32). En effet, une élévation des ROS aboutit la phosphorylation d'eIF2α, lequel inhibe le NMD (Gardner, 2008; Lejeune, 2017; Nickless *et al*., 2017; Popp et Maquat, 2018; Wang *et al*., 2011a). D'ailleurs le niveau de ROS est inversement proportionnel à l'activité du NMD, étant plus élevé dans les cellules Ca Ski que les cellules SiHa (Williams *et al*., 2014). Par conséquent, d'autres expériences s'avèrent nécessaires pour caractériser le mécanisme d'inhibition du NMD par HPV16, notamment en discernant l'effet d'E6 de celui E6*I.

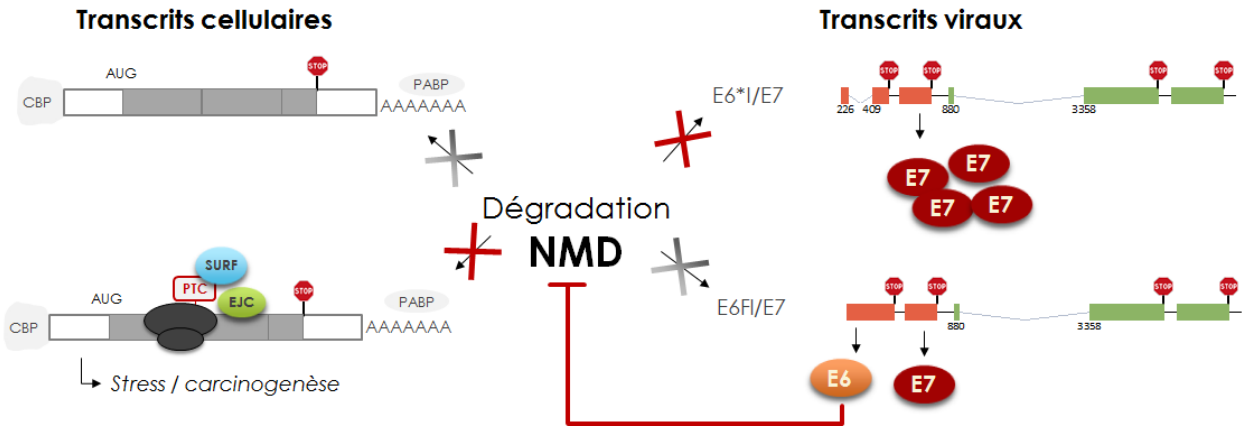


Figure 32. Interaction entre HPV16 et le NMD.

HPV16 serait donc le premier virus à ADN qui inhibe le NMD entre autre *via* la protéine E6/E6*I. L'inhibition du NMD permettrait donc de protéger les transcrits viraux de la dégradation mais aussi d'altérer l'expression de gènes cellulaires induit par le NMD, étant donné le rôle du NMD dans le contrôle de l'expression génique. D'ailleurs, nous avons observé une augmentation de l'expression de cibles de NMD, NRP1 (Neuropilin 1) et GABARAPL1 (GABA Type A Receptor Associated Protein Like 1) dans les cellules HPV16+. En outre, il serait important de déterminer le rôle de cette nouvelle activité d'E6 dans cycle viral et la carcinogenèse induite par les HPV-HR. De façon intéressante, Banister et collaborateurs, ont identifié des mutations de TP53, PTEN, P300 et SMG7, dans des cancers du col de l'utérus dans lesquels les HPV sont transcriptionnellement inactifs. Dans

ces tumeurs où E6 et E7 ne sont pas exprimés, ils ont identifié des mutations dans les cibles d'E6 et E7 (Banister *et al*., 2017). SMG7 étant un facteur du NMD, cela suggère l'importance de la voie du NMD dans la carcinogenèse du col de l'utérus. Par ailleurs, les protéines E6 et E7 pourraient aussi inhiber le NMD, induisant l'augmentation des transporteurs ABC (ATP Binding Cassette) (Rigalli *et al*., 2018). Cela pourrait permettre une meilleure adaptation cellulaire aux stress, mais aussi aux chimiothérapies.

Il bien connu qu'E6 et E7 collaborent pour mener à la transformation cellulaire. De plus, la co-expression d'E6 et E7 contribue à renforcer leur effets respectifs (Mittal et Banks, 2017; Moody et Laimins, 2010). Il serait aussi très intéressant d'étudier les rôles d'E6 avec E7 dans l'inhibition du NMD. La persistance des infections à HPV est liée non seulement à l'intégration mais aussi à l'échappement aux mécanismes de lutte anti-viraux. Si l'inhibition du NMD permet d'éluder la restriction qu'il induit vis-à-vis de l'expression des protéines virales, une thérapie visant à restaurer son activité pourrait contrecarrer l'infection voire la carcinogène dépendante des HPV-HR. Enfin, il sera aussi intéressant de déterminer le rôle d'une telle activité chez les HPV-HR mais aussi HPV-BR.

4. Régulation de l'épissage au sein de l'ORF E6

Comme nous l'avons vu, l'épissage d'E6 est important, non seulement pour réguler l'expression d'E6 mais aussi pour moduler la balance E6/E6*I. Alors que les mécanismes de régulation de l'épissage au sein des ORF E1, E4, L1 et L2 ont été bien caractérisés, peu d'informations ont été fournies en ce qui concerne les ORF E6 et E7 (Graham et Faizo, 2017). En particulier, les séquences *cis*-régulatrices potentielles, de type ESE, ESS, ISE ou ISS, impliquées dans la régulation de l'épissage d'E6 n'ont pas été identifiées. Notre travail a consisté à tenter d'identifier de tels éléments, en étudiant notamment des mutations modifiant l'épissage d'E6 d'HPV16. Bien que nous ayons découvert des mutations associées à un changement d'épissage, nos travaux s'avèrent trop préliminaires pour identifier des séquences *cis*-régulatrices d'épissage. Des études expérimentales seront menées pour caractériser plus amplement l'effet de ces mutations sur l'épissage d'E6 d'HPV16. Une fois les éléments régulateurs identifiés, une caractérisation des protéines *trans*-régulatrices d'épissage qui s'y fixent sera effectuée. Un ciblage de ces facteurs permettrait de moduler l'épissage d'E6 et d'altérer la balance d'expression d'E6, E7 et E6*I.

5. Conclusion générale

Notre travail a consisté à tenter d'approfondir les connaissances sur les fonctions d'E6 des HPV-HR, notamment d'HPV16, ainsi que les mécanismes impliqués dans la régulation de l'expression des gènes viraux. Nous avons pu montrer que l'expression d'E6*I avait un effet au niveau cellulaire indépendant d'E6. Nous avons aussi mis en évidence de nouvelles propriétés d'E6, notamment son effet inhibiteur sur le NMD. La régulation de l'expression de nombreux gènes cellulaires, impliqués dans le maintien de l'homéostasie et la réponse aux stress, suggère une importance de ce nouveau rôle d'E6 dans la carcinogenèse associée aux HPV-HR.

En particulier, cette propriété s'inscrit dans un nouveau cercle d'interaction entre HPV16 et le NMD. Le NMD peut réguler l'expression des transcrits d'HPV16, lequel module aussi son activité. La régulation par le NMD se situe au niveau posttranscriptionnel, et pourrait participer à la régulation de la balance E6/E6*I, dont l'importance commence de plus en plus à être mise en évidence.

Enfin, les connaissances apportées par notre travail soulèvent de nombreuses questions fascinantes sur le monde des ARN, et son importance dans le cycle viral des HPV-HR ainsi que la carcinogenèse qui y est associée.

Références Bibliographiques

Adachi, T., Togashi, H., Suzuki, A., Kasai, S., Ito, J., Sugahara, K., and Kawata, S. (2005). NAD(P)H oxidase plays a crucial role in PDGF-induced proliferation of hepatic stellate cells. Hepatol. *41*, 1272–1281.

Ahn, S., Kim, J., and Hwang, J. (2013). CK2-mediated TEL2 phosphorylation augments nonsense-mediated mRNA decay (NMD) by increase of SMG1 stability. Biochim. Biophys. Acta *1829*, 1047–1055.

Ajamian, L., Abrahamyan, L., Milev, M., Ivanov, P.V., Kulozik, A.E., Gehring, N.H., and Mouland, A.J. (2008). Unexpected roles for UPF1 in HIV-1 RNA metabolism and translation. RNA *14*, 914–927.

Ajamian, L., Abel, K., Rao, S., Vyboh, K., García-de-Gracia, F., Soto-Rifo, R., Kulozik, A.E., Gehring, N.H., and Mouland, A.J. (2015). HIV-1 Recruits UPF1 but Excludes UPF2 to Promote Nucleocytoplasmic Export of the Genomic RNA. Biomolecules *5*, 2808–2839.

Ajiro, M., and Zheng, Z.-M. (2015). E6^E7, a novel splice isoform protein of human papillomavirus 16, stabilizes viral E6 and E7 oncoproteins via HSP90 and GRP78. MBio *6*, e02068-02014.

Ajiro, M., Jia, R., Zhang, L., Liu, X., and Zheng, Z.-M. (2012). Intron definition and a branch site adenosine at nt 385 control RNA splicing of HPV16 E6*I and E7 expression. PloS One *7*, 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.

Alexandrov, A., Colognori, D., and Steitz, J.A. (2011). Human eIF4AIII interacts with an eIF4G-like partner, NOM1, revealing an evolutionarily conserved function outside the exon junction complex. Genes Dev. *25*, 1078–1090.

Alexandrov, A., Colognori, D., Shu, M.-D., and Steitz, J.A. (2012). Human spliceosomal protein CWC22 plays a role in coupling splicing to exon junction complex deposition and nonsense-mediated decay. Proc. Natl. Acad. Sci. U. S. A. *109*, 21313–21318.

Amrani, N., Ganesan, R., Kervestin, S., Mangus, D.A., Ghosh, S., and Jacobson, A. (2004). A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. Nature *432*, 112–118.

Andersen, C.B.F., Ballut, L., Johansen, J.S., Chamieh, H., Nielsen, K.H., Oliveira, C.L.P., Pedersen, J.S., Séraphin, B., Le Hir, H., and Andersen, G.R. (2006). Structure of the exon junction core complex with a trapped DEAD-box ATPase bound to RNA. Science *313*, 1968–1972.

Andrews, N., Stowe, J., and Miller, E. (2017). No increased risk of Guillain-Barré syndrome after human papilloma virus vaccine: A self-controlled case-series study in England. Vaccine *35*, 1729–1732.

Änkö, M.-L. (2014). Regulation of gene expression programmes by serine-arginine rich splicing factors. Semin. Cell Dev. Biol. *32*, 11–21.

Ansari, T., Brimer, N., and Vande Pol, S.B. (2012). Peptide interactions stabilize and restructure human papillomavirus type 16 E6 to interact with p53. J. Virol. *86*, 11386–11391.

Arenz, A., Ziemann, F., Mayer, C., Wittig, A., Dreffke, K., Preising, S., Wagner, S., Klussmann, J.-P., Engenhart-Cabillic, R., and Wittekindt, C. (2014). Increased radiosensitivity of HPV-positive head and neck cancer cell lines due to cell cycle dysregulation and induction of apoptosis. Strahlenther. Onkol. Organ Dtsch. Rontgengesellschaft Al *190*, 839–846.

Arnheim-Dahlström, L., Pasternak, B., Svanström, H., Sparén, P., and Hviid, A. (2013). Autoimmune, neurological, and venous thromboembolic adverse events after immunisation of adolescent girls with quadrivalent human papillomavirus vaccine in Denmark and Sweden: cohort study. BMJ *347*, f5906.

Ashton-Beaucage, D., Udell, C.M., Lavoie, H., Baril, C., Lefrançois, M., Chagnon, P., Gendron, P., Caron-Lizotte, O., Bonneil, E., Thibault, P., et al. (2010). The exon junction complex controls the splicing of MAPK and other long intron-containing transcripts in Drosophila. Cell *143*, 251–262.

Asselta, R., Duga, S., Spena, S., Santagostino, E., Peyvandi, F., Piseddu, G., Targhetta, R., Malcovati, M., Mannucci, P.M., and Tenchini, M.L. (2001). Congenital afibrinogenemia: mutations leading to premature termination codons in fibrinogen A alpha-chain gene are not associated with the decay of the mutant mRNAs. Blood *98*, 3685–3692.

Baguet, A., Degot, S., Cougot, N., Bertrand, E., Chenard, M.-P., Wendling, C., Kessler, P., Le Hir, H., Rio, M.-C., and Tomasetto, C. (2007). The exon-junction-complex-component metastatic lymph node 51 functions in stress-granule assembly. J. Cell Sci. *120*, 2774–2784.

Balistreri, G., Horvath, P., Schweingruber, C., Zünd, D., McInerney, G., Merits, A., Mühlemann, O., Azzalin, C., and Helenius, A. (2014). The host nonsense-mediated mRNA decay pathway restricts Mammalian RNA virus replication. Cell Host Microbe *16*, 403–411.

Ballut, L., Marchadier, B., Baguet, A., Tomasetto, C., Séraphin, B., and Le Hir, H. (2005). The exon junction core complex is locked onto RNA by inhibition of eIF4AIII ATPase activity. Nat. Struct. Mol. Biol. *12*, 861–869.

Banister, C.E., Liu, C., Pirisi, L., Creek, K.E., and Buckhaults, P.J. (2017). Identification and characterization of HPV-independent cervical cancers. Oncotarget *8*, 13375–13386.

Barbosa, I., Haque, N., Fiorini, F., Barrandon, C., Tomasetto, C., Blanchette, M., and Le Hir, H. (2012). Human CWC22 escorts the helicase eIF4AIII to spliceosomes and promotes exon junction complex assembly. Nat. Struct. Mol. Biol. *19*, 983–990.

Bateman, J.F., Freddi, S., Nattrass, G., and Savarirayan, R. (2003). Tissue-specific RNA surveillance? Nonsense-mediated mRNA decay causes collagen X haploinsufficiency in Schmid metaphyseal chondrodysplasia cartilage. Hum. Mol. Genet. *12*, 217–225.

Bello, J.O.M., Nieva, L.O., Paredes, A.C., Gonzalez, A.M.F., Zavaleta, L.R., and Lizano, M. (2015). Regulation of the Wnt/β-Catenin Signaling Pathway by Human Papillomavirus E6 and E7 Oncoproteins. Viruses *7*, 4734–4755.

Bergvall, M., Melendy, T., and Archambault, J. (2013). The E1 proteins. Virology *445*, 35–56.

Bernard, H.-U. (2013). Regulatory elements in the viral genome. Virology *445*, 197–204.

Bernard, H.-U., Burk, R.D., Chen, Z., van Doorslaer, K., zur Hausen, H., and de Villiers, E.-M. (2010). Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. Virology *401*, 70–79.

Bhosale, P.G., Pandey, M., Cristea, S., Shah, M., Patil, A., Beerenwinkel, N., Schäffer, A.A., and Mahimkar, M.B. (2017). Recurring Amplification at 11q22.1-q22.2 Locus Plays an Important Role in Lymph Node Metastasis and Radioresistance in OSCC. Sci. Rep. *7*, 16051.

Bienkowska-Haba, M., Luszczek, W., Keiffer, T.R., Guion, L.G.M., DiGiuseppe, S., Scott, R.S., and Sapp, M. (2017). Incoming human papillomavirus 16 genome is lost in PML protein-deficient HaCaT keratinocytes. Cell. Microbiol. *19*.

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.

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.

Bodily, J.M., Hennigan, C., Wrobel, G.A., and Rodriguez, C.M. (2013). Regulation of the human papillomavirus type 16 late promoter by E7 and the cell cycle. Virology *443*, 11–19.

Boehm, V., and Gehring, N.H. (2016). Exon Junction Complexes: Supervising the Gene Expression Assembly Line. Trends Genet. TIG *32*, 724–735.

Bono, F., Ebert, J., Unterholzner, L., Güttler, T., Izaurralde, E., and Conti, E. (2004). Molecular insights into the interaction of PYM with the Mago-Y14 core of the exon junction complex. EMBO Rep. *5*, 304–310.

Bono, F., Ebert, J., Lorentzen, E., and Conti, E. (2006). The crystal structure of the exon junction complex reveals how it maintains a stable grip on mRNA. Cell *126*, 713–725.

Bono, F., Cook, A.G., Grünwald, M., Ebert, J., and Conti, E. (2010). Nuclear import mechanism of the EJC component Mago-Y14 revealed by structural studies of importin 13. Mol. Cell *37*, 211–222.

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*, E3255-3264.

Borbély, A.A., Murvai, M., Kónya, J., Beck, Z., Gergely, L., Li, F., and Veress, G. (2006). Effects of human papillomavirus type 16 oncoproteins on survivin gene expression. J. Gen. Virol. *87*, 287–294.

Braunstein, T.H., Madsen, B.S., Gavnholt, B., Rosenstierne, M.W., Johnsen, C.K., and Norrild, B. (1999). Identification of a new promoter in the early region of the human papillomavirus type 16 genome. J. Gen. Virol. *80 (Pt 12)*, 3241–3250.

Bravo, I.G., and Félez-Sánchez, M. (2015). Papillomaviruses: Viral evolution, cancer and evolutionary medicine. Evol. Med. Public Health *2015*, 32–51.

Bruno, I.G., Karam, R., Huang, L., Bhardwaj, A., Lou, C.H., Shum, E.Y., Song, H.-W., Corbett, M.A., Gifford, W.D., Gecz, J., et al. (2011). Identification of a microRNA that activates gene expression by repressing nonsense-mediated RNA decay. Mol. Cell *42*, 500–510.

Buchwald, G., Schüssler, S., Basquin, C., Le Hir, H., and Conti, E. (2013). Crystal structure of the human eIF4AIII-CWC22 complex shows how a DEAD-box protein is inhibited by a MIF4G domain. Proc. Natl. Acad. Sci. U. S. A. *110*, E4611-4618.

Budanov, A.V. (2014). The role of tumor suppressor p53 in the antioxidant defense and metabolism. Subcell. Biochem. *85*, 337–358.

Budanov, A.V., Sablina, A.A., Feinstein, E., Koonin, E.V., and Chumakov, P.M. (2004). Regeneration of peroxiredoxins by p53-regulated sestrins, homologs of bacterial AhpD. Science *304*, 596–600.

Bühler, M., Steiner, S., Mohn, F., Paillusson, A., and Mühlemann, O. (2006). EJC-independent degradation of nonsense immunoglobulin-mu mRNA depends on 3' UTR length. Nat. Struct. Mol. Biol. *13*, 462–464.

Cai, Q., Lv, L., Shao, Q., Li, X., and Dian, A. (2013). Human papillomavirus early proteins and apoptosis. Arch. Gynecol. Obstet. *287*, 541–548.

Camini, F.C., da Silva Caetano, C.C., Almeida, L.T., and de Brito Magalhães, C.L. (2017). Implications of oxidative stress on viral pathogenesis. Arch. Virol. *162*, 907–917.

Cao, L., Qi, L., Zhang, L., Song, W., Yu, Y., Xu, C., Li, L., Guo, Y., Yang, L., Liu, C., et al. (2017). Human nonsense-mediated RNA decay regulates EMT by targeting the TGF-ß signaling pathway in lung adenocarcinoma. Cancer Lett. *403*, 246–259.

Cartegni, L., Chew, S.L., and Krainer, A.R. (2002). Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nat. Rev. Genet. *3*, 285–298.

Chamieh, H., Ballut, L., Bonneau, F., and Le Hir, H. (2008). NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity. Nat. Struct. Mol. Biol. *15*, 85–93.

Chan, D., Weng, Y.M., Graham, H.K., Sillence, D.O., and Bateman, J.F. (1998). A nonsense mutation in the carboxyl-terminal domain of type X collagen causes haploinsufficiency in schmid metaphyseal chondrodysplasia. J. Clin. Invest. *101*, 1490–1499.

Chan, W.-K., Huang, L., Gudikote, J.P., Chang, Y.-F., Imam, J.S., MacLean, J.A., and Wilkinson, M.F. (2007). An alternative branch of the nonsense-mediated decay pathway. EMBO J. *26*, 1820–1830.

Chang, L., Li, C., Guo, T., Wang, H., Ma, W., Yuan, Y., Liu, Q., Ye, Q., and Liu, Z. (2016). The human RNA surveillance factor UPF1 regulates tumorigenesis by targeting Smad7 in hepatocellular carcinoma. J. Exp. Clin. Cancer Res. CR *35*, 8.

Chapin, A., Hu, H., Rynearson, S.G., Hollien, J., Yandell, M., and Metzstein, M.M. (2014). In vivo determination of direct targets of the nonsense-mediated decay pathway in Drosophila. G3 Bethesda Md *4*, 485–496.

Chazal, P.-E., Daguenet, E., Wendling, C., Ulryck, N., Tomasetto, C., Sargueil, B., and Le Hir, H. (2013). EJC core component MLN51 interacts with eIF3 and activates translation. Proc. Natl. Acad. Sci. U. S. A. *110*, 5903–5908.

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, 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., Kong, X., Zhang, Z., Chen, W., Chen, J., Li, H., Cao, W., Ge, Y., and Fang, S. (2014b). Alpha-2-macroglobulin as a radioprotective agent: a review. Chin. J. Cancer Res. *26*, 611–621.

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.

Cheng, H., Dufu, K., Lee, C.-S., Hsu, J.L., Dias, A., and Reed, R. (2006). Human mRNA export machinery recruited to the 5' end of mRNA. Cell *127*, 1389–1400.

Cheunim, T., Zhang, J., Milligan, S.G., McPhillips, M.G., and Graham, S.V. (2008). The alternative splicing factor hnRNP A1 is up-regulated during virus-infected epithelial cell differentiation and binds the human papillomavirus type 16 late regulatory element. Virus Res. *131*, 189–198.

Chibon, F. (2013). Cancer gene expression signatures - the rise and fall? Eur. J. Cancer *49*, 2000–2009.

Cho, H., Han, S., Choe, J., Park, S.G., Choi, S.S., and Kim, Y.K. (2013). SMG5-PNRC2 is functionally dominant compared with SMG5-SMG7 in mammalian nonsense-mediated mRNA decay. Nucleic Acids Res. *41*, 1319–1328.

Choe, J., Ryu, I., Park, O.H., Park, J., Cho, H., Yoo, J.S., Chi, S.W., Kim, M.K., Song, H.K., and Kim, Y.K. (2014). eIF4AIII enhances translation of nuclear cap-binding complex-bound mRNAs by promoting disruption of secondary structures in 5'UTR. Proc. Natl. Acad. Sci. U. S. A. *111*, E4577-4586.

Choi, D.-H., Kim, J.-H., Seo, J.-H., Lee, J., Choi, W.S., and Kim, Y.-S. (2014). Matrix metalloproteinase-3 causes dopaminergic neuronal death through Nox1-regenerated oxidative stress. PloS One *9*, e115954.

Chopjitt, P., Pientong, C., Bumrungthai, S., Kongyingyoes, B., and Ekalaksananan, T. (2015). Activities of E6 Protein of Human Papillomavirus 16 Asian Variant on miR-21 Up-regulation and Expression of Human Immune Response Genes. Asian Pac. J. Cancer Prev. *16*, 3961–3968.

Chuang, T.-W., Chang, W.-L., Lee, K.-M., and Tarn, W.-Y. (2013). The RNA-binding protein Y14 inhibits mRNA decapping and modulates processing body formation. Mol. Biol. Cell *24*, 1–13.

Collier, B., Goobar-Larsson, L., Sokolowski, M., and Schwartz, S. (1998). Translational inhibition in vitro of human papillomavirus type 16 L2 mRNA mediated through interaction with heterogenous ribonucleoprotein K and poly(rC)-binding proteins 1 and 2. J. Biol. Chem. *273*, 22648–22656.

Colombo, M., Karousis, E.D., Bourquin, J., Bruggmann, R., and Mühlemann, O. (2017). Transcriptome-wide identification of NMD-targeted human mRNAs reveals extensive redundancy between SMG6- and SMG7-mediated degradation pathways. RNA *23*, 189–201.

Cougot, N., Daguenet, E., Baguet, A., Cavalier, A., Thomas, D., Bellaud, P., Fautrel, A., Godey, F., Bertrand, E., Tomasetto, C., et al. (2014). Overexpression of MLN51 triggers P-body disassembly and formation of a new type of RNA granules. J. Cell Sci. *127*, 4692–4701.

Cricca, M., Venturoli, S., Leo, E., Costa, S., Musiani, M., and Zerbini, M. (2009). 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.

Crow, J.M. (2012). HPV: The global burden. Nature *488*, S2-3.

Cumming, S.A., Chuen-Im, T., Zhang, J., and Graham, S.V. (2009). The RNA stability regulator HuR regulates L1 protein expression in vivo in differentiating cervical epithelial cells. Virology *383*, 142–149.

Czech, B., and Hannon, G.J. (2011). Small RNA sorting: matchmaking for Argonautes. Nat. Rev. Genet. *12*, 19–31.

Daguenet, E., Baguet, A., Degot, S., Schmidt, U., Alpy, F., Wendling, C., Spiegelhalter, C., Kessler, P., Rio, M.-C., Le Hir, H., et al. (2012). Perispeckles are major assembly sites for the exon junction core complex. Mol. Biol. Cell *23*, 1765–1782.

Dalstein, V., Riethmuller, D., Prétet, J.-L., Le Bail Carval, K., Sautière, J.-L., Carbillet, J.-P., Kantelip, B., Schaal, J.-P., and Mougin, C. (2003). Persistence and load of high-risk HPV are predictors for development of high-grade cervical lesions: a longitudinal French cohort study. Int. J. Cancer *106*, 396–403.

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.

D'Costa, Z.J., Jolly, C., Androphy, E.J., Mercer, A., Matthews, C.M., and Hibma, M.H. (2012). Transcriptional repression of E-cadherin by human papillomavirus type 16 E6. PloS One *7*, e48954.

De Cicco, M., Rahim, M.S.A., and Dames, S.A. (2015). Regulation of the Target of Rapamycin and Other Phosphatidylinositol 3-Kinase-Related Kinases by Membrane Targeting. Membranes *5*, 553–575.

De la Cruz-Hernández, E., García-Carrancá, A., Mohar-Betancourt, A., Dueñas-González, A., Contreras-Paredes, A., Pérez-Cardenas, E., Herrera-Goepfert, R., and Lizano-Soberón, M. (2005). Differential splicing of E6 within human papillomavirus type 18 variants and functional consequences. J. Gen. Virol. *86*, 2459–2468.

De Vuyst, H., Clifford, G.M., Nascimento, M.C., Madeleine, M.M., and Franceschi, S. (2009). 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.

De-Castro Arce, J., Göckel-Krzikalla, E., and Rösl, F. (2012). Silencing of multi-copy HPV16 by viral self-methylation and chromatin occlusion: a model for epigenetic virus-host interaction. Hum. Mol. Genet. *21*, 1693–1705.

Degot, S., Régnier, C.H., Wendling, C., Chenard, M.-P., Rio, M.-C., and Tomasetto, C. (2002). Metastatic Lymph Node 51, a novel nucleo-cytoplasmic protein overexpressed in breast cancer. Oncogene *21*, 4422–4434.

Degot, S., Le Hir, H., Alpy, F., Kedinger, V., Stoll, I., Wendling, C., Seraphin, B., Rio, M.-C., and Tomasetto, C. (2004). 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.

Denecke, J., Kranz, C., Kemming, D., Koch, H.-G., and Marquardt, T. (2004). An activated 5' cryptic splice site in the human ALG3 gene generates a premature termination codon insensitive to nonsense-mediated mRNA decay in a new case of congenital disorder of glycosylation type Id (CDG-Id). Hum. Mutat. *23*, 477–486.

Desmet, F.-O., Hamroun, D., Lalande, M., Collod-Béroud, G., Claustres, M., and Béroud, C. (2009). Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res. *37*, e67.

Dhanjal, S., Kajitani, N., Glahder, J., Mossberg, A.-K., Johansson, C., and Schwartz, S. (2015). Heterogeneous Nuclear Ribonucleoprotein C Proteins Interact with the Human Papillomavirus Type 16 (HPV16) Early 3'-Untranslated Region and Alleviate Suppression of HPV16 Late L1 mRNA Splicing. J. Biol. Chem. *290*, 13354–13371.

Di Giammartino, D.C., and Manley, J.L. (2014). New links between mRNA polyadenylation and diverse nuclear pathways. Mol. Cells *37*, 644–649.

Diebold, B.A., and Bokoch, G.M. (2001). Molecular basis for Rac2 regulation of phagocyte NADPH oxidase. Nat. Immunol. *2*, 211–215.

Diem, M.D., Chan, C.C., Younis, I., and Dreyfuss, G. (2007). PYM binds the cytoplasmic exon-junction complex and ribosomes to enhance translation of spliced mRNAs. Nat. Struct. Mol. Biol. *14*, 1173–1179.

Dillon, S., Sasagawa, T., Crawford, A., Prestidge, J., Inder, M.K., Jerram, J., Mercer, A.A., and Hibma, M. (2007). Resolution of cervical dysplasia is associated with T-cell proliferative responses to human papillomavirus type 16 E2. J. Gen. Virol. *88*, 803–813.

Dong, G., Broker, T.R., and Chow, L.T. (1994). Human papillomavirus type 11 E2 proteins repress the homologous E6 promoter by interfering with the binding of host transcription factors to adjacent elements. J. Virol. *68*, 1115–1127.

Doorbar, J. (2005). The papillomavirus life cycle. J. Clin. Virol. *32 Suppl 1*, S7-15.

Doorbar, J. (2013). Latent papillomavirus infections and their regulation. Curr. Opin. Virol. *3*, 416–421.

Doorbar, J. (2016). Model systems of human papillomavirus-associated disease. J. Pathol. *238*, 166–179.

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.

Dostie, J., and Dreyfuss, G. (2002). Translation is required to remove Y14 from mRNAs in the cytoplasm. Curr. Biol. CB *12*, 1060–1067.

Du, J., Chen, G.G., Vlantis, A.C., Chan, P.K.S., Tsang, R.K.Y., and van Hasselt, C.A. (2004). Resistance to apoptosis of HPV 16-infected laryngeal cancer cells is associated with decreased Bak and increased Bcl-2 expression. Cancer Lett. *205*, 81–88.

Durand, S., Cougot, N., Mahuteau-Betzer, F., Nguyen, C.-H., Grierson, D.S., Bertrand, E., Tazi, J., and Lejeune, F. (2007). Inhibition of nonsense-mediated mRNA decay (NMD) by a new chemical molecule reveals the dynamic of NMD factors in P-bodies. J. Cell Biol. *178*, 1145–1160.

Durzynska, J., Lesniewicz, K., and Poreba, E. (2017). Human papillomaviruses in epigenetic regulations. Mutat. Res. Rev. Mutat. Res. *772*, 36–50.

Eberle, A.B., Lykke-Andersen, S., Mühlemann, O., and Jensen, T.H. (2009). SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. Nat. Struct. Mol. Biol. *16*, 49–55.

Einstein, M.H., Baron, M., Levin, M.J., Chatterjee, A., Fox, B., Scholar, S., Rosen, J., Chakhtoura, N., Lebacq, M., van der Most, R., et al. (2011). Comparison of the immunogenicity of the human papillomavirus (HPV)-16/18 vaccine and the HPV-6/11/16/18 vaccine for oncogenic non-vaccine types HPV-31 and HPV-45 in healthy women aged 18-45 years. Hum. Vaccin. *7*, 1359–1373.

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.

Fedorova, M., Vinokurova, S., Pavlova, L., Komel'kov, A., Korolenkova, L., Kisseljov, F., and Kisseljova, N. (2016). Human papillomavirus types 16 E1 mRNA is transcribed from P14 early promoter in cervical neoplasms. Virology *488*, 196–201.

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., 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. (2014). The small splice variant of HPV16 E6, E6, reduces tumor formation in cervical carcinoma xenografts. Virology *450–451*, 153–164.

Foresta, C., Garolla, A., Parisi, S., Ghezzi, M., Bertoldo, A., Di Nisio, A., and De Toni, L. (2015). HPV prophylactic vaccination in males improves the clearance of semen infection. EBioMedicine *2*, 1487–1493.

Forman, D., de Martel, C., Lacey, C.J., Soerjomataram, I., Lortet-Tieulent, J., Bruni, L., Vignat, J., Ferlay, J., Bray, F., Plummer, M., et al. (2012). Global burden of human papillomavirus and related diseases. Vaccine *30 Suppl 5*, F12-23.

Franco, E.L., Villa, L.L., Sobrinho, J.P., Prado, J.M., Rousseau, M.C., Désy, M., and Rohan, T.E. (1999). Epidemiology of acquisition and clearance of cervical human papillomavirus infection in women from a high-risk area for cervical cancer. J. Infect. Dis. *180*, 1415–1423.

Fribourg, S., Gatfield, D., Izaurralde, E., and Conti, E. (2003). A novel mode of RBD-protein recognition in the Y14-Mago complex. Nat. Struct. Biol. *10*, 433–439.

Fukumura, K., Wakabayashi, S., Kataoka, N., Sakamoto, H., Suzuki, Y., Nakai, K., Mayeda, A., and Inoue, K. (2016). 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*.

Galloway, D.A., Gewin, L.C., Myers, H., Luo, W., Grandori, C., Katzenellenbogen, R.A., and McDougall, J.K. (2005). Regulation of telomerase by human papillomaviruses. Cold Spring Harb. Symp. Quant. Biol. *70*, 209–215.

Gao, P., Zhang, H., Dinavahi, R., Li, F., Xiang, Y., Raman, V., Bhujwalla, Z.M., Felsher, D.W., Cheng, L., Pevsner, J., et al. (2007). HIF-dependent antitumorigenic effect of antioxidants in vivo. Cancer Cell *12*, 230–238.

Gao, Q., Srinivasan, S., Boyer, S.N., Wazer, D.E., and Band, V. (1999). The E6 oncoproteins of high-risk papillomaviruses bind to a novel putative GAP protein, E6TP1, and target it for degradation. Mol. Cell. Biol. *19*, 733–744.

Garcia, D., Garcia, S., and Voinnet, O. (2014). Nonsense-mediated decay serves as a general viral restriction mechanism in plants. Cell Host Microbe *16*, 391–402.

Gardiol, D., Kühne, C., Glaunsinger, B., Lee, S.S., Javier, R., and Banks, L. (1999). Oncogenic human papillomavirus E6 proteins target the discs large tumour suppressor for proteasome-mediated degradation. Oncogene *18*, 5487–5496.

Gardner, L.B. (2008). Hypoxic inhibition of nonsense-mediated RNA decay regulates gene expression and the integrated stress response. Mol. Cell. Biol. *28*, 3729–3741.

Garnett, T.O., Filippova, M., and Duerksen-Hughes, P.J. (2006). Accelerated degradation of FADD and procaspase 8 in cells expressing human papilloma virus 16 E6 impairs TRAIL-mediated apoptosis. Cell Death Differ. *13*, 1915–1926.

Gatfield, D., and Izaurralde, E. (2002). REF1/Aly and the additional exon junction complex proteins are dispensable for nuclear mRNA export. J. Cell Biol. *159*, 579–588.

Ge, Z., Quek, B.L., Beemon, K.L., and Hogg, J.R. (2016). Polypyrimidine tract binding protein 1 protects mRNAs from recognition by the nonsense-mediated mRNA decay pathway. ELife *5*.

Gehring, N.H., Kunz, J.B., Neu-Yilik, G., Breit, S., Viegas, M.H., Hentze, M.W., and Kulozik, A.E. (2005). Exon-junction complex components specify distinct routes of nonsense-mediated mRNA decay with differential cofactor requirements. Mol. Cell *20*, 65–75.

Gehring, N.H., Lamprinaki, S., Hentze, M.W., and Kulozik, A.E. (2009a). The hierarchy of exon-junction complex assembly by the spliceosome explains key features of mammalian nonsense-mediated mRNA decay. PLoS Biol. *7*, e1000120.

Gehring, N.H., Lamprinaki, S., Kulozik, A.E., and Hentze, M.W. (2009b). Disassembly of exon junction complexes by PYM. Cell *137*, 536–548.

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.

Gewandter, J.S., Bambara, R.A., and O'Reilly, M.A. (2011). The RNA surveillance protein SMG1 activates p53 in response to DNA double-strand breaks but not exogenously oxidized mRNA. Cell Cycle Georget. Tex *10*, 2561–2567.

Gewin, L., Myers, H., Kiyono, T., and Galloway, D.A. (2004). Identification of a novel telomerase repressor that interacts with the human papillomavirus type-16 E6/E6-AP complex. Genes Dev. *18*, 2269–2282.

Glahder, J.A., Hansen, C.N., Vinther, J., Madsen, B.S., and Norrild, B. (2003). A promoter within the E6 ORF of human papillomavirus type 16 contributes to the expression of the E7 oncoprotein from a monocistronic mRNA. J. Gen. Virol. *84*, 3429–3441.

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.

Goetz, A.E., and Wilkinson, M. (2017). Stress and the nonsense-mediated RNA decay pathway. Cell. Mol. Life Sci. CMLS *74*, 3509–3531.

Gong, P., Zhao, M., and He, C. (2014). Slow co-evolution of the MAGO and Y14 protein families is required for the maintenance of their obligate heterodimerization mode. PloS One *9*, e84842.

Goodman, A. (2015). HPV testing as a screen for cervical cancer. BMJ *350*, h2372.

Goraczniak, R., and Gunderson, S.I. (2008). The regulatory element in the 3'-untranslated region of human papillomavirus 16 inhibits expression by binding CUG-binding protein 1. J. Biol. Chem. *283*, 2286–2296.

Grabowska, A.K., and Riemer, A.B. (2012). The invisible enemy - how human papillomaviruses avoid recognition and clearance by the host immune system. Open Virol. J. *6*, 249–256.

Graham, S.V. (2010). Human papillomavirus: gene expression, regulation and prospects for novel diagnostic methods and antiviral therapies. Future Microbiol. *5*, 1493–1506.

Graham, S.V. (2017). The human papillomavirus replication cycle, and its links to cancer progression: a comprehensive review. Clin. Sci. Lond. Engl. 1979 *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.

Groves, I.J., Knight, E.L.A., Ang, Q.Y., Scarpini, C.G., and Coleman, N. (2016). HPV16 oncogene expression levels during early cervical carcinogenesis are determined by the balance of epigenetic chromatin modifications at the integrated virus genome. Oncogene *35*, 4773–4786.

Grozdanov, P., Hadjidekova, S., Dimova, I., Nikolova, I., Toncheva, D., Ganchev, G., Zlatkov, V., and Galabov, A.S. (2016). Characterization of genomic changes in the cervical pre-cancerous lesions and tumors induced by different types of human papillomaviruses. Virusdisease *27*, 271–276.

Gu, W., Ding, J., Wang, X., de Kluyver, R.L., Saunders, N.A., Frazer, I.H., and Zhao, K.-N. (2007). Generalized substitution of isoencoding codons shortens the duration of papillomavirus L1 protein expression in transiently gene-transfected keratinocytes due to cell differentiation. Nucleic Acids Res. *35*, 4820–4832.

Gu, W., An, J., Ye, P., Zhao, K.-N., and Antonsson, A. (2011). Prediction of conserved microRNAs from skin and mucosal human papillomaviruses. Arch. Virol. *156*, 1161–1171.

Gubanova, E., Brown, B., Ivanov, S.V., Helleday, T., Mills, G.B., Yarbrough, W.G., and Issaeva, N. (2012). Downregulation of SMG-1 in HPV-positive head and neck squamous cell carcinoma due to promoter hypermethylation correlates with improved survival. Clin. Cancer Res. *18*, 1257–1267.

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.

Han, D., Gao, X., Wang, M., Qiao, Y., Xu, Y., Yang, J., Dong, N., He, J., Sun, Q., Lv, G., et al. (2016). Long noncoding RNA H19 indicates a poor prognosis of colorectal cancer and promotes tumor growth by recruiting and binding to eIF4A3. Oncotarget *7*, 22159–22173.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646–674.

Hanns, E., Job, S., Coliat, P., Wasylyk, C., Ramolu, L., Pencreach, E., Suarez-Carmona, M., Herfs, M., Ledrappier, S., Macabre, C., et al. (2015). Human Papillomavirus-related tumours of the oropharynx display a lower tumour hypoxia signature. Oral Oncol. *51*, 848–856.

Harden, M.E., and Munger, K. (2017). Perturbation of DROSHA and DICER expression by human papillomavirus 16 oncoproteins. Virology *507*, 192–198.

Hauer, C., Sieber, J., Schwarzl, T., Hollerer, I., Curk, T., Alleaume, A.-M., Hentze, M.W., and Kulozik, A.E. (2016). Exon Junction Complexes Show a Distributional Bias toward Alternatively Spliced mRNAs and against mRNAs Coding for Ribosomal Proteins. Cell Rep. *16*, 1588–1603.

Hayashi, R., Handler, D., Ish-Horowicz, D., and Brennecke, J. (2014). The exon junction complex is required for definition and excision of neighboring introns in Drosophila. Genes Dev. *28*, 1772–1785.

He, F., and Jacobson, A. (2015). Nonsense-Mediated mRNA Decay: Degradation of Defective Transcripts Is Only Part of the Story. Annu. Rev. Genet. *49*, 339–366.

He, J., Ji, Y., Li, A., Zhang, Q., Song, W., Li, Y., Huang, H., Qian, J., Zhai, A., Yu, X., et al. (2014). MiR-122 directly inhibits human papillomavirus E6 gene and enhances interferon signaling through blocking suppressor of cytokine signaling 1 in SiHa cells. PloS One *9*, e108410.

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.

Hellner, K., Mar, J., Fang, F., Quackenbush, J., and Münger, K. (2009). HPV16 E7 oncogene expression in normal human epithelial cells causes molecular changes indicative of an epithelial to mesenchymal transition. Virology *391*, 57–63.

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. U. S. A. *109*, 10516–10521.

Herfs, M., Soong, T.R., Delvenne, P., and Crum, C.P. (2017). Deciphering the Multifactorial Susceptibility of Mucosal Junction Cells to HPV Infection and Related Carcinogenesis. Viruses *9*.

Hinnebusch, A.G., and Lorsch, J.R. (2012). The mechanism of eukaryotic translation initiation: new insights and challenges. Cold Spring Harb. Perspect. Biol. *4*.

Hu, Y., Ye, F., Lu, W., Hong, D., Wan, X., and Xie, X. (2009). HPV16 E6-induced and E6AP-dependent inhibition of the transcriptional coactivator hADA3 in human cervical carcinoma cells. Cancer Invest. *27*, 298–306.

Huang, L., Lou, C.-H., Chan, W., Shum, E.Y., Shao, A., Stone, E., Karam, R., Song, H.-W., and Wilkinson, M.F. (2011). RNA homeostasis governed by cell type-specific and branched feedback loops acting on NMD. Mol. Cell *43*, 950–961.

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. U. S. A. *102*, 11492–11497.

Huntzinger, E., Kashima, I., Fauser, M., Saulière, J., and Izaurralde, E. (2008). SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. RNA *14*, 2609–2617.

Hwang, J., and Maquat, L.E. (2011). Nonsense-mediated mRNA decay (NMD) in animal embryogenesis: to die or not to die, that is the question. Curr. Opin. Genet. Dev. *21*, 422–430.

Hwang, S.G., Lee, D., Kim, J., Seo, T., and Choe, J. (2002). Human papillomavirus type 16 E7 binds to E2F1 and activates E2F1-driven transcription in a retinoblastoma protein-independent manner. J. Biol. Chem. *277*, 2923–2930.

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.

Imamachi, N., Tani, H., and Akimitsu, N. (2012). Up-frameshift protein 1 (UPF1): multitalented entertainer in RNA decay. Drug Discov. Ther. *6*, 55–61.

Isken, O., Kim, Y.K., Hosoda, N., Mayeur, G.L., Hershey, J.W.B., and Maquat, L.E. (2008). Upf1 phosphorylation triggers translational repression during nonsense-mediated mRNA decay. Cell *133*, 314–327.

Isok-Paas, H., Männik, A., Ustav, E., and Ustav, M. (2015). The transcription map of HPV11 in U2OS cells adequately reflects the initial and stable replication phases of the viral genome. Virol. J. *12*, 59.

Jaisamrarn, U., Castellsagué, X., Garland, S.M., Naud, P., Palmroth, J., Del Rosario-Raymundo, M.R., Wheeler, C.M., Salmerón, J., Chow, S.-N., Apter, D., et al. (2013). Natural history of progression of HPV infection to cervical lesion or clearance: analysis of the control arm of the large, randomised PATRICIA study. PloS One *8*, e79260.

James, M.A., Lee, J.H., and Klingelhutz, A.J. (2006). Human papillomavirus type 16 E6 activates NF-kappaB, induces cIAP-2 expression, and protects against apoptosis in a PDZ binding motif-dependent manner. J. Virol. *80*, 5301–5307.

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.

Jha, S., Vande Pol, S., Banerjee, N.S., Dutta, A.B., Chow, L.T., and Dutta, A. (2010). Destabilization of TIP60 by human papillomavirus E6 results in attenuation of TIP60-dependent transcriptional regulation and apoptotic pathway. Mol. Cell *38*, 700–711.

Jia, R., Liu, X., Tao, M., Kruhlak, M., Guo, M., Meyers, C., Baker, C.C., and Zheng, Z.-M. (2009). Control of the papillomavirus early-to-late switch by differentially expressed SRp20. J. Virol. *83*, 167–180.

Jin, L., and Xu, Z.-X. (2015). Recent advances in the study of HPV-associated carcinogenesis. Virol. Sin. *30*, 101–106.

Johannsen, E., and Lambert, P.F. (2013). Epigenetics of human papillomaviruses. Virology *445*, 205–212.

Johansson, C., and Schwartz, S. (2013). 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.

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 14-3-3ζ. 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.

Kashima, I., Yamashita, A., Izumi, N., Kataoka, N., Morishita, R., Hoshino, S., Ohno, M., Dreyfuss, G., and Ohno, S. (2006). Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. Genes Dev. *20*, 355–367.

Kashima, I., Jonas, S., Jayachandran, U., Buchwald, G., Conti, E., Lupas, A.N., and Izaurralde, E. (2010). SMG6 interacts with the exon junction complex via two conserved EJC-binding motifs (EBMs) required for nonsense-mediated mRNA decay. Genes Dev. *24*, 2440–2450.

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.

Katzenellenbogen, R.A., Vliet-Gregg, P., Xu, M., and Galloway, D.A. (2009). NFX1-123 increases hTERT expression and telomerase activity posttranscriptionally in human papillomavirus type 16 E6 keratinocytes. J. Virol. *83*, 6446–6456.

Katzenellenbogen, R.A., Vliet-Gregg, P., Xu, M., and Galloway, D.A. (2010). Cytoplasmic poly(A) binding proteins regulate telomerase activity and cell growth in human papillomavirus type 16 E6-expressing keratinocytes. J. Virol. *84*, 12934–12944.

Kessis, T.D., Slebos, R.J., Nelson, W.G., Kastan, M.B., Plunkett, B.S., Han, S.M., Lorincz, A.T., Hedrick, L., and Cho, K.R. (1993). Human papillomavirus 16 E6 expression disrupts the p53-mediated cellular response to DNA damage. Proc. Natl. Acad. Sci. U. S. A. *90*, 3988–3992.

Kessler, T.A. (2017). Cervical Cancer: Prevention and Early Detection. Semin. Oncol. Nurs. *33*, 172–183.

Kim, K., Garner-Hamrick, P.A., Fisher, C., Lee, D., and Lambert, P.F. (2003). Methylation patterns of papillomavirus DNA, its influence on E2 function, and implications in viral infection. J. Virol. *77*, 12450–12459.

Klymenko, T., Hernandez-Lopez, H., MacDonald, A.I., Bodily, J.M., and Graham, S.V. (2016). Human Papillomavirus E2 Regulates SRSF3 (SRp20) To Promote Capsid Protein Expression in Infected Differentiated Keratinocytes. J. Virol. *90*, 5047–5058.

Klymenko, T., Gu, Q., Herbert, I., Stevenson, A., Iliev, V., Watkins, G., Pollock, C., Bhatia, R., Cuschieri, K., Herzyk, P., et al. (2017). RNASeq analysis of differentiated keratinocytes reveals a massive response to late events during human papillomavirus type 16 infection, including loss of epithelial barrier function. J. Virol. *91*, e01001-17.

Korzeniewski, N., Spardy, N., Duensing, A., and Duensing, S. (2011). Genomic instability and cancer: lessons learned from human papillomaviruses. Cancer Lett. *305*, 113–122.

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. Cytopathol. *18*, 290–299.

Kota, K.P., Wagner, S.R., Huerta, E., Underwood, J.M., and Nickerson, J.A. (2008). Binding of ATP to UAP56 is necessary for mRNA export. J. Cell Sci. *121*, 1526–1537.

Kozak, M. (2001). Constraints on reinitiation of translation in mammals. Nucleic Acids Res. *29*, 5226–5232.

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.

Kukimoto, I., and Kanda, T. (2001). Displacement of YY1 by differentiation-specific transcription factor hSkn-1a activates the P(670) promoter of human papillomavirus type 16. J. Virol. *75*, 9302–9311.

Kukimoto, I., Takeuchi, T., and Kanda, T. (2006). CCAAT/enhancer binding protein beta binds to and activates the P670 promoter of human papillomavirus type 16. Virology *346*, 98–107.

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., Turek, L.P., and Haugen, T.H. (2008). Functional mapping of the human papillomavirus type 16 E1 cistron. J. Virol. *82*, 10724–10734.

Lace, M.J., Yamakawa, Y., Ushikai, M., Anson, J.R., Haugen, T.H., and Turek, L.P. (2009). Cellular factor YY1 downregulates the human papillomavirus 16 E6/E7 promoter, P97, in vivo and in vitro from a negative element overlapping the transcription-initiation site. J. Gen. Virol. *90*, 2402–2412.

Laurson, J., Khan, S., Chung, R., Cross, K., and Raj, K. (2010). Epigenetic repression of E-cadherin by human papillomavirus 16 E7 protein. Carcinogenesis *31*, 918–926.

Le Hir, H., Moore, M.J., and Maquat, L.E. (2000a). Pre-mRNA splicing alters mRNP composition: evidence for stable association of proteins at exon-exon junctions. Genes Dev. *14*, 1098–1108.

Le Hir, H., Izaurralde, E., Maquat, L.E., and Moore, M.J. (2000b). The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. EMBO J. *19*, 6860–6869.

Le Hir, H., Gatfield, D., Izaurralde, E., and Moore, M.J. (2001). 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.

Le Hir, H., Saulière, J., and Wang, Z. (2016). The exon junction complex as a node of post-transcriptional networks. Nat. Rev. Mol. Cell Biol. *17*, 41–54.

LeBlanc, J.J., and Beemon, K.L. (2004). Unspliced Rous sarcoma virus genomic RNAs are translated and subjected to nonsense-mediated mRNA decay before packaging. J. Virol. *78*, 5139–5146.

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

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.

Lehtinen, M., Paavonen, J., Wheeler, C.M., Jaisamrarn, U., Garland, S.M., Castellsagué, X., Skinner, S.R., Apter, D., Naud, P., Salmerón, J., et al. (2012). Overall efficacy of HPV-16/18 AS04-adjuvanted vaccine against grade 3 or greater cervical intraepithelial neoplasia: 4-year end-of-study analysis of the randomised, double-blind PATRICIA trial. Lancet Oncol. *13*, 89–99.

Lejeune, F. (2017). Nonsense-mediated mRNA decay at the crossroads of many cellular pathways. BMB Rep. *50*, 175–185.

Lejeune, F., Ishigaki, Y., Li, X., and Maquat, L.E. (2002). 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.

Li, X., Johansson, C., Cardoso Palacios, C., Mossberg, A., Dhanjal, S., Bergvall, M., and Schwartz, S. (2013a). Eight nucleotide substitutions inhibit splicing to HPV-16 3'-splice site SA3358 and reduce the

efficiency by which HPV-16 increases the life span of primary human keratinocytes. PloS One *8*, e72776.

Li, X., Johansson, C., Glahder, J., Mossberg, A.-K., and Schwartz, S. (2013b). Suppression of HPV-16 late L1 5'-splice site SD3632 by binding of hnRNP D proteins and hnRNP A2/B1 to upstream AUAGUA RNA motifs. Nucleic Acids Res. *41*, 10488–10508.

Li, X., Xu, Q., Wu, Y., Li, J., Tang, D., Han, L., and Fan, Q. (2014). A CCL2/ROS autoregulation loop is critical for cancer-associated fibroblasts-enhanced tumor growth of oral squamous cell carcinoma. Carcinogenesis *35*, 1362–1370.

Li, Z., Vuong, J.K., Zhang, M., Stork, C., and Zheng, S. (2017). Inhibition of nonsense-mediated RNA decay by ER stress. RNA *23*, 378–394.

Linde, L., Boelz, S., Neu-Yilik, G., Kulozik, A.E., and Kerem, B. (2007). The efficiency of nonsense-mediated mRNA decay is an inherent character and varies among different cells. Eur. J. Hum. Genet. *15*, 1156–1162.

Lindeboom, R.G.H., Supek, F., and Lehner, B. (2016). The rules and impact of nonsense-mediated mRNA decay in human cancers. Nat. Genet. *48*, 1112–1118.

Liu, C., Karam, R., Zhou, Y., Su, F., Ji, Y., Li, G., Xu, G., Lu, L., Wang, C., Song, M., et al. (2014a). The UPF1 RNA surveillance gene is commonly mutated in pancreatic adenosquamous carcinoma. Nat. Med. *20*, 596–598.

Liu, H.C., Chen, G.G., Vlantis, A.C., Tse, G.M., Chan, A.T.C., and van Hasselt, C.A. (2008). Inhibition of apoptosis in human laryngeal cancer cells by E6 and E7 oncoproteins of human papillomavirus 16. J. Cell. Biochem. *103*, 1125–1143.

Liu, M., Li, Y., Liu, A., Li, R., Su, Y., Du, J., Li, C., and Zhu, A.J. (2016). The exon junction complex regulates the splicing of cell polarity gene dlg1 to control Wingless signaling in development. ELife *5*.

Liu, W., Gao, G., Hu, X., Wang, Y., Schwarz, J.K., Chen, J.J., Grigsby, P.W., and Wang, X. (2014b). Activation of miR-9 by human papillomavirus in cervical cancer. Oncotarget *5*, 11620–11630.

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. U. S. A. *106*, 18780–18785.

Liu, Y., Wang, Y., Chen, C., Zhang, J., Qian, W., Dong, Y., Liu, Z., Zhang, X., Wang, X., Zhang, Z., et al. (2017). LSD1 binds to HPV16 E7 and promotes the epithelial-mesenchymal transition in cervical cancer by demethylating histones at the Vimentin promoter. Oncotarget *8*, 11329–11342.

Loh, B., Jonas, S., and Izaurralde, E. (2013). The SMG5-SMG7 heterodimer directly recruits the CCR4-NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2. Genes Dev. *27*, 2125–2138.

Longman, D., Johnstone, I.L., and Cáceres, J.F. (2003). The Ref/Aly proteins are dispensable for mRNA export and development in Caenorhabditis elegans. RNA *9*, 881–891.

Lu, Z., Hu, X., Li, Y., Zheng, L., Zhou, Y., Jiang, H., Ning, T., Basang, Z., Zhang, C., and Ke, Y. (2004). Human papillomavirus 16 E6 oncoprotein interferences with insulin signaling pathway by binding to tuberin. J. Biol. Chem. *279*, 35664–35670.

Luco, R.F., Allo, M., Schor, I.E., Kornblihtt, A.R., and Misteli, T. (2011). Epigenetics in alternative pre-mRNA splicing. Cell *144*, 16–26.

Luo, M.L., Zhou, Z., Magni, K., Christoforides, C., Rappsilber, J., Mann, M., and Reed, R. (2001). Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly. Nature *413*, 644–647.

Ma, X.M., Yoon, S.-O., Richardson, C.J., Jülich, K., and Blenis, J. (2008). SKAR links pre-mRNA splicing to mTOR/S6K1-mediated enhanced translation efficiency of spliced mRNAs. Cell *133*, 303–313.

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.

Mallon, R.G., Wojciechowicz, D., and Defendi, V. (1987). DNA-binding activity of papillomavirus proteins. J. Virol. *61*, 1655–1660.

Malone, C.D., Mestdagh, C., Akhtar, J., Kreim, N., Deinhard, P., Sachidanandam, R., Treisman, J., and Roignant, J.-Y. (2014). The exon junction complex controls transposable element activity by ensuring faithful splicing of the piwi transcript. Genes Dev. *28*, 1786–1799.

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.

Mao, H., Brown, H.E., and Silver, D.L. (2017). Mouse models of Casc3 reveal developmental functions distinct from other components of the exon junction complex. RNA *23*, 23–31.

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., Martines-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.

Martinez-Zapien, D., Ruiz, F.X., Poirson, J., Mitschler, A., Ramirez, J., Forster, A., Cousido-Siah, A., Masson, M., Vande Pol, S., Podjarny, A., et al. (2016). Structure of the E6/E6AP/p53 complex required for HPV-mediated degradation of p53. Nature *529*, 541–545.

Matsuda, D., Hosoda, N., Kim, Y.K., and Maquat, L.E. (2007). Failsafe nonsense-mediated mRNA decay does not detectably target eIF4E-bound mRNA. Nat. Struct. Mol. Biol. *14*, 974–979.

McBride, A.A. (2013). The papillomavirus E2 proteins. Virology *445*, 57–79.

McBride, A.A. (2017). Mechanisms and strategies of papillomavirus replication. Biol. Chem. *398*, 919–927.

McFarlane, M., MacDonald, A.I., Stevenson, A., and Graham, S.V. (2015). Human papillomavirus type 16 oncoprotein expression is controlled by the cellular splicing factor SRSF2 (SC35). J. Virol. *89*, 5276-87.

McGlincy, N.J., Tan, L.-Y., Paul, N., Zavolan, M., Lilley, K.S., and Smith, C.W.J. (2010). Expression proteomics of UPF1 knockdown in HeLa cells reveals autoregulation of hnRNP A2/B1 mediated by alternative splicing resulting in nonsense-mediated mRNA decay. BMC Genomics *11*, 565.

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.

McPhillips, M.G., Veerapraditsin, T., Cumming, S.A., Karali, D., Milligan, S.G., Boner, W., Morgan, I.M., and Graham, S.V. (2004). SF2/ASF binds the human papillomavirus type 16 late RNA control element and is regulated during differentiation of virus-infected epithelial cells. J. Virol. *78*, 10598–10605.

Mejilla, A., Li, E., and Sadowski, C.A. (2017). Human papilloma virus (HPV) vaccination: Questions and answers. Can. Pharm. J. *150*, 306–315.

Mendell, J.T., Sharifi, N.A., Meyers, J.L., Martinez-Murillo, F., and Dietz, H.C. (2004). Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. Nat. Genet. *36*, 1073–1078.

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.

Metze, S., Herzog, V.A., Ruepp, M.-D., and Mühlemann, O. (2013). Comparison of EJC-enhanced and EJC-independent NMD in human cells reveals two partially redundant degradation pathways. RNA *19*, 1432–1448.

Michelle, L., Cloutier, A., Toutant, J., Shkreta, L., Thibault, P., Durand, M., Garneau, D., Gendron, D., Lapointe, E., Couture, S., et al. (2012). Proteins associated with the exon junction complex also control the alternative splicing of apoptotic regulators. Mol. Cell. Biol. *32*, 954–967.

Miranda, S., Chaignot, C., Collin, C., Dray-Spira, R., Weill, A., and Zureik, M. (2017). Human papillomavirus vaccination and risk of autoimmune diseases: A large cohort study of over 2million young girls in France. Vaccine *35*, 4761–4768.

Mirghani, H., Amen, F., Tao, Y., Deutsch, E., and Levy, A. (2015). Increased radiosensitivity of HPV-positive head and neck cancers: Molecular basis and therapeutic perspectives. Cancer Treat. Rev. *41*, 844–852.

Mishler, D.M., Christ, A.B., and Steitz, J.A. (2008). Flexibility in the site of exon junction complex deposition revealed by functional group and RNA secondary structure alterations in the splicing substrate. RNA *14*, 2657–2670.

Mittal, S., and Banks, L. (2017). Molecular mechanisms underlying human papillomavirus E6 and E7 oncoprotein-induced cell transformation. Mutat. Res. *772*, 23–35.

Mocquet, V., Neusiedler, J., Rende, F., Cluet, D., Robin, J.-P., Terme, J.-M., Duc Dodon, M., Wittmann, J., Morris, C., Le Hir, H., et al. (2012). 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.

Monnier-Benoit, S., Mauny, F., Riethmuller, D., Guerrini, J.-S., Căpîlna, M., Félix, S., Seillès, E., Mougin, C., and Prétet, J.-L. (2006). Immunohistochemical analysis of CD4+ and CD8+ T-cell subsets in high risk human papillomavirus-associated pre-malignant and malignant lesions of the uterine cervix. Gynecol. Oncol. *102*, 22–31.

Moody, C.A., and Laimins, L.A. (2010). Human papillomavirus oncoproteins: pathways to transformation. Nat. Rev. Cancer *10*, 550–560.

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.

Mori, S., Takeuchi, T., Ishii, Y., Yugawa, T., Kiyono, T., Nishina, H., and Kukimoto, I. (2017). Human Papillomavirus 16 E6 Upregulates APOBEC3B via the TEAD Transcription Factor. J. Virol. *91*, e02413-16.

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.

Murphy-Marshman, H., Quensel, K., Shi-Wen, X., Barnfield, R., Kelly, J., Peidl, A., Stratton, R.J., and Leask, A. (2017). Antioxidants and NOX1/NOX4 inhibition blocks TGFβ1-induced CCN2 and α-SMA expression in dermal and gingival fibroblasts. PloS One *12*, e0186740.

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.

Nakagawa, M., Stites, D.P., Patel, S., Farhat, S., Scott, M., Hills, N.K., Palefsky, J.M., and Moscicki, A.B. (2000). Persistence of human papillomavirus type 16 infection is associated with lack of cytotoxic T lymphocyte response to the E6 antigens. J. Infect. Dis. *182*, 595–598.

Nakano, K., Ando, T., Yamagishi, M., Yokoyama, K., Ishida, T., Ohsugi, T., Tanaka, Y., Brighty, D.W., and Watanabe, T. (2013). Viral interference with host mRNA surveillance, the nonsense-mediated mRNA decay (NMD) pathway, through a new function of HTLV-1 Rex: implications for retroviral replication. Microbes Infect. *15*, 491–505.

Narisawa-Saito, M., Handa, K., Yugawa, T., Ohno, S., Fujita, M., and Kiyono, T. (2007). HPV16 E6-mediated stabilization of ErbB2 in neoplastic transformation of human cervical keratinocytes. Oncogene *26*, 2988–2996.

Neerman-Arbez, M., Germanos-Haddad, M., Tzanidakis, K., Vu, D., Deutsch, S., David, A., Morris, M.A., and de Moerloose, P. (2004). Expression and analysis of a split premature termination codon in FGG responsible for congenital afibrinogenemia: escape from RNA surveillance mechanisms in transfected cells. Blood *104*, 3618–3623.

Nguyen, C.L., and Münger, K. (2009). Human papillomavirus E7 protein deregulates mitosis via an association with nuclear mitotic apparatus protein 1. J. Virol. *83*, 1700–1707.

Nguyen, C.L., Eichwald, C., Nibert, M.L., and Münger, K. (2007). Human papillomavirus type 16 E7 oncoprotein associates with the centrosomal component gamma-tubulin. J. Virol. *81*, 13533–13543.

Nguyen, C.L., McLaughlin-Drubin, M.E., and Münger, K. (2008). Delocalization of the microtubule motor Dynein from mitotic spindles by the human papillomavirus E7 oncoprotein is not sufficient for induction of multipolar mitoses. Cancer Res. *68*, 8715–8722.

Nguyen, L.S., Wilkinson, M.F., and Gecz, J. (2014). Nonsense-mediated mRNA decay: inter-individual variability and human disease. Neurosci. Biobehav. Rev. *46 Pt 2*, 175–186.

Nickless, A., Jackson, E., Marasa, J., Nugent, P., Mercer, R.W., Piwnica-Worms, D., and You, Z. (2014). Intracellular calcium regulates nonsense-mediated mRNA decay. Nat. Med. *20*, 961–966.

Nickless, A., Bailis, J.M., and You, Z. (2017). Control of gene expression through the nonsense-mediated RNA decay pathway. Cell Biosci. *7*, 26.

Nickson, C.M., Moori, P., Carter, R.J., Rubbi, C.P., and Parsons, J.L. (2017). Misregulation of DNA damage repair pathways in HPV-positive head and neck squamous cell carcinoma contributes to cellular radiosensitivity. Oncotarget *8*, 29963–29975.

Nieborowska-Skorska, M., Kopinski, P.K., Ray, R., Hoser, G., Ngaba, D., Flis, S., Cramer, K., Reddy, M.M., Koptyra, M., Penserga, T., et al. (2012). Rac2-MRC-cIII-generated ROS cause genomic instability in chronic myeloid leukemia stem cells and primitive progenitors. Blood *119*, 4253–4263.

Nielsen, K.H., Chamieh, H., Andersen, C.B.F., Fredslund, F., Hamborg, K., Le Hir, H., and Andersen, G.R. (2009). Mechanism of ATP turnover inhibition in the EJC. RNA *15*, 67–75.

Nilsson, K., Norberg, C., Mossberg, A.-K., and Schwartz, S. (2018). HPV16 E5 is produced from an HPV16 early mRNA spliced from SD226 to SA3358. Virus Res. *244*, 128–136.

Nojima, T., Hirose, T., Kimura, H., and Hagiwara, M. (2007). The interaction between cap-binding complex and RNA export factor is required for intronless mRNA export. J. Biol. Chem. *282*, 15645–15651.

Nott, A., Le Hir, H., and Moore, M.J. (2004). Splicing enhances translation in mammalian cells: an additional function of the exon junction complex. Genes Dev. *18*, 210–222.

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.

O'Connor, M.J., Stünkel, W., Koh, C.H., Zimmermann, H., and Bernard, H.U. (2000). The differentiation-specific factor CDP/Cut represses transcription and replication of human papillomaviruses through a conserved silencing element. J. Virol. *74*, 401–410.

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.

Ohnishi, T., Yamashita, A., Kashima, I., Schell, T., Anders, K.R., Grimson, A., Hachiya, T., Hentze, M.W., Anderson, P., and Ohno, S. (2003). Phosphorylation of hUPF1 induces formation of mRNA surveillance complexes containing hSMG-5 and hSMG-7. Mol. Cell *12*, 1187–1200.

Okada-Katsuhata, Y., Yamashita, A., Kutsuzawa, K., Izumi, N., Hirahara, F., and Ohno, S. (2012). N- and C-terminal Upf1 phosphorylations create binding platforms for SMG-6 and SMG-5:SMG-7 during NMD. Nucleic Acids Res. *40*, 1251–1266.

Oliver, A.W., He, X., Borthwick, K., Donne, A.J., Hampson, L., and Hampson, I.N. (2011). The HPV16 E6 binding protein Tip-1 interacts with ARHGEF16, which activates Cdc42. Br. J. Cancer *104*, 324–331.

Ottinger, M., Smith, J.A., Schweiger, M.-R., Robbins, D., Powell, M.L.C., You, J., and Howley, P.M. (2009). Cell-type specific transcriptional activities among different papillomavirus long control regions and their regulation by E2. Virology *395*, 161–171.

Palacios, I.M. (2013). Nonsense-mediated mRNA decay: from mechanistic insights to impacts on human health. Brief. Funct. Genomics *12*, 25–36.

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.

Paracha, U.Z., Fatima, K., Alqahtani, M., Chaudhary, A., Abuzenadah, A., Damanhouri, G., and Qadri, I. (2013). Oxidative stress and hepatitis C virus. Virol. J. *10*, 251.

Paris, C., Pentland, I., Groves, I., Roberts, D.C., Powis, S.J., Coleman, N., Roberts, S., and Parish, J.L. (2015). CCCTC-binding factor recruitment to the early region of the human papillomavirus 18 genome regulates viral oncogene expression. J. Virol. *89*, 4770–4785.

Park, J.S., Rhyu, J.W., Kim, C.J., Kim, H.S., Lee, S.Y., Kwon, Y.I., Namkoong, S.E., Sin, H.S., and Um, S.J. (2003). Neoplastic change of squamo-columnar junction in uterine cervix and vaginal epithelium by exogenous estrogen in hpv-18 URR E6/E7 transgenic mice. Gynecol. Oncol. *89*, 360–368.

Pastuszak-Lewandoska, D., Bartosińska-Dyc, A., Migdalska-Sęk, M., Czarnecka, K.H., Nawrot, E., Domańska, D., Szyłło, K., and Brzeziańska, E. (2014). HPV16 E6*II gene expression in intraepithelial cervical lesions as an indicator of neoplastic grade: a pilot study. Med. Oncol. *31*, 842.

Patel, S., and Chiplunkar, S. (2009). Host immune responses to cervical cancer. Curr. Opin. Obstet. Gynecol. *21*, 54–59.

Patel, D., Huang, S.M., Baglia, L.A., and McCance, D.J. (1999). The E6 protein of human papillomavirus type 16 binds to and inhibits co-activation by CBP and p300. EMBO J. *18*, 5061–5072.

Pereverzev, A.P., Gurskaya, N.G., Ermakova, G.V., Kudryavtseva, E.I., Markina, N.M., Kotlobay, A.A., Lukyanov, S.A., Zaraisky, A.G., and Lukyanov, K.A. (2015). Method for quantitative analysis of nonsense-mediated mRNA decay at the single cell level. Sci. Rep. *5*, 7729.

Pett, M., and Coleman, N. (2007). Integration of high-risk human papillomavirus: a key event in cervical carcinogenesis? J. Pathol. *212*, 356–367.

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. (1997). 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.

Piva, F., Giulietti, M., Burini, A.B., and Principato, G. (2012). SpliceAid 2: a database of human splicing factors expression data and RNA target motifs. Hum. Mutat. *33*, 81–85.

Popp, M.W., and Maquat, L.E. (2018). Nonsense-mediated mRNA Decay and Cancer. Curr. Opin. Genet. Dev. *48*, 44–50.

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.

Quek, B.L., and Beemon, K. (2014). Retroviral strategy to stabilize viral RNA. Curr. Opin. Microbiol. *18*, 78–82.

Radisky, D.C., Levy, D.D., Littlepage, L.E., Liu, H., Nelson, C.M., Fata, J.E., Leake, D., Godden, E.L., Albertson, D.G., Nieto, M.A., et al. (2005). Rac1b and reactive oxygen species mediate MMP-3-induced EMT and genomic instability. Nature *436*, 123–127.

Ramage, H.R., Kumar, G.R., Verschueren, E., Johnson, J.R., Von Dollen, J., Johnson, T., Newton, B., Shah, P., Horner, J., Krogan, N.J., et al. (2015). A combined proteomics/genomics approach links hepatitis C virus infection with nonsense-mediated mRNA decay. Mol. Cell *57*, 329–340.

Rebbapragada, I., and Lykke-Andersen, J. (2009). Execution of nonsense-mediated mRNA decay: what defines a substrate? Curr. Opin. Cell Biol. *21*, 394–402.

Reiser, J., Hurst, J., Voges, M., Krauss, P., Münch, P., Iftner, T., and Stubenrauch, F. (2011). High-risk human papillomaviruses repress constitutive kappa interferon transcription via E6 to prevent pathogen recognition receptor and antiviral-gene expression. J. Virol. *85*, 11372–11380.

Reshi, M.L., Su, Y.-C., and Hong, J.-R. (2014). RNA Viruses: ROS-Mediated Cell Death. Int. J. Cell Biol. *2014*, 467452.

Reuter, S., Gupta, S.C., Chaturvedi, M.M., and Aggarwal, B.B. (2010). Oxidative stress, inflammation, and cancer: how are they linked? Free Radic. Biol. Med. *49*, 1603–1616.

Rieckmann, T., Tribius, S., Grob, T.J., Meyer, F., Busch, C.-J., Petersen, C., Dikomey, E., and Kriegs, M. (2013). HNSCC cell lines positive for HPV and p16 possess higher cellular radiosensitivity due to an impaired DSB repair capacity. Radiother. Oncol. J. *107*, 242–246.

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.

Riley, T., Sontag, E., Chen, P., and Levine, A. (2008). Transcriptional control of human p53-regulated genes. Nat. Rev. Mol. Cell Biol. *9*, 402–412.

Ristriani, T., Masson, M., Nominé, Y., Laurent, C., Lefevre, J.F., Weiss, E., and Travé, G. (2000). HPV oncoprotein E6 is a structure-dependent DNA-binding protein that recognizes four-way junctions. J. Mol. Biol. *296*, 1189–1203.

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.

Roignant, J.-Y., and Treisman, J.E. (2010). Exon junction complex subunits are required to splice Drosophila MAP kinase, a large heterochromatic gene. Cell *143*, 238–250.

Roman, A., and Munger, K. (2013). The papillomavirus E7 proteins. Virology *445*, 138–168.

Rosenberger, S., De-Castro Arce, J., 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.

Rosenstierne, M.W., Vinther, J., Hansen, C.N., Prydsoe, M., and Norrild, B. (2003). Identification and characterization of a cluster of transcription start sites located in the E6 ORF of human papillomavirus type 16. J. Gen. Virol. *84*, 2909–2920.

Ryser, M.D., Myers, E.R., and Durrett, R. (2015). HPV clearance and the neglected role of stochasticity. PLoS Comput. Biol. *11*, e1004113.

Sablina, A.A., Budanov, A.V., Ilyinskaya, G.V., Agapova, L.S., Kravchenko, J.E., and Chumakov, P.M. (2005). The antioxidant function of the p53 tumor suppressor. Nat. Med. *11*, 1306–1313.

Saulière, J., Haque, N., Harms, S., Barbosa, I., Blanchette, M., and Le Hir, H. (2010). The exon junction complex differentially marks spliced junctions. Nat. Struct. Mol. Biol. *17*, 1269–1271.

Saulière, J., Murigneux, V., Wang, Z., Marquenet, E., Barbosa, I., Le Tonquèze, O., Audic, Y., Paillard, L., Roest Crollius, H., and Le Hir, H. (2012). CLIP-seq of eIF4AIII reveals transcriptome-wide mapping of the human exon junction complex. Nat. Struct. Mol. Biol. *19*, 1124–1131.

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.

Schieber, M., and Chandel, N.S. (2014). ROS function in redox signaling and oxidative stress. Curr. Biol. CB *24*, R453-462.

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*, 16086.

Schmitt, M., Dalstein, V., Waterboer, T., Clavel, C., Gissmann, L., and Pawlita, M. (2010). Diagnosing cervical cancer and high-grade precursors by HPV16 transcription patterns. Cancer Res. *70*, 249–256.

Schmitt, M., Dalstein, V., Waterboer, T., Clavel, C., Gissmann, L., and Pawlita, M. (2011). The HPV16 transcriptome in cervical lesions of different grades. Mol. Cell. Probes *25*, 260–265.

Schneider-Gädicke, A., Kaul, S., Schwarz, E., Gausepohl, H., Frank, R., and Bastert, G. (1988). 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.

van de Schootbrugge, C., Schults, E.M.J., Bussink, J., Span, P.N., Grénman, R., Pruijn, G.J.M., Kaanders, J.H.A.M., and Boelens, W.C. (2014). Effect of hypoxia on the expression of αB-crystallin in head and neck squamous cell carcinoma. BMC Cancer *14*, 252.

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.

Severino, A., Abbruzzese, C., Manente, L., Valderas, A.A., Mattarocci, S., Federico, A., Starace, G., Chersi, A., Mileo, A.M., and Paggi, M.G. (2007). Human papillomavirus-16 E7 interacts with Siva-1 and modulates apoptosis in HaCaT human immortalized keratinocytes. J. Cell. Physiol. *212*, 118–125.

Shi, Y., and Manley, J.L. (2015). The end of the message: multiple protein-RNA interactions define the mRNA polyadenylation site. Genes Dev. *29*, 889–897.

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.

Sichero, L., Sobrinho, J.S., and Villa, L.L. (2012). Identification of novel cellular transcription factors that regulate early promoters of human papillomavirus types 18 and 16. J. Infect. Dis. *206*, 867–874.

Singh, G., Kucukural, A., Cenik, C., Leszyk, J.D., Shaffer, S.A., Weng, Z., and Moore, M.J. (2012). The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. Cell *151*, 750–764.

Singh, K.K., Wachsmuth, L., Kulozik, A.E., and Gehring, N.H. (2013). Two mammalian MAGOH genes contribute to exon junction complex composition and nonsense-mediated decay. RNA Biol. *10*, 1291–1298.

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.

Slebos, R.J.C., Yi, Y., Ely, K., Carter, J., Evjen, A., Zhang, X., Shyr, Y., Murphy, B.M., Cmelak, A.J., Burkey, B.B., et al. (2006). Gene expression differences associated with human papillomavirus status in head and neck squamous cell carcinoma. Clin. Cancer Res. *12*, 701–709.

Sloan, K.E., Gleizes, P.-E., and Bohnsack, M.T. (2016). Nucleocytoplasmic Transport of RNAs and RNA-Protein Complexes. J. Mol. Biol. *428*, 2040–2059.

Smith, J.A., Haberstroh, F.S., White, E.A., Livingston, D.M., DeCaprio, J.A., and Howley, P.M. (2014). SMCX and components of the TIP60 complex contribute to E2 regulation of the HPV E6/E7 promoter. Virology *468–470*, 311–321.

Smola, S., Trimble, C., and Stern, P.L. (2017). Human papillomavirus-driven immune deviation: challenge and novel opportunity for immunotherapy. Ther. Adv. Vaccines *5*, 69–82.

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.

Sokolowski, M., Tan, W., Jellne, M., and Schwartz, S. (1998). mRNA instability elements in the human papillomavirus type 16 L2 coding region. J. Virol. *72*, 1504–1515.

Somberg, M., and Schwartz, S. (2010). Multiple ASF/SF2 sites in the human papillomavirus type 16 (HPV-16) E4-coding region promote splicing to the most commonly used 3'-splice site on the HPV-16 genome. J. Virol. *84*, 8219–8230.

Somberg, M., Li, X., Johansson, C., Orru, B., Chang, R., Rush, M., Fay, J., Ryan, F., and Schwartz, S. (2011). Serine/arginine-rich protein 30c activates human papillomavirus type 16 L1 mRNA expression via a bimodal mechanism. J. Gen. Virol. *92*, 2411–2421.

Song, L., Liu, S., Zeng, S., Zhang, L., and Li, X. (2015). miR-375 Modulates Radiosensitivity of HR-HPV-Positive Cervical Cancer Cells by Targeting UBE3A through the p53 Pathway. Med. Sci. Monit. *21*, 2210–2217.

Songock, W.K., Kim, S.-M., and Bodily, J.M. (2017a). The human papillomavirus E7 oncoprotein as a regulator of transcription. Virus Res. *231*, 56–75.

Songock, W.K., Scott, M.L., and Bodily, J.M. (2017b). Regulation of the human papillomavirus type 16 late promoter by transcriptional elongation. Virology *507*, 179–191.

Spangle, J.M., and Münger, K. (2010). The human papillomavirus type 16 E6 oncoprotein activates mTORC1 signaling and increases protein synthesis. J. Virol. *84*, 9398–9407.

Spangle, J.M., and Munger, K. (2013). The HPV16 E6 oncoprotein causes prolonged receptor protein tyrosine kinase signaling and enhances internalization of phosphorylated receptor species. PLoS Pathog. *9*, e1003237.

Spanos, W.C., Hoover, A., Harris, G.F., Wu, S., Strand, G.L., Anderson, M.E., Klingelhutz, 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.

Spector, D.L., and Lamond, A.I. (2011). Nuclear speckles. Cold Spring Harb. Perspect. Biol. *3*.

Srivenugopal, K.S., and Ali-Osman, F. (2002). The DNA repair protein, O(6)-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.

Steckelberg, A.-L., Boehm, V., Gromadzka, A.M., and Gehring, N.H. (2012). CWC22 connects pre-mRNA splicing and exon junction complex assembly. Cell Rep. *2*, 454–461.

Steckelberg, A.-L., Altmueller, J., Dieterich, C., and Gehring, N.H. (2015). CWC22-dependent pre-mRNA splicing and eIF4A3 binding enables global deposition of exon junction complexes. Nucleic Acids Res. *43*, 4687–4700.

Stockklausner, C., Breit, S., Neu-Yilik, G., Echner, N., Hentze, M.W., Kulozik, A.E., and Gehring, N.H. (2006). The uORF-containing thrombopoietin mRNA escapes nonsense-mediated decay (NMD). Nucleic Acids Res. *34*, 2355–2363.

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.

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.

Stünkel, W., Huang, Z., Tan, S.H., O'Connor, M.J., and Bernard, H.U. (2000). Nuclear matrix attachment regions of human papillomavirus type 16 repress or activate the E6 promoter, depending on the physical state of the viral DNA. J. Virol. *74*, 2489–2501.

Sureau, A., Gattoni, R., Dooghe, Y., Stévenin, J., and Soret, J. (2001). SC35 autoregulates its expression by promoting splicing events that destabilize its mRNAs. EMBO J. *20*, 1785–1796.

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, S.H., Bartsch, D., Schwarz, E., and Bernard, H.U. (1998). Nuclear matrix attachment regions of human papillomavirus type 16 point toward conservation of these genomic elements in all genital papillomaviruses. J. Virol. *72*, 3610–3622.

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.

Teng, I.-F., and Wilson, S.A. (2013). Mapping interactions between mRNA export factors in living cells. PloS One *8*, e67676.

Thomas, M.C., and Chiang, C.-M. (2005). E6 oncoprotein represses p53-dependent gene activation via inhibition of protein acetylation independently of inducing p53 degradation. Mol. Cell *17*, 251–264.

Thomas, M., Massimi, P., Jenkins, J., and Banks, L. (1995). HPV-18 E6 mediated inhibition of p53 DNA binding activity is independent of E6 induced degradation. Oncogene *10*, 261–268.

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.

Thompson, D.A., Belinsky, G., Chang, T.H., Jones, D.L., Schlegel, R., and Münger, K. (1997). The human papillomavirus-16 E6 oncoprotein decreases the vigilance of mitotic checkpoints. Oncogene *15*, 3025–3035.

Thompson, D.A., Zacny, V., Belinsky, G.S., Classon, M., Jones, D.L., Schlegel, R., and Münger, K. (2001). The HPV E7 oncoprotein inhibits tumor necrosis factor alpha-mediated apoptosis in normal human fibroblasts. Oncogene *20*, 3629–3640.

Tommasino, M. (2017). The biology of beta human papillomaviruses. Virus Res. *231*, 128–138.

Tong, X., and Howley, P.M. (1997). The bovine papillomavirus E6 oncoprotein interacts with paxillin and disrupts the actin cytoskeleton. Proc. Natl. Acad. Sci. U. S. A. *94*, 4412–4417.

Toots, M., Männik, A., Kivi, G., Ustav, M., 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., Neidigh, J.W., Fodor, N., and Duerksen-Hughes, P.J. (2008). The interaction between human papillomavirus type 16 and FADD is mediated by a novel E6 binding domain. J. Virol. *82*, 9600–9614.

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.

Underbrink, M.P., Dupuis, C., Wang, J., and Tyring, S.K. (2016). E6 proteins from low-risk human papillomavirus types 6 and 11 are able to protect keratinocytes from apoptosis via Bak degradation. J. Gen. Virol. *97*, 715–724.

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.

Valencia, P., Dias, A.P., and Reed, R. (2008). Splicing promotes rapid and efficient mRNA export in mammalian cells. Proc. Natl. Acad. Sci. U. S. A. *105*, 3386–3391.

Van Doorslaer, K. (2013). Evolution of the papillomaviridae. Virology *445*, 11–20.

Van Doorslaer, K., Li, Z., Xirasagar, S., Maes, P., Kaminsky, D., Liou, D., Sun, Q., Kaur, R., Huyen, Y., and McBride, A.A. (2017). The Papillomavirus Episteme: a major update to the papillomavirus sequence database. Nucleic Acids Res. *45*, D499–D506.

Vande Pol, S.B., and Klingelhutz, A.J. (2013). Papillomavirus E6 oncoproteins. Virology *445*, 115–137.

Veldman, T., Horikawa, I., Barrett, J.C., and Schlegel, R. (2001). Transcriptional activation of the telomerase hTERT gene by human papillomavirus type 16 E6 oncoprotein. J. Virol. *75*, 4467–4472.

Veldman, T., Liu, X., Yuan, H., and Schlegel, R. (2003). Human papillomavirus E6 and Myc proteins associate in vivo and bind to and cooperatively activate the telomerase reverse transcriptase promoter. Proc. Natl. Acad. Sci. U. S. A. *100*, 8211–8216.

Viegas, M.H., Gehring, N.H., Breit, S., Hentze, M.W., and Kulozik, A.E. (2007). The abundance of RNPS1, a protein component of the exon junction complex, can determine the variability in efficiency of the Nonsense Mediated Decay pathway. Nucleic Acids Res. *35*, 4542–4551.

Vieira, V.C., Leonard, B., White, E.A., Starrett, G.J., Temiz, N.A., Lorenz, L.D., Lee, D., Soares, M.A., Lambert, P.F., Howley, P.M., et al. (2014). Human papillomavirus E6 triggers upregulation of the antiviral and cancer genomic DNA deaminase APOBEC3B. MBio *5*.

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.

Vinther, J., Rosenstierne, M.W., Kristiansen, K., and Norrild, B. (2005). The 3' region of human papillomavirus type 16 early mRNAs decrease expression. BMC Infect. Dis. *5*, 83.

Wahl, M.C., Will, C.L., and Lührmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. Cell *136*, 701–718.

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*.

Wallin, K.L., Wiklund, F., Angström, T., Bergman, F., Stendahl, U., Wadell, G., Hallmans, G., and Dillner, J. (1999). Type-specific persistence of human papillomavirus DNA before the development of invasive cervical cancer. N. Engl. J. Med. *341*, 1633–1638.

Wang, Z., and Burge, C.B. (2008). Splicing regulation: from a parts list of regulatory elements to an integrated splicing code. RNA *14*, 802–813.

Wang, D., Wengrod, J., and Gardner, L.B. (2011a). Overexpression of the c-myc oncogene inhibits nonsense-mediated RNA decay in B lymphocytes. J. Biol. Chem. *286*, 40038–40043.

Wang, D., Zavadil, J., Martin, L., Parisi, F., Friedman, E., Levy, D., Harding, H., Ron, D., and Gardner, L.B. (2011b). Inhibition of nonsense-mediated RNA decay by the tumor microenvironment promotes tumorigenesis. Mol. Cell. Biol. *31*, 3670–3680.

Wang, S., Pang, T., Gao, M., Kang, H., Ding, W., Sun, X., Zhao, Y., Zhu, W., Tang, X., Yao, Y., et al. (2013). HPV E6 induces eIF4E transcription to promote the proliferation and migration of cervical cancer. FEBS Lett. *587*, 690–697.

Wang, S.S., Schiffman, M., Herrero, R., Carreon, J., Hildesheim, A., Rodriguez, A.C., Bratti, M.C., Sherman, M.E., Morales, J., Guillen, D., et al. (2004). Determinants of human papillomavirus 16 serological conversion and persistence in a population-based cohort of 10 000 women in Costa Rica. Br. J. Cancer *91*, 1269–1274.

Wang, X., Liu, H., Ge, H., Ajiro, M., Sharma, N.R., Meyers, C., Morozov, P., Tuschl, T., Klar, A., Court, D., et al. (2017). Viral DNA Replication Orientation and hnRNPs Regulate Transcription of the Human Papillomavirus 18 Late Promoter. MBio *8*.

Wang, Z., Murigneux, V., and Le Hir, H. (2014). Transcriptome-wide modulation of splicing by the exon junction complex. Genome Biol. *15*, 551.

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.

Weinberg, F., Hamanaka, R., Wheaton, W.W., Weinberg, S., Joseph, J., Lopez, M., Kalyanaraman, B., Mutlu, G.M., Budinger, G.R.S., and Chandel, N.S. (2010). Mitochondrial metabolism and ROS generation are essential for Kras-mediated tumorigenicity. Proc. Natl. Acad. Sci. U. S. A. *107*, 8788–8793.

Weng, S.-L., Huang, K.-Y., Weng, J.T.-Y., Hung, F.-Y., Chang, T.-H., and Lee, T.-Y. (2017). Genome-wide discovery of viral microRNAs based on phylogenetic analysis and structural evolution of various human papillomavirus subtypes. Brief. Bioinform. *19*, 1102-1114.

Westrich, J.A., Warren, C.J., and Pyeon, D. (2017). Evasion of host immune defenses by human papillomavirus. Virus Res. *231*, 21–33.

Wheeler, C.M., Castellsagué, X., Garland, S.M., Szarewski, A., Paavonen, J., Naud, P., Salmerón, J., Chow, S.-N., Apter, D., Kitchener, H., et al. (2012). Cross-protective efficacy of HPV-16/18 AS04-adjuvanted vaccine against cervical infection and precancer caused by non-vaccine oncogenic HPV types: 4-year end-of-study analysis of the randomised, double-blind PATRICIA trial. Lancet Oncol. *13*, 100–110.

Wiklund, L., Sokolowski, M., Carlsson, A., Rush, M., and Schwartz, S. (2002). Inhibition of translation by UAUUUAU and UAUUUUUAU motifs of the AU-rich RNA instability element in the HPV-1 late 3' untranslated region. J. Biol. Chem. *277*, 40462–40471.

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. (2011). 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.

Wittmann, J., Hol, E.M., and Jäck, H.-M. (2006). hUPF2 silencing identifies physiologic substrates of mammalian nonsense-mediated mRNA decay. Mol. Cell. Biol. *26*, 1272–1287.

Woo, Y.L., van den Hende, M., Sterling, J.C., Coleman, N., Crawford, R.A.F., Kwappenberg, K.M.C., Stanley, M.A., and van der Burg, S.H. (2010). A prospective study on the natural course of low-grade squamous intraepithelial lesions and the presence of HPV16 E2-, E6- and E7-specific T-cell responses. Int. J. Cancer *126*, 133–141.

Woodward, L.A., Mabin, J.W., Gangras, P., and Singh, G. (2017). The exon junction complex: a lifelong guardian of mRNA fate. Wiley Interdiscip. Rev. RNA *8*.

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.

Yaginuma, Y., Yoshimoto, M., Eguchi, A., Tokuda, A., and Takahashi, S. (2015). The human papillomavirus18 E7 protein inhibits CENP-C binding to α-satellite DNA. Virus Res. *205*, 27–32.

Yamashita, A., Izumi, N., Kashima, I., Ohnishi, T., Saari, B., Katsuhata, Y., Muramatsu, R., Morita, T., Iwamatsu, A., Hachiya, T., et al. (2009). SMG-8 and SMG-9, two novel subunits of the SMG-1 complex, regulate remodeling of the mRNA surveillance complex during nonsense-mediated mRNA decay. Genes Dev. *23*, 1091–1105.

Yepiskoposyan, H., Aeschimann, F., Nilsson, D., Okoniewski, M., and Mühlemann, O. (2011). Autoregulation of the nonsense-mediated mRNA decay pathway in human cells. RNA *17*, 2108–2118.

You, J. (2010). Papillomavirus interaction with cellular chromatin. Biochim. Biophys. Acta *1799*, 192–199.

Zanier, K., Nominé, Y., Charbonnier, S., Ruhlmann, C., Schultz, P., Schweizer, J., and Travé, G. (2007). Formation of well-defined soluble aggregates upon fusion to MBP is a generic property of E6 proteins from various human papillomavirus species. Protein Expr. Purif. *51*, 59–70.

Zanier, K., ould M'hamed ould Sidi, A., 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. Lond. *20*, 604–617.

Zhang, Z., and Krainer, A.R. (2007). Splicing remodels messenger ribonucleoprotein architecture via eIF4A3-dependent and -independent recruitment of exon junction complex components. Proc. Natl. Acad. Sci. U. S. A. *104*, 11574–11579.

Zhang, E., Feng, X., Liu, F., Zhang, P., Liang, J., and Tang, X. (2014). Roles of PI3K/Akt and c-Jun signaling pathways in human papillomavirus type 16 oncoprotein-induced HIF-1α, VEGF, and IL-8 expression and in vitro angiogenesis in non-small cell lung cancer cells. PloS One *9*, e103440.

Zhang, S., Leng, T., Zhang, Q., Zhao, Q., Nie, X., and Yang, L. (2018). Sanguinarine inhibits epithelial ovarian cancer development via regulating long non-coding RNA CASC2-EIF4A3 axis and/or inhibiting NF-κB signaling or PI3K/AKT/mTOR pathway. Biomed. Pharmacother. *102*, 302–308.

Zhao, K.-N., Liu, W.J., and Frazer, I.H. (2003). Codon usage bias and A+T content variation in human papillomavirus genomes. Virus Res. *98*, 95–104.

Zhao, X., Rush, M., and Schwartz, S. (2004). Identification of an hnRNP A1-dependent splicing silencer in the human papillomavirus type 16 L1 coding region that prevents premature expression of the late L1 gene. J. Virol. *78*, 10888–10905.

Zhao, X., Oberg, D., Rush, M., Fay, J., Lambkin, H., and Schwartz, S. (2005). A 57-nucleotide upstream early polyadenylation element in human papillomavirus type 16 interacts with hFip1, CstF-64, hnRNP C1/C2, and polypyrimidine tract binding protein. J. Virol. *79*, 4270–4288.

Zheng, Z.-M., and Baker, C.C. (2006). Papillomavirus genome structure, expression, and post-transcriptional regulation. Front. Biosci. J. Virtual Libr. *11*, 2286–2302.

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.

Zhou, J., Liu, W.J., Peng, S.W., Sun, X.Y., and Frazer, I. (1999). Papillomavirus capsid protein expression level depends on the match between codon usage and tRNA availability. J. Virol. *73*, 4972–4982.

Zhou, Z., Luo, M.J., Straesser, K., Katahira, J., Hurt, E., and Reed, R. (2000). The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. Nature *407*, 401–405.

Zimmermann, H., Degenkolbe, R., Bernard, H.U., and O'Connor, M.J. (1999). The human papillomavirus type 16 E6 oncoprotein can down-regulate p53 activity by targeting the transcriptional coactivator CBP/p300. J. Virol. *73*, 6209–6219.

Zünd, D., Gruber, A.R., Zavolan, M., and Mühlemann, O. (2013). Translation-dependent displacement of UPF1 from coding sequences causes its enrichment in 3' UTRs. Nat. Struct. Mol. Biol. *20*, 936–943.

Annexe 1 : liste des transcrits dérégulés dans les cellules U-E6All et U-E6*I.

Table S1. Differential expression analysis of transcripts between U-E6All vs. U-CTL cell lines.

Transcrit_id	Gene_id	Gene_name	A FPKM	B FPKM	Log2(FC)
ENST00000009589	ENSG00000008988	RPS20	2361,34	1102,28	-1,09911
ENST00000025301	ENSG00000023516	AKAP11	2,62378	6,9545	1,4063
ENST00000196551	ENSG00000083845	RPS5	1951,1	908,197	-1,10321
ENST00000200453	ENSG00000087074	PPP1R15A	20,7169	9,38523	-1,14235
ENST00000206380	ENSG00000091947	TMEM101	49,7023	21,7173	-1,19447
ENST00000215754	ENSG00000240972	MIF	563,804	195,905	-1,52504
ENST00000216862	ENSG00000019186	CYP24A1	2,62609	7,60931	1,53485
ENST00000217233	ENSG00000101255	TRIB3	38,2	10,2855	-1,89296
ENST00000218230	ENSG00000102109	PCSK1N	27,2429	9,92949	-1,45609
ENST00000221847	ENSG00000105246	EBI3	8,20531	1,8818	-2,12445
ENST00000221975	ENSG00000105372	RPS19	5160,69	1973,7	-1,38666
ENST00000222157	ENSG00000105550	FGF21	8,69545	0,700782	-3,63322
ENST00000222573	ENSG00000105855	ITGB8	2,74129	7,31338	1,41568
ENST00000223095	ENSG00000106366	SERPINE1	58,0148	28,3812	-1,03149
ENST00000225831	ENSG00000108691	CCL2	70,9086	25,8834	-1,45393
ENST00000228841	ENSG00000111245	MYL2	12,9444	2,9515	-2,13281
ENST00000230050	ENSG00000112306	RPS12	2834,93	1457,82	-0,959503
ENST00000230431	ENSG00000112667	DNPH1	56,2093	22,2783	-1,33517
ENST00000232375	ENSG00000114268	PFKFB4	14,2003	5,58867	-1,34535
ENST00000232564	ENSG00000114450	GNB4	3,66609	9,61124	1,39048
ENST00000232975	ENSG00000114854	TNNC1	92,872	24,1297	-1,94443
ENST00000233055	ENSG00000085449	WDFY1	7,58419	18,8853	1,3162
ENST00000233813	ENSG00000115461	IGFBP5	13,0762	30,705	1,23153
ENST00000234961	ENSG00000116329	OPRD1	0,0624199	0,384235	2,62191
ENST00000238497	ENSG00000119541	VPS4B	6,84318	15,2561	1,15665
ENST00000240093	ENSG00000104290	FZD3	1,36475	2,93609	1,10525
ENST00000242208	ENSG00000122641	INHBA	7,59196	18,9797	1,32191
ENST00000242719	ENSG00000123091	RNF11	14,1038	30,3561	1,1059
ENST00000243347	ENSG00000123610	TNFAIP6	0,21817	2,92225	3,74356
ENST00000245458	ENSG00000213741	RPS29	4649,18	1482,79	-1,64866
ENST00000245479	ENSG00000125398	SOX9	16,6628	35,7413	1,10096
ENST00000245903	ENSG00000125726	CD70	50,6619	12,7414	-1,99137
ENST00000246802	ENSG00000105373	GLTSCR2	177,927	46,5098	-1,93568
ENST00000247225	ENSG00000126821	SGPP1	7,14119	19,7861	1,47025
ENST00000248879	ENSG00000128185	DGCR6L	70,4003	29,9047	-1,23521
ENST00000250498	ENSG00000129562	DAD1	205,966	107,768	-0,934482
ENST00000251047	ENSG00000074695	LMAN1	27,7427	73,0824	1,39741
ENST00000251081	ENSG00000134755	DSC2	1,75768	4,03449	1,19871
ENST00000251453	ENSG00000105193	RPS16	2165,2	954,635	-1,18148
ENST00000251646	ENSG00000131738	KRT33B	1,62782	0,394929	-2,04328
ENST00000252245	ENSG00000170454	KRT75	4,81887	20,3312	2,07693
ENST00000252593	ENSG00000130303	BST2	25,3006	71,3998	1,49675
ENST00000252809	ENSG00000130513	GDF15	26,8643	0,822639	-5,02929
ENST00000253047	ENSG00000130748	TMEM160	22,4379	6,99892	-1,68074
ENST00000253063	ENSG00000130766	SESN2	18,0727	6,28334	-1,52421
ENST00000253407	ENSG00000131094	C1QL1	148,825	62,3478	-1,2552
ENST00000253788	ENSG00000131469	RPL27	3107,87	1255,06	-1,30816
ENST00000254508	ENSG00000132182	NUP210	9,85384	20,7028	1,07107
ENST00000255305	ENSG00000132953	XPO4	2,21474	4,65262	1,0709
ENST00000256108	ENSG00000133731	IMPA1	4,63976	15,4139	1,7321
ENST00000256646	ENSG00000134250	NOTCH2	15,1214	28,1767	0,897907
ENST00000257190	ENSG00000134758	RNF138	3,04334	14,2121	2,22339
ENST00000257287	ENSG00000174799	CEP135	5,94445	13,3327	1,16535

Transcrit_id	Gene_id	Gene_name	A FPKM	B FPKM	Log2(FC)
ENST00000009589	ENSG00000008988	RPS20	2361,34	1102,28	-1,09911
ENST00000025301	ENSG00000023516	AKAP11	2,62378	6,9545	1,4063
ENST00000196551	ENSG00000083845	RPS5	1951,1	908,197	-1,10321
ENST00000200453	ENSG00000087074	PPP1R15A	20,7169	9,38523	-1,14235
ENST00000206380	ENSG00000091947	TMEM101	49,7023	21,7173	-1,19447
ENST00000215754	ENSG00000240972	MIF	563,804	195,905	-1,52504
ENST00000216862	ENSG00000019186	CYP24A1	2,62609	7,60931	1,53485
ENST00000217233	ENSG00000101255	TRIB3	38,2	10,2855	-1,89296
ENST00000218230	ENSG00000102109	PCSK1N	27,2429	9,92949	-1,45609
ENST00000221847	ENSG00000105246	EBI3	8,20531	1,8818	-2,12445
ENST00000221975	ENSG00000105372	RPS19	5160,69	1973,7	-1,38666
ENST00000222157	ENSG00000105550	FGF21	8,69545	0,700782	-3,63322
ENST00000222573	ENSG00000105855	ITGB8	2,74129	7,31338	1,41568
ENST00000223095	ENSG00000106366	SERPINE1	58,0148	28,3812	-1,03149
ENST00000225831	ENSG00000108691	CCL2	70,9086	25,8834	-1,45393
ENST00000228841	ENSG00000111245	MYL2	12,9444	2,9515	-2,13281
ENST00000230050	ENSG00000112306	RPS12	2834,93	1457,82	-0,959503
ENST00000230431	ENSG00000112667	DNPH1	56,2093	22,2783	-1,33517
ENST00000232375	ENSG00000114268	PFKFB4	14,2003	5,58867	-1,34535
ENST00000232564	ENSG00000114450	GNB4	3,66609	9,61124	1,39048
ENST00000232975	ENSG00000114854	TNNC1	92,872	24,1297	-1,94443
ENST00000233055	ENSG00000085449	WDFY1	7,58419	18,8853	1,3162
ENST00000233813	ENSG00000115461	IGFBP5	13,0762	30,705	1,23153
ENST00000234961	ENSG00000116329	OPRD1	0,0624199	0,384235	2,62191
ENST00000238497	ENSG00000119541	VPS4B	6,84318	15,2561	1,15665
ENST00000240093	ENSG00000104290	FZD3	1,36475	2,93609	1,10525
ENST00000242208	ENSG00000122641	INHBA	7,59196	18,9797	1,32191
ENST00000242719	ENSG00000123091	RNF11	14,1038	30,3561	1,1059
ENST00000243347	ENSG00000123610	TNFAIP6	0,21817	2,92225	3,74356
ENST00000245458	ENSG00000213741	RPS29	4649,18	1482,79	-1,64866
ENST00000245479	ENSG00000125398	SOX9	16,6628	35,7413	1,10096
ENST00000245903	ENSG00000125726	CD70	50,6619	12,7414	-1,99137
ENST00000246802	ENSG00000105373	GLTSCR2	177,927	46,5098	-1,93568
ENST00000247225	ENSG00000126821	SGPP1	7,14119	19,7861	1,47025
ENST00000248879	ENSG00000128185	DGCR6L	70,4003	29,9047	-1,23521
ENST00000250498	ENSG00000129562	DAD1	205,966	107,768	-0,934482
ENST00000251047	ENSG00000074695	LMAN1	27,7427	73,0824	1,39741
ENST00000251081	ENSG00000134755	DSC2	1,75768	4,03449	1,19871
ENST00000251453	ENSG00000105193	RPS16	2165,2	954,635	-1,18148
ENST00000251646	ENSG00000131738	KRT33B	1,62782	0,394929	-2,04328
ENST00000252245	ENSG00000170454	KRT75	4,81887	20,3312	2,07693
ENST00000252593	ENSG00000130303	BST2	25,3006	71,3998	1,49675
ENST00000252809	ENSG00000130513	GDF15	26,8643	0,822639	-5,02929
ENST00000253047	ENSG00000130748	TMEM160	22,4379	6,99892	-1,68074
ENST00000253063	ENSG00000130766	SESN2	18,0727	6,28334	-1,52421
ENST00000253407	ENSG00000131094	C1QL1	148,825	62,3478	-1,2552
ENST00000253788	ENSG00000131469	RPL27	3107,87	1255,06	-1,30816
ENST00000254508	ENSG00000132182	NUP210	9,85384	20,7028	1,07107
ENST00000255305	ENSG00000132953	XPO4	2,21474	4,65262	1,0709
ENST00000256108	ENSG00000133731	IMPA1	4,63976	15,4139	1,7321
ENST00000256646	ENSG00000134250	NOTCH2	15,1214	28,1767	0,897907
ENST00000257190	ENSG00000134758	RNF138	3,04334	14,2121	2,22339
ENST00000257287	ENSG00000174799	CEP135	5,94445	13,3327	1,16535

ENST00000257836	ENSG00000135378	PRRG4	0,204121	0,710718	1,79985
ENST00000258399	ENSG00000135913	USP37	3,45319	7,60386	1,1388
ENST00000258534	ENSG00000136048	DRAM1	7,23276	2,90684	-1,31509
ENST00000258955	ENSG00000136444	RSAD1	36,629	18,7035	-0,969677
ENST00000259008	ENSG00000136492	BRIP1	7,07586	16,7512	1,24329
ENST00000259874	ENSG00000137331	IER3	119,751	49,528	-1,27372
ENST00000260227	ENSG00000137673	MMP7	2,652	0,736549	-1,84822
ENST00000260526	ENSG00000137962	ARHGAP29	26,9372	54,3463	1,01258
ENST00000260810	ENSG00000163781	TOPBP1	18,5598	38,529	1,05377
ENST00000261187	ENSG00000118596	SLC16A7	0,725478	1,75029	1,27059
ENST00000261304	ENSG00000054983	GALC	11,4905	22,3907	0,96246
ENST00000261402	ENSG00000074590	NUAK1	7,81495	14,0819	0,849537
ENST00000261491	ENSG00000102780	DGKH	0,372709	1,11427	1,57997
ENST00000261590	ENSG00000046604	DSG2	14,494	33,9954	1,22988
ENST00000261637	ENSG00000120800	UTP20	3,71644	7,32248	0,97841
ENST00000262077	ENSG00000124789	NUP153	7,36842	15,6063	1,0827
ENST00000262096	ENSG00000104219	ZDHHC2	7,29502	15,0705	1,04675
ENST00000262120	ENSG00000128791	TWSG1	7,61294	25,8717	1,76485
ENST00000262209	ENSG00000104321	TRPA1	0,95242	3,48547	1,87168
ENST00000262244	ENSG00000120162	MOB3B	0,349335	1,26027	1,85105
ENST00000262464	ENSG00000138829	FBN2	15,33	29,3085	0,934961
ENST00000263339	ENSG00000115008	IL1A	1,58383	0,213162	-2,89339
ENST00000263368	ENSG00000090013	BLVRB	220,699	106,884	-1,04603
ENST00000263735	ENSG00000119888	EPCAM	2,66692	9,36105	1,8115
ENST00000263918	ENSG00000115808	STRN	5,88547	11,9589	1,02285
ENST00000264245	ENSG00000031081	ARHGAP31	1,67267	3,51811	1,07265
ENST00000264930	ENSG00000113504	SLC12A7	26,4053	74,1357	1,48934
ENST00000265000	ENSG00000109586	GALNT7	14,8213	30,7188	1,05145
ENST00000265361	ENSG00000075223	SEMA3C	7,2386	20,4264	1,49665
ENST00000265462	ENSG00000126432	PRDX5	475,982	249,368	-0,93263
ENST00000265471	ENSG00000149541	B3GAT3	36,8256	15,2791	-1,26915
ENST00000265734	ENSG00000105810	CDK6	7,26005	16,9389	1,22229
ENST00000265742	ENSG00000001629	ANKIB1	7,19239	14,0401	0,96501
ENST00000266712	ENSG00000139324	TMTC3	3,64937	8,76276	1,26374
ENST00000267163	ENSG00000139687	RB1	4,83712	13,9492	1,52796
ENST00000269195	ENSG00000141429	GALNT1	11,4161	30,5133	1,41837
ENST00000269593	ENSG00000141753	IGFBP4	54,2251	15,2364	-1,83144
ENST00000270112	ENSG00000142149	HUNK	1,94057	4,20427	1,11537
ENST00000271417	ENSG00000143195	ILDR2	2,9707	5,92239	0,995377
ENST00000272163	ENSG00000143815	LBR	36,3183	75,8861	1,06314
ENST00000272164	ENSG00000143816	WNT9A	0,239865	0,871212	1,8608
ENST00000272641	ENSG00000144227	NXPH2	0,391613	1,60973	2,03932
ENST00000274255	ENSG00000145604	SKP2	12,5455	26,9753	1,10447
ENST00000274457	ENSG00000145780	FEM1C	2,8958	6,12291	1,08025
ENST00000275034	ENSG00000146247	PHIP	3,56955	9,72463	1,4459
ENST00000278193	ENSG00000148943	LIN7C	3,85701	10,1582	1,39708
ENST00000280481	ENSG00000150893	FREM2	0,330287	0,887907	1,42669
ENST00000281043	ENSG00000134323	MYCN	0,492919	2,96486	2,58854
ENST00000281928	ENSG00000123066	MED13L	2,43643	5,60839	1,20282
ENST00000282561	ENSG00000152661	GJA1	17,3322	35,4101	1,0307
ENST00000282570	ENSG00000087338	GMCL1	6,97521	14,8966	1,09468
ENST00000285667	ENSG00000155304	HSPA13	21,7757	43,0935	0,984755
ENST00000286364	ENSG00000155903	RASA2	3,21384	7,3884	1,20096
ENST00000286657	ENSG00000156140	ADAMTS3	3,79748	8,69025	1,19436
ENST00000286800	ENSG00000156273	BACH1	5,71812	12,4365	1,12096
ENST00000287394	ENSG00000156802	ATAD2	40,5016	91,828	1,18095

ENST00000287474	ENSG00000156869	FRRS1	3,43385	8,28964	1,27148
ENST00000288266	ENSG00000157500	APPL1	4,68234	9,93246	1,08492
ENST00000290551	ENSG00000159388	BTG2	12,8624	4,90688	-1,39028
ENST00000293195	ENSG00000161513	FDXR	70,8722	14,8959	-2,25031
ENST00000295087	ENSG00000162980	ARL5A	2,04736	6,91896	1,75679
ENST00000295238	ENSG00000082258	CCNT2	4,4156	11,5722	1,38999
ENST00000295522	ENSG00000163347	CLDN1	12,4357	3,54281	-1,81152
ENST00000295628	ENSG00000163428	LRRC58	9,71273	22,6579	1,22207
ENST00000295888	ENSG00000163625	WDFY3	2,63475	6,91083	1,39119
ENST00000295927	ENSG00000163661	PTX3	29,4043	6,41391	-2,19675
ENST00000295955	ENSG00000163682	RPL9	1245,35	497,47	-1,32387
ENST00000296755	ENSG00000131711	MAP1B	9,88294	19,1058	0,950997
ENST00000296871	ENSG00000164400	CSF2	25,2856	0,981138	-4,68772
ENST00000297350	ENSG00000164761	TNFRSF11B	6,20631	19,3122	1,63771
ENST00000297632	ENSG00000164983	TMEM65	3,08873	10,2243	1,72692
ENST00000299238	ENSG00000166189	HPS6	10,4509	4,88202	-1,09808
ENST00000299441	ENSG00000166341	DCHS1	3,11519	8,07441	1,37404
ENST00000299608	ENSG00000166479	TMX3	10,8925	25,0323	1,20046
ENST00000301246	ENSG00000167644	C19orf33	162,234	48,928	-1,72934
ENST00000301587	ENSG00000167863	ATP5H	578,408	272,553	-1,08555
ENST00000301998	ENSG00000170340	B3GNT2	3,77307	10,7694	1,51313
ENST00000302103	ENSG00000172461	FUT9	0,334449	0,970678	1,53721
ENST00000302418	ENSG00000170759	KIF5B	41,1497	84,8631	1,04426
ENST00000302424	ENSG00000171206	TRIM8	29,0531	13,2594	-1,13167
ENST00000303004	ENSG00000172216	CEBPB	45,1448	16,8247	-1,42398
ENST00000303115	ENSG00000168685	IL7R	34,169	15,5762	-1,13334
ENST00000304312	ENSG00000169020	ATP5I	343,566	159,16	-1,11011
ENST00000304710	ENSG00000170367	CST5	0,67085	7,82568	3,54415
ENST00000307365	ENSG00000168209	DDIT4	53,8279	7,44369	-2,85426
ENST00000307407	ENSG00000169429	CXCL8	20,2814	0,355918	-5,83247
ENST00000307712	ENSG00000100934	SEC23A	9,16947	22,6047	1,30171
ENST00000308330	ENSG00000174106	LEMD3	5,9689	12,1519	1,02564
ENST00000308742	ENSG00000175548	ALG10B	1,85071	5,10392	1,46353
ENST00000309244	ENSG00000105379	ETFB	100,033	42,5758	-1,23237
ENST00000309558	ENSG00000173705	SUSD5	2,97764	6,44814	1,11471
ENST00000310160	ENSG00000174951	FUT1	2,95584	0,6519	-2,18085
ENST00000310454	ENSG00000115159	GPD2	6,36377	15,3349	1,26886
ENST00000310836	ENSG00000174607	UGT8	5,00387	12,7459	1,34892
ENST00000311208	ENSG00000128422	KRT17	267,004	98,6833	-1,43598
ENST00000311322	ENSG00000175445	LPL	2,8815	10,4244	1,85507
ENST00000312493	ENSG00000173517	PEAK1	3,84612	7,77209	1,0149
ENST00000312579	ENSG00000175602	CCDC85B	78,585	20,6369	-1,92903
ENST00000314191	ENSG00000253729	PRKDC	26,2721	51,8507	0,980835
ENST00000315367	ENSG00000175895	PLEKHF2	2,21587	5,35178	1,27214
ENST00000318602	ENSG00000175899	A2M	0,741666	3,13406	2,07919
ENST00000319420	ENSG00000180730	SHISA2	0,418647	3,32768	2,99071
ENST00000320152	ENSG00000099256	PRTFDC1	3,80661	10,8573	1,51209
ENST00000320334	ENSG00000116774	OLFML3	24,9494	51,607	1,04856
ENST00000320876	ENSG00000101596	SMCHD1	7,91021	20,8607	1,399
ENST00000321037	ENSG00000001561	ENPP4	4,18443	9,80081	1,22787
ENST00000321437	ENSG00000166128	RAB8B	6,45673	14,3494	1,15211
ENST00000321731	ENSG00000179008	C14orf39	1,95757	6,0687	1,63233
ENST00000323076	ENSG00000136167	LCP1	4,48712	1,01839	-2,1395
ENST00000323468	ENSG00000178202	KDELC2	9,58966	28,1111	1,55159
ENST00000324453	ENSG00000180354	MTURN	2,64198	8,56993	1,69767
ENST00000324607	ENSG00000172197	MBOAT1	1,68342	4,96107	1,55926

ENST00000325110	ENSG00000112972	HMGCS1	25,4483	62,9189	1,30593
ENST00000325612	ENSG00000180211	RP1-278E11.3	3,49626	0,552966	-2,66055
ENST00000326441	ENSG00000198300	PEG3	3,29624	8,65796	1,3932
ENST00000326505	ENSG00000176597	B3GNT5	2,50535	8,44389	1,7529
ENST00000326793	ENSG00000114331	ACAP2	4,62203	9,90262	1,09928
ENST00000327040	ENSG00000156011	PSD3	1,69087	4,22196	1,32015
ENST00000327490	ENSG00000185262	UBALD2	42,3584	13,3413	-1,66675
ENST00000327741	ENSG00000205426	KRT81	535,585	250,068	-1,0988
ENST00000330236	ENSG00000198315	ZKSCAN8	3,8521	7,79512	1,01693
ENST00000331837	ENSG00000182636	NDN	70,2765	34,8235	-1,01298
ENST00000332180	ENSG00000136051	KIAA1033	3,78998	8,54052	1,17213
ENST00000332246	ENSG00000147408	CSGALNACT1	0,527182	2,70921	2,3615
ENST00000333129	ENSG00000182263	FIGN	2,20397	5,29039	1,26327
ENST00000333916	ENSG00000120159	CAAP1	3,28591	13,367	2,02432
ENST00000334678	ENSG00000187051	RPS19BP1	168,104	75,04	-1,16362
ENST00000334828	ENSG00000171314	PGAM1	352,046	170,554	-1,04554
ENST00000335012	ENSG00000187068	C3orf70	0,473095	1,78067	1,91221
ENST00000336138	ENSG00000130193	THEM6	11,4663	4,43635	-1,36996
ENST00000337190	ENSG00000123933	MXD4	5,90392	2,16808	-1,44525
ENST00000338435	ENSG00000115419	GLS	20,6181	88,4411	2,10081
ENST00000339750	ENSG00000176406	RIMS2	1,98151	4,84729	1,29058
ENST00000340553	ENSG00000170802	FOXN2	3,34309	10,3575	1,63142
ENST00000341864	ENSG00000187325	TAF9B	10,7515	22,8279	1,08626
ENST00000341866	ENSG00000189275	LINC01164	0,541715	0,0549843	-3,30044
ENST00000343677	ENSG00000187837	HIST1H1C	448,013	163,662	-1,45282
ENST00000343986	ENSG00000171858	RPS21	3596,88	1646,81	-1,12707
ENST00000347616	ENSG00000018236	CNTN1	0,892055	5,86885	2,71787
ENST00000347624	ENSG00000172061	LRRC15	1,22892	4,38976	1,83676
ENST00000354275	ENSG00000135549	PKIB	17,2838	59,4702	1,78275
ENST00000354817	ENSG00000197601	FAR1	13,3884	24,8882	0,894482
ENST00000356860	ENSG00000153956	CACNA2D1	2,16496	8,19719	1,92079
ENST00000357232	ENSG00000197410	DCHS2	1,00869	3,21081	1,67046
ENST00000358296	ENSG00000197020	ZNF100	0,181801	0,948593	2,38343
ENST00000358435	ENSG00000196683	TOMM7	214,419	102,254	-1,06827
ENST00000358465	ENSG00000197323	TRIM33	5,15276	10,8704	1,07699
ENST00000358697	ENSG00000198205	ZXDA	0,909826	3,5877	1,97939
ENST00000358812	ENSG00000172469	MANEA	2,07696	5,26616	1,34228
ENST00000358912	ENSG00000197329	PELI1	1,55378	4,65024	1,58152
ENST00000359314	ENSG00000198087	CD2AP	7,91143	17,739	1,16492
ENST00000359358	ENSG00000197705	KLHL14	0,126516	0,689226	2,44566
ENST00000359478	ENSG00000197614	MFAP5	3,56936	0,472404	-2,91757
ENST00000360612	ENSG00000198355	PIM3	44,9619	21,7569	-1,04723
ENST00000360647	ENSG00000163848	ZNF148	2,28335	6,33704	1,47265
ENST00000360982	ENSG00000061676	NCKAP1	31,1312	64,3624	1,04786
ENST00000361138	ENSG00000198887	SMC5	7,08628	17,3609	1,29274
ENST00000361157	ENSG00000126709	IFI6	41,6718	12,0837	-1,78601
ENST00000361228	ENSG00000198774	RASSF9	0,716464	1,88562	1,39607
ENST00000361371	ENSG00000198862	LTN1	7,9759	17,3227	1,11895
ENST00000361439	ENSG00000169189	NSMCE1	59,194	27,9269	-1,08379
ENST00000361575	ENSG00000198918	RPL39	1666,44	765,213	-1,12284
ENST00000361758	ENSG00000112531	QKI	12,5951	30,6392	1,28252
ENST00000361956	ENSG00000198732	SMOC1	1,97058	28,8473	3,87174
ENST00000363046	ENSG00000277027	RMRP	10,0887	69,9092	2,79274
ENST00000364215	ENSG00000201085	RNU6-173P	37,5802	0	-inf
ENST00000366481	ENSG00000162722	TRIM58	3,88047	9,09363	1,22863
ENST00000366553	ENSG00000203668	CHML	10,1593	23,2806	1,19633

ENST00000367104	ENSG00000122335	SERAC1	1,49656	3,96221	1,40465
ENST00000367194	ENSG00000170498	KISS1	18,6511	0,336384	-5,79301
ENST00000367311	ENSG00000174307	PHLDA3	25,5279	7,53721	-1,75997
ENST00000367405	ENSG00000177888	ZBTB41	1,29394	4,04305	1,64367
ENST00000367429	ENSG00000000971	CFH	1,00052	3,81491	1,9309
ENST00000367545	ENSG00000152818	UTRN	4,77568	12,8935	1,43287
ENST00000367558	ENSG00000143333	RGS16	3,39962	8,19448	1,26928
ENST00000367662	ENSG00000116183	PAPPA2	0,90833	0,268612	-1,75769
ENST00000368012	ENSG00000143217	NECTIN4	1,19324	0,0865407	-3,78536
ENST00000368468	ENSG00000111885	MAN1A1	11,5492	30,0023	1,37728
ENST00000368567	ENSG00000177954	RPS27	4592,41	2417,38	-0,925805
ENST00000368719	ENSG00000197956	S100A6	637,009	282,792	-1,17157
ENST00000369278	ENSG00000014123	UFL1	4,3181	10,4268	1,27183
ENST00000369409	ENSG00000092621	PHGDH	252,199	142,861	-0,819954
ENST00000369449	ENSG00000155962	CLIC2	0,42683	5,60718	3,71554
ENST00000369454	ENSG00000155961	RAB39B	3,94322	1,48812	-1,40589
ENST00000369577	ENSG00000188994	ZNF292	2,23885	5,31598	1,24758
ENST00000369697	ENSG00000171385	KCND3	0,178594	0,744453	2,0595
ENST00000369713	ENSG00000148834	GSTO1	89,6033	30,0585	-1,57578
ENST00000369937	ENSG00000107874	CUEDC2	71,7207	25,6023	-1,48612
ENST00000370223	ENSG00000107816	LZTS2	43,5844	19,8408	-1,13534
ENST00000370278	ENSG00000221867	MAGEA3	7,42173	0,410745	-4,17544
ENST00000370521	ENSG00000065243	PKN2	9,52531	19,3388	1,02166
ENST00000370664	ENSG00000165886	UBTD1	22,5104	9,80112	-1,19957
ENST00000370742	ENSG00000162618	ADGRL4	1,4486	7,04482	2,2819
ENST00000370763	ENSG00000162616	DNAJB4	4,53253	10,8927	1,26498
ENST00000370828	ENSG00000076716	GPC4	5,60877	13,2644	1,2418
ENST00000370978	ENSG00000056277	ZNF280C	1,03022	2,71222	1,39652
ENST00000371001	ENSG00000196074	SYCP2	0,495438	4,82799	3,28465
ENST00000371443	ENSG00000182154	MRPL41	65,7643	28,5866	-1,20197
ENST00000372106	ENSG00000101888	NXT2	2,03533	5,91233	1,53847
ENST00000372680	ENSG00000204065	TCEAL5	0,351713	0	-inf
ENST00000372733	ENSG00000124145	SDC4	125,745	51,6997	-1,28227
ENST00000373087	ENSG00000163874	ZC3H12A	10,9353	4,82074	-1,18166
ENST00000373113	ENSG00000095970	TREM2	0,516427	0	-inf
ENST00000373644	ENSG00000138336	TET1	0,368339	1,28901	1,80715
ENST00000373659	ENSG00000186130	ZBTB6	1,30543	3,35404	1,36137
ENST00000373664	ENSG00000080839	RBL1	3,93043	8,27518	1,0741
ENST00000373789	ENSG00000182010	RTKN2	3,52237	8,90123	1,33746
ENST00000373935	ENSG00000151151	IPMK	1,81208	4,80013	1,40543
ENST00000374379	ENSG00000169504	CLIC4	61,2427	113,496	0,890035
ENST00000374557	ENSG00000095203	EPB41L4B	2,00908	5,46135	1,44272
ENST00000374580	ENSG00000204217	BMPR2	4,58019	9,58537	1,06543
ENST00000374888	ENSG00000198455	ZXDB	3,01131	8,69946	1,53054
ENST00000375256	ENSG00000165244	ZNF367	7,78414	19,8912	1,35352
ENST00000376150	ENSG00000068985	PAGE1	0,416872	82,3361	7,62578
ENST00000376447	ENSG00000165105	RASEF	1,60505	4,14995	1,37048
ENST00000376958	ENSG00000152749	GPR180	2,3294	5,1531	1,14548
ENST00000377200	ENSG00000188647	PTAR1	3,41088	8,29931	1,28285
ENST00000377275	ENSG00000165997	ARL5B	4,36316	9,65998	1,14665
ENST00000378164	ENSG00000064313	TAF2	8,57406	18,011	1,07083
ENST00000378667	ENSG00000164405	UQCRQ	1070,61	529,247	-1,01642
ENST00000378858	ENSG00000125872	LRRN4	0,205372	1,04778	2,35103
ENST00000378930	ENSG00000157625	TAB3	2,16013	6,82169	1,65901
ENST00000379389	ENSG00000187608	ISG15	50,0932	12,9983	-1,94629
ENST00000380138	ENSG00000164176	EDIL3	31,9992	72,902	1,18792

ENST00000380152	ENSG00000139618	BRCA2	2,03402	5,6627	1,47715
ENST00000380672	ENSG00000173068	BNC2	1,12016	2,62929	1,23097
ENST00000380916	ENSG00000175893	ZDHHC21	1,6609	3,60291	1,11719
ENST00000381280	ENSG00000198732	SMOC1	1,71428	15,5729	3,18336
ENST00000381401	ENSG00000169100	SLC25A6	100,257	47,6355	-1,0736
ENST00000381913	ENSG00000167723	TRPV3	0,0960037	0,921817	3,26332
ENST00000389463	ENSG00000112769	LAMA4	0,318848	2,27411	2,83436
ENST00000391753	ENSG00000170889	RPS9	1336,77	591,214	-1,177
ENST00000391812	ENSG00000167747	C19orf48	77,3175	25,873	-1,57935
ENST00000391857	ENSG00000142541	RPL13A	1492,38	473,693	-1,65559
ENST00000392647	ENSG00000213160	KLHL23	5,84164	12,421	1,08834
ENST00000392730	ENSG00000215030	RPL13P12	110,367	44,0839	-1,32399
ENST00000393063	ENSG00000100697	DICER1	2,58599	7,84967	1,60192
ENST00000393249	ENSG00000091039	OSBPL8	1,97982	7,25652	1,87391
ENST00000393676	ENSG00000110195	FOLR1	14,7837	2,93343	-2,33335
ENST00000393752	ENSG00000184845	DRD1	1,83105	4,41373	1,26932
ENST00000394001	ENSG00000131737	KRT34	18,3307	1,99212	-3,20189
ENST00000394144	ENSG00000036054	TBC1D23	12,3375	39,9456	1,69499
ENST00000394329	ENSG00000196159	FAT4	0,659437	1,63332	1,3085
ENST00000394332	ENSG00000125691	RPL23	1695,31	695,172	-1,28611
ENST00000394684	ENSG00000164023	SGMS2	2,78192	8,7575	1,65444
ENST00000395603	ENSG00000085224	ATRX	5,80873	13,6741	1,23515
ENST00000395738	ENSG00000128165	ADM2	6,9891	3,03001	-1,20578
ENST00000396680	ENSG00000128272	ATF4	282,52	134,833	-1,06718
ENST00000397128	ENSG00000132356	PRKAA1	7,34935	19,6704	1,42034
ENST00000397786	ENSG00000108510	MED13	8,75495	18,098	1,04766
ENST00000398316	ENSG00000102531	FNDC3A	3,08767	8,13482	1,39759
ENST00000399302	ENSG00000060749	QSER1	7,70315	17,3804	1,17394
ENST00000400286	ENSG00000184564	SLITRK6	0,173936	1,77314	3,34968
ENST00000401609	ENSG00000100316	RPL3	1669,73	735,251	-1,18331
ENST00000404735	ENSG00000143947	RPS27A	2072,73	1019,05	-1,0243
ENST00000407540	ENSG00000132031	MATN3	4,73465	11,5571	1,28745
ENST00000407925	ENSG00000049323	LTBP1	1,92326	9,97233	2,37438
ENST00000408925	ENSG00000221821	C6orf226	29,0774	10,1076	-1,52446
ENST00000408965	ENSG00000221869	CEBPD	8,73927	2,38808	-1,87166
ENST00000409914	ENSG00000196604	POTEF	0,571136	2,8803	2,33431
ENST00000409947	ENSG00000169967	MAP3K2	3,59925	8,47597	1,23568
ENST00000411416	ENSG00000145287	PLAC8	53,5808	13,945	-1,94197
ENST00000411975	ENSG00000179869	ABCA13	1,34689	3,95075	1,55249
ENST00000412590	ENSG00000171483	SSX6	0	0,246196	inf
ENST00000416195	ENSG00000172845	SP3	13,8829	31,4278	1,17874
ENST00000417088	ENSG00000233927	RPS28	1473,69	608,668	-1,27571
ENST00000421491	ENSG00000254535	PABPC4L	1,29344	3,44647	1,41391
ENST00000431979	ENSG00000224559	LINC01087	1,57609	4,51511	1,51841
ENST00000435466	ENSG00000226965	AC003088.1	0,293355	0	-inf
ENST00000435607	ENSG00000007314	SCN4A	0,0341752	0,828441	4,59938
ENST00000439590	ENSG00000172809	RPL38	1407,76	509,552	-1,4661
ENST00000440460	ENSG00000150764	DIXDC1	1,09038	3,58749	1,71815
ENST00000441095	ENSG00000233922	AL133493.2	0,109269	1,02162	3,2249
ENST00000441328	ENSG00000240583	AQP1	5,37768	12,4792	1,21447
ENST00000442697	ENSG00000164031	DNAJB14	3,44701	9,30156	1,43213
ENST00000443617	ENSG00000103657	HERC1	2,58081	5,1644	1,00078
ENST00000443836	ENSG00000228208	C1orf143	3,79749	0,402547	-3,23782
ENST00000444165	ENSG00000223495	RP11-763B22.7	0	1,01513	inf
ENST00000444980	ENSG00000231715	COX6CP2	2,03984	0	-inf
ENST00000446231	ENSG00000157106	SMG1	6,64287	14,1645	1,0924

ENST00000448775	ENSG00000166250	CLMP	1,57542	4,70485	1,57841
ENST00000448864	ENSG00000182899	RPL35A	1391,35	641,545	-1,11687
ENST00000450251	ENSG00000145390	USP53	1,94852	6,00113	1,62285
ENST00000450471	ENSG00000138395	CDK15	1,16492	0,313224	-1,89496
ENST00000451531	ENSG00000196834	POTEI	2,19113	7,78274	1,8286
ENST00000452400	ENSG00000140396	NCOA2	2,28384	6,00045	1,39361
ENST00000455127	ENSG00000225663	MCRIP1	56,6965	12,3917	-2,19389
ENST00000456586	ENSG00000197756	RPL37A	4458,8	2240,93	-0,992556
ENST00000458537	ENSG00000136603	SKIL	4,06108	12,5171	1,62396
ENST00000463836	ENSG00000163584	RPL22L1	576,722	186,968	-1,62508
ENST00000466501	ENSG00000237550	RPL9P9	415,293	169,609	-1,29192
ENST00000473756	ENSG00000240476	LINC00973	78,0645	28,8751	-1,43484
ENST00000477922	ENSG00000257743	MGAM2	0,0719669	0,508129	2,81979
ENST00000479017	ENSG00000162244	RPL29	925,218	351,927	-1,39452
ENST00000482325	ENSG00000135968	GCC2	5,67116	12,6374	1,15598
ENST00000484859	ENSG00000241860	RP11-34P13.13	1,11837	3,48195	1,6385
ENST00000486888	ENSG00000243635	RP11-281P11.1	0,469297	0	-inf
ENST00000487273	ENSG00000168615	ADAM9	58,6594	123,071	1,06906
ENST00000491934	ENSG00000269926	RP11-442H21.2	133,486	22,0439	-2,59824
ENST00000495707	ENSG00000196584	XRCC2	2,31597	5,43076	1,22954
ENST00000503486	ENSG00000250420	AACSP1	3,01881	0,624299	-2,27367
ENST00000504087	ENSG00000151458	ANKRD50	2,68855	6,95594	1,37142
ENST00000504562	ENSG00000145592	RPL37	2494,05	1101,82	-1,17861
ENST00000504573	ENSG00000248367	CTB-129O4.1	12,3045	3,59639	-1,77456
ENST00000511306	ENSG00000183206	POTEC	0,485168	1,7409	1,84327
ENST00000517927	ENSG00000253522	MIR3142HG	3,56708	0,2653	-3,74905
ENST00000518973	ENSG00000253288	RP11-238K6.1	0,678652	2,6472	1,96372
ENST00000519656	ENSG00000132294	EFR3A	14,6628	32,3139	1,13999
ENST00000520572	ENSG00000253741	CTD-2292P10.4	1,8016	5,48819	1,60705
ENST00000521466	ENSG00000164587	RPS14	1632,22	533,341	-1,6137
ENST00000521891	ENSG00000091656	ZFHX4	1,26974	2,86946	1,17625
ENST00000522551	ENSG00000233913	RPL10P9	132,637	53,6565	-1,30565
ENST00000524013	ENSG00000136997	MYC	49,1327	17,8115	-1,46387
ENST00000525217	ENSG00000255026	RP11-326C3.2	0,959997	11,5303	3,58626
ENST00000530760	ENSG00000255099	RP11-25D10.2	0,14161	0,537807	1,92517
ENST00000539065	ENSG00000136536	MARCH7	9,06406	22,7204	1,32576
ENST00000540845	ENSG00000172059	KLF11	2,96432	8,67673	1,54945
ENST00000543571	ENSG00000131023	LATS1	2,53151	6,7128	1,40692
ENST00000544848	ENSG00000246705	H2AFJ	106,597	27,197	-1,97065
ENST00000545203	ENSG00000185483	ROR1	1,87731	11,158	2,57135
ENST00000545294	ENSG00000152332	UHMK1	4,6343	12,8852	1,47529
ENST00000546136	ENSG00000134532	SOX5	0,30648	1,35296	2,14225
ENST00000553612	ENSG00000165417	GTF2A1	3,69688	7,51872	1,02418
ENST00000557158	ENSG00000258585	ZNF519P3	0,211241	1,42588	2,75489
ENST00000559447	ENSG00000181827	RFX7	1,74681	5,80459	1,73247
ENST00000562945	ENSG00000260073	RP11-488I20.9	0,680409	2,56387	1,91385
ENST00000563192	ENSG00000260260	SNHG19	451,713	124,393	-1,8605
ENST00000564152	ENSG00000260708	CTA-29F11.1	38,4685	16,5017	-1,22106
ENST00000565120	ENSG00000261514	RP11-527L4.2	1,90638	8,31899	2,12557
ENST00000565283	ENSG00000260025	RP11-490M8.1	40,7221	12,1567	-1,74406
ENST00000566986	ENSG00000260563	RP13-516M14.1	12,3684	3,59711	-1,78175
ENST00000570269	ENSG00000259976	MIR568	1,81276	4,99699	1,46287
ENST00000579374	ENSG00000108298	RPL19	4866,34	1906,62	-1,35182
ENST00000580094	ENSG00000204177	BMS1P1	2,08934	4,96717	1,24938
ENST00000582401	ENSG00000265972	TXNIP	7,14656	1,4091	-2,34248
ENST00000587427	ENSG00000167925	GHDC	13,6009	5,07182	-1,42313

ENST00000590478	ENSG00000266947	RP11-799D4.4	1,0203	0,228372	-2,15953
ENST00000596358	ENSG00000142546	NOSIP	85,1862	32,8596	-1,3743
ENST00000599898	ENSG00000105640	RPL18A	1747,96	706,555	-1,3068
ENST00000601306	ENSG00000142534	RPS11	3339,98	1572,18	-1,08707
ENST00000602461	ENSG00000270069	MIR222HG	29,7404	11,0045	-1,43434
ENST00000604437	ENSG00000270392	PFN1P2	1,51635	5,89617	1,95918
ENST00000606645	ENSG00000272449	RP3-395M20.12	1,80378	0,424014	-2,08884
ENST00000609661	ENSG00000229373	LINC00452	8,74144	21,2103	1,27882
ENST00000610979	ENSG00000273619	RP5-908M14.9	6,72562	2,04488	-1,71765
ENST00000611049	ENSG00000152208	GRID2	2,81915	7,85845	1,47899
ENST00000611847	ENSG00000204186	ZDBF2	1,20138	3,21781	1,42138
ENST00000612365	ENSG00000276851	RP5-875H18.9	7,48751	2,30079	-1,70236
ENST00000613130	ENSG00000278626	RP11-467L19.16	0,4636	0	-inf
ENST00000613549	ENSG00000272636	DOC2B	0,0237605	0,513015	4,43236
ENST00000613650	ENSG00000132563	REEP2	11,403	4,01841	-1,50472
ENST00000615852	ENSG00000273576	RP11-390P24.1	12,0803	4,15865	-1,53847
ENST00000616921	ENSG00000176624	MEX3C	11,1233	22,296	1,00319
ENST00000617961	ENSG00000128965	CHAC1	20,3677	4,50351	-2,17716
ENST00000617998	ENSG00000181222	POLR2A	7,98522	16,4795	1,04527
ENST00000619432	ENSG00000274213	RP11-670E13.6	23,6785	7,06291	-1,74524
ENST00000620384	ENSG00000189057	FAM111B	10,7451	39,3714	1,87347
ENST00000620890	ENSG00000233276	GPX1	332,737	140,043	-1,24851
ENST00000621052	ENSG00000274827	LINC01297	15,1668	28,9804	0,934163
ENST00000623678	ENSG00000279998	CH507-210P18.1	0,12389	2,13796	4,10911
ENST00000623708	ENSG00000279795	RP11-586K12.1	0,812081	3,00611	1,8882
ENST00000623959	ENSG00000279184	RP3-323A16.1	0,623619	1,31728	1,07883
ENST00000624675	ENSG00000279561	RP11-4L24.4	0,861085	3,1796	1,88462
ENST00000635788	ENSG00000283431	RP11-1217F2.22	0,420958	2,5715	2,61086
ENST00000636484	ENSG00000283293	RN7SK	9,77872	89,6762	3,19701
ENST00000637938	ENSG00000153404	PLEKHG4B	0,67593	2,04116	1,59445

Table S2. Differential expression analysis of transcripts between U-E6*I vs. U-CTL cell lines.

Transcrit_id	Gene_id	Gene_name	A FPKM	D FPKM	Log2(FC)
ENST00000225831	ENSG00000108691	CCL2	70,9086	229,65	1,6954
ENST00000232375	ENSG00000114268	PFKFB4	14,2003	6,07868	-1,22409
ENST00000232975	ENSG00000114854	TNNC1	92,872	193,504	1,05905
ENST00000239223	ENSG00000120129	DUSP1	19,6896	42,8732	1,12264
ENST00000243347	ENSG00000123610	TNFAIP6	0,21817	2,86396	3,71449
ENST00000249071	ENSG00000128340	RAC2	4,6224	18,1016	1,9694
ENST00000252245	ENSG00000170454	KRT75	4,81887	40,316	3,06458
ENST00000252456	ENSG00000130176	CNN1	2,61032	10,2604	1,97478
ENST00000255409	ENSG00000133048	CHI3L1	2,68545	9,94964	1,88948
ENST00000260302	ENSG00000137745	MMP13	0,777232	0,107929	-2,84826
ENST00000261172	ENSG00000083782	EPYC	0,334646	0	-inf
ENST00000263233	ENSG00000102003	SYP	2,73063	6,75861	1,30749
ENST00000269593	ENSG00000141753	IGFBP4	54,2251	22,4485	-1,27234
ENST00000286049	ENSG00000155622	XAGE2	0,734121	4,82605	2,71675
ENST00000290271	ENSG00000159167	STC1	5,51846	0,919621	-2,58515
ENST00000294984	ENSG00000162892	IL24	1,12028	10,2676	3,19617
ENST00000299855	ENSG00000149968	MMP3	7,12382	22,8589	1,68204
ENST00000300976	ENSG00000167487	KLHL26	0,315902	1,27282	2,01048
ENST00000303045	ENSG00000169436	COL22A1	3,51536	9,15944	1,38159
ENST00000303329	ENSG00000172379	ARNT2	1,91336	4,49728	1,23294
ENST00000307128	ENSG00000172399	MYOZ2	0,0708259	0,544479	2,94253
ENST00000318602	ENSG00000175899	A2M	0,741666	3,85185	2,37671
ENST00000331163	ENSG00000100311	PDGFB	3,71655	10,8482	1,54542
ENST00000335119	ENSG00000185972	CCIN	0,517542	1,80604	1,80309
ENST00000366714	ENSG00000198835	GJC2	0,781879	6,49633	3,05461
ENST00000367976	ENSG00000118523	CTGF	118,54	236,484	0,996368
ENST00000369704	ENSG00000120051	CFAP58	2,75898	7,94273	1,5255
ENST00000370742	ENSG00000162618	ADGRL4	1,4486	4,88083	1,75247
ENST00000372476	ENSG00000066056	TIE1	0,163939	1,30108	2,98849
ENST00000373008	ENSG00000122884	P4HA1	38,7768	14,4828	-1,42085
ENST00000374045	ENSG00000181634	TNFSF15	0,197881	1,15246	2,54201
ENST00000374955	ENSG00000158639	PAGE5	0,472348	19,5859	5,37382
ENST00000375978	ENSG00000179772	FOXS1	1,71713	8,74701	2,34879
ENST00000376150	ENSG00000068985	PAGE1	0,416872	29,992	6,16883
ENST00000376468	ENSG00000120937	NPPB	0,936778	4,57856	2,28912
ENST00000381280	ENSG00000198732	SMOC1	1,71428	10,4807	2,61206
ENST00000382349	ENSG00000205922	ONECUT3	0,389932	0,105208	-1,88999
ENST00000423218	ENSG00000137809	ITGA11	0,0781795	0,406507	2,37842
ENST00000425951	ENSG00000244155	CYP4F34P	0,771632	4,08092	2,40291
ENST00000433460	ENSG00000230109	RP11-275N1.1	0,445077	0	-inf
ENST00000442437	ENSG00000224659	GAGE12J	0	5,87656	inf
ENST00000448656	ENSG00000153237	CCDC148	2,58906	8,77587	1,76111
ENST00000451052	ENSG00000175302	ANKRD30BP1	0,101832	0,731134	2,84395
ENST00000452841	ENSG00000235500	SNX19P2	0,250609	0,896833	1,8394
ENST00000495688	ENSG00000143851	PTPN7	0,706745	3,37844	2,2571
ENST00000498616	ENSG00000244128	LINC01322	0,682091	5,92429	3,11861
ENST00000527950	ENSG00000109846	CRYAB	14,4714	53,5444	1,88753
ENST00000547963	ENSG00000249550	LINC01234	1,03097	3,86572	1,90674
ENST00000558480	ENSG00000058335	RASGRF1	4,27756	12,1228	1,50286
ENST00000595748	ENSG00000231412	CTC-490G23.2	0	0,636608	inf
ENST00000609911	ENSG00000273415	RP11-702B10.2	0,154108	1,37656	3,15905
ENST00000615826	ENSG00000275874	PICSAR	10,4787	26,9794	1,3644
ENST00000623678	ENSG00000279998	CH507-210P18.1	0,12389	1,44036	3,5393

Annexe 2 : Liste des communications

Affichées :

- Meznad K, Morel A, Missey A, Mougin C, Prétet JL, Aubin F, Baguet A. *Implication du facteur cellulaire eIF4III dans l'expression des gènes viraux d'HPV16.* 10th SifrARN meeting, Toulouse, France, 2016.

Orales :

- Meznad K, Morel A, Mougin C, Prétet JL, Aubin F, Baguet A. *Régulation de la stabilité des transcrits d'HPV16.* 10èmes Journées Scientifiques de l'AFIPP, Sainte-Foy-Lès-Lyon, France, 2017.

- Meznad K, Morel A, Mougin C, Prétet JL, Aubin F, Baguet A *Impact of Exon Junction Complex factors on HPV16 E6 and E7 oncoproteins expression*. 23ème Forum des jeunes chercheurs, Dijon, France, 2017.

- Meznad K, Morel A, Mougin C, Prétet JL, Aubin F, Baguet A. *Régulation post-transcriptionnelle des transcrits d'HPV16 par le complexe de jonction des exons*. 10ème Forum du Cancéropôle du Grand-Est, Nancy, France, 2016.

- Meznad K, Morel A, Missey A, Mourareau C, Jung AC, Dalstein V, Clavel C, Prétet JL, Mougin C, Aubin F, Baguet A. *Implication du complexe de jonction des exons dans la régulation d'expression des gènes d'HPV16*. 9èmes Journées Scientifiques de l'AFIPP, Loches, France, 2016.

- Meznad K, Morel A, Missey A, Mougin C, Aubin F, Baguet A. *Le Nonsense-mediated mRNA decay : surveillance des ARNm viraux d'HPV16*. 22ème Forum des jeunes chercheurs, Besançon, France, 2016.

- Meznad K, Morel A, Mougin C, Prétet JL, Aubin F, Baguet A. *Régulation des transcrits viraux d'HPV16 par le NMD*. Journée Scientifique de la SFR FED 4234, Besançon, France, 2016.

- Meznad K, Morel A, Missey A, Mougin C, Aubin F, Prétet JL, Baguet A. *Régulation post-transcriptionnelle de l'expression des transcrits d'HPV16*. Société de Biologie de Besançon, France, 2015.

- Meznad K, Missey A, Morel A, Mougin C, Aubin F, Baguet A. *Impact de l'ARN hélicase eIF4AIII sur l'expression de l'oncoprotéine E6 d'HPV16*. 8èmes Journées Scientifiques de l'AFIPP, Rouilly-Sacey, France, 2015.

Abstract

Human papillomaviruses (HPV) are double strand DNA viruses that infect skin and mucosa. HPV infections, although mostly asymptomatic, can sometimes give rise to cancer. According to their carcinogenic potential, we distinguish low-risk HPVs (lr-HPV) causing benign lesions, and high-risk HPV (hr-HPV) responsible for the appearance of numerous anogenital and some head and neck squamous-cell cancers. Among the hr-HPV, HPV16 is the most prevalent. Hr-HPV-induced carcinogenesis is linked with the expression of the viral oncoproteins, E6 and E7, which deregulate many cellular processes. Viral gene expression, performed by the host cell machine, is finely regulated particularly at the post-transcriptional level. Besides, alternative splicing generates about twenty viral transcripts, leading to the expression of viral proteins. The splicing within the E6 open reading frame that generates an E6*I mRNA only in hr-HPV, but not in the lr-HPV, suggests its involvement in hr-HPV-induced carcinogenesis. However, the biological role of E6*I protein produced by HPV-HR is still controversial.

In order to better understand the mechanisms of hr-HPV-induced carcinogenesis, we have interested in: (i) the study of the biological functions of the E6*I isoform, and (ii) the mechanisms involved in the regulation of E6 and E7 protein expression.

To study the biological role of HPV16 E6*I, we used RNA sequencing to identify targets deregulated by its ectopic expression. Expression of HPV16 E6 and E6*I isoforms in negative HPV cells deregulate several transcripts involved in biological processes related to viral gene expression, viral carcinogenesis, signal transduction and translation. The expression of E6*I alone, deregulates transcripts involved in the organization of the extracellular matrix, signaling pathways and cell adhesion. Interestingly, it was shown that the genes deregulated by E6*I expression are commonly affected by the intracellular level of ROS (reactive oxygen species). These results support the role of E6*I in increasing ROS production. The ROS-associated oxidative stress could favor viral genome integration with that of the host cell, a characteristic of hr-induced carcinogenesis. In sum, E6*I may have an oncogenic role independent of E6, and intervene in the carcinogenesis associated with hr-HPV.

We also studied the role of the exon junction complex (EJC) in the posttranscriptional regulation of E6 and E7 expression. EJC is a multiprotein complex deposited on mRNAs *via* splicing, thus influencing their fate. We have shown that a factor of EJC, eukaryotic initiation factor 4A3 (eIF4A3), binds to viral mRNAs. Moreover, we have observed that the components of the EJC affected, in different ways, the expression of E6 and E7. Finally, we also studied the effect of nonsense-mediated mRNA decay (NMD), a mechanism linked to the EJC, on the expression of E6 and E7. Our results suggest that not only NMD inhibits the expression of E6 and E7, but we have also observed that HPV16 E6 protein reduces NMD activity. This inhibition would allow HPV16 to have control over its transcripts but also to affect NMD cellular targets. Given the involvement of NMD-regulated genes in the maintenance of cellular homeostasis and adaptation, it would be interesting to understand the role of this new E6 activity in carcinogenesis associated with HPV-HR.

Résumé

Les papillomavirus humains (HPV) sont des virus à ADN double brin à tropisme cutanéoqui infectent la peau et les muqueuses. Les infections par les HPV, bien que majoritairement asymptomatiques, peuvent parfois générer des cancers. Selon leur pouvoir carcinogène, on distingue les HPV à bas risque oncogène (HPV-BR) provoquant des lésions bénignes, et les HPV à haut risque (HPV-HR) responsables de l'apparition de nombreux cancers ano-génitaux et de certains cancers des voies aéro-digestives supérieures. Parmi les HPV-HR, HPV16 est le plus prévalent. La carcinogenèse induite par les HPV-HR est liée à l'expression des protéines virales E6 et E7, qui dérégulent de nombreux processus cellulaires. L'expression des gènes viraux, réalisée par la machinerie de la cellule hôte, est finement régulée particulièrement au niveau posttranscriptionnel. En outre, l'épissage alternatif génère une vingtaine de transcrits viraux, permettant l'expression des protéines virales. L'épissage au sein de la région codant E6 permettant de former l'isoforme E6*I est présent uniquement chez les HPV-HR, mais pas chez les HPV-BR, ce qui suggère son implication dans la carcinogenèse induite par les HPV-HR. Toutefois, le rôle biologique de la protéine E6*I produite par les HPV-HR est encore controversé.

Afin de mieux comprendre les mécanismes de la carcinogenèse induite par les HPV-HR, nous nous sommes intéressés à : (i) l'étude des fonctions biologiques de l'isoforme E6*I, et (ii) aux mécanismes impliqués dans la régulation de l'expression des oncoprotéines E6 et E7.

Pour étudier le rôle biologique d'E6*I d'HPV16, nous avons utilisé le séquençage de l'ARN afin d'identifier des cibles dérégulées par son expression ectopique. L'expression des isoformes E6 et E6*I d'HPV16 dans des cellules HPV négatives dérégule des transcrits impliqués dans des processus biologiques relatifs à l'expression des gènes viraux, la carcinogenèse virale, la transduction du signal et la traduction. L'expression d'E6*I seule, dérégule des transcrits impliqués dans l'organisation de la matrice extracellulaire, des voies de signalisation et d'adhérence cellulaire. De façon intéressante, il a été montré que ces gènes dérégulés par l'expression d'E6*I sont communément affecté par le niveau intracellulaire de ROS (espèces réactives de l'oxygène). Cela corrobore le rôle d'E6*I dans l'augmentation de la production de ROS. Le stress oxydatif associé aux ROS pourrait favoriser l'intégration du génome viral à celui de la cellule hôte, caractéristique de carcinogenèse associée aux HPV-HR. En somme, E6*I pourrait avoir un rôle oncogénique indépendant de celui d'E6, et interviendrait dans la carcinogenèse associée aux HPV-HR.

Nous avons aussi étudié le rôle du complexe de jonction des exons (EJC), dans la régulation posttranscriptionnelle de l'expression d'E6 et E7. L'EJC est un complexe multiprotéique déposé sur les ARNm *via* l'épissage influençant ainsi leur devenir. Nous avons montré qu'un facteur de l'EJC, eukaryotic initiation factor 4A3 (eIF4A3), se liait aux ARNm viraux. Par ailleurs, nous avons observé que les composants de l'EJC affectent, certes de différentes façons, l'expression d'E6 et E7. Enfin, nous avons aussi étudié l'effet du nonsense-mediated mRNA decay (NMD), un mécanisme lié à l'EJC, sur l'expression d'E6 et E7. Non seulement nos résultats suggèrent que le NMD inhibe l'expression d'E6 et E7, mais nous avons aussi observé que la protéine E6 d'HPV16 réduit l'activité du NMD. Cette inhibition permettrait à HPV16 d'avoir un contrôle sur ses transcrits mais d'affecter aussi des cibles cellulaires du NMD. Etant donné l'implication des gènes régulés par le NMD dans le maintien de l'homéostasie et l'adaptation cellulaires, il serait intéressant d'appréhender le rôle de cette nouvelle activité d'E6 dans la carcinogenèse associée aux HPV-HR.