

Université de Montréal

# **Analyse fonctionnelle des polymorphismes du virus du papillome humain de type 33**

par Jennifer Élisabeth Alvarez Orellana

Département de biochimie et médecine moléculaire  
Faculté de médecine

Thèse présentée à la Faculté de médecine  
en vue de l'obtention du grade de doctorat en biochimie  
option génétique moléculaire

Avril, 2016

© Jennifer Élisabeth Alvarez Orellana, 2016



Université de Montréal  
Faculté des études supérieures et postdoctorales

Cette thèse intitulée :  
Analyse fonctionnelle des polymorphismes du virus du papillome humain de type 33

Présentée par :  
Jennifer Élisabeth Alvarez Orellana

a été évaluée par un jury composé des personnes suivantes :

Dr Gerardo Ferbeyre, président-rapporteur  
Dr Jacques Archambault, directeur de recherche  
Dr François Coutlée, co-directeur de recherche  
Dr André Tremblay, membre du jury  
Dre Anne Gatignol, examinatrice externe  
Dr Guy Lemay, représentant du doyen de la FES



## Résumé

L'infection au virus du papillome humain oncogénique (VPH-HR) est associée au développement du cancer du col de l'utérus. L'expression des oncogènes viraux E6 et E7, qui favorisent la progression vers le cancer, est régulée par la région de contrôle longue ou LCR qui contient des sites de liaison pour plusieurs facteurs de transcription cellulaires et pour le répresseur viral E2. L'expression des oncogènes E6 et E7 est souvent dérégulée dans les cancers du col de l'utérus, une résultante de l'intégration du génome viral qui détruit fréquemment le cadre de lecture du gène E2. Plusieurs isolats ou variantes du VPH de type 33 (VPH33) avec une séquence nucléotidique du LCR différente du prototype ont été identifiés dans des spécimens de femmes infectées par le VPH et présentant des lésions intraépithéliales au niveau du col utérin. Deux polymorphismes, une délétion de 79 pb et la transversion C7732G, ont été associés à la persistance de l'infection et à un plus grand risque de développer une lésion intraépithéliale de haut-grade, respectivement. L'effet fonctionnel des polymorphismes du LCR du VPH33 sur l'expression génique virale, en particulier de la délétion de 79 pb et de C7732G, n'a jamais été caractérisé.

Dans une première étude, l'activité transcriptionnelle des variantes de LCR du VPH33 a été comparée à celle du prototype dans les lignées cellulaires C33A et HeLa par des études de gènes rapporteurs suite à leur introduction en amont du gène de la Renilla luciférase. Nos résultats indiquent que l'activité transcriptionnelle des variantes du LCR du VPH33 reflète leur classification phylogénétique; les variantes issues de la lignée A2 étant les plus actives dans les C33A que le prototype du groupe A1. Des analyses par mutagenèse dirigée ont établi que l'activité accrue des LCRs de ces variantes est due à plusieurs variations, incluant la délétion de 79 pb et C7732G, et dont leur effet individuel est assez modeste. Dans les cellules HeLa, les niveaux de transcription des variantes de la lignée A2 étaient similaires à ceux observés chez le prototype. Cet effet a été attribué à la variation A7879G qui induit la répression du LCR dans ces cellules. Ces résultats indiquent que la combinaison de plusieurs variations dites faibles altère significativement l'activité transcriptionnelle du LCR du VPH33.

D'un autre côté, les résultats rapportés dans notre deuxième étude indiquent que la mutation d'un élément important pour la régulation de la transcription virale peut tout autant altérer l'activité du LCR. La variante rare LCR10 de la lignée B montre une activité transcriptionnelle accrue comparée au prototype. Des analyses par mutagenèse dirigée ont identifié la variation T7791C comme étant responsable de l'activité augmentée du LCR10. L'effet stimulateur de ce polymorphisme est dû au fait qu'un site répresseur pour le facteur cellulaire C/EBP $\beta$  est aboli, tel que déterminé par des essais de liaison à l'ADN in vitro et de gène rapporteur. Un deuxième site C/EBP $\beta$  a également été identifié, cette fois-ci, agissant comme activateur de l'expression génique.

Finalement, dans une troisième étude, nous montrons que C/EBP $\beta$  est un régulateur commun de l'expression des gènes chez les VPHs oncogéniques. Effectivement, plusieurs sites C/EBP $\beta$  sont retrouvés dans la région amplificatrice du LCR des VPHs à haut-risque tel que déterminé par des prédictions basées sur des matrices d'énergie et de fréquence. Certains de ces éléments sont similaires aux sites répresseurs et activateurs retrouvés dans le VPH33. Dans l'ensemble, nos études démontrent l'importance de caractériser l'effet individuel des polymorphismes du LCR pour mieux comprendre les mécanismes de modulation de la transcription des VPHs. De plus, nous montrons que C/EBP $\beta$  joue un rôle central autant dans la répression que dans l'activation de l'expression des gènes viraux.

**Mots-clés** : VPH33, LCR, variante, polymorphisme, phylogénie, luciférase, C/EBP $\beta$ , polarisation de la fluorescence, prédiction, matrice énergétique, matrice de fréquence.

## Abstract

The infection by the oncogenic human papillomavirus (HR-HPV) is associated with the development of cervical cancer. The expression of the E6 and E7 viral oncogenes favors progression towards cancer and is driven by the long control region (LCR) which contains binding sites for both cellular factors and the viral E2 protein. The expression of these oncogenes is often upregulated in cervical cancer due to the integration of the viral genome which frequently disrupts the open reading frame of the repressor E2 gene. Several isolates or variants of HPV type 33 (HPV33) with a slightly different nucleotide sequence were previously identified in clinical samples from women infected with HPV33 and presenting squamous intraepithelial lesions of the uterus. Two polymorphisms, a 79-bp deletion and the C7732G transversion, were previously associated with an increased risk of persistence of infection and of developing high-grade lesions, respectively. The functional effect of polymorphisms in the HPV33 LCR on viral gene expression has never been assessed in particular for the 79-bp deletion and C7732G.

In the first study presented here, the transcriptional activity of the HPV33 LCR variants were compared to that of the prototype in C33A and HeLa cell lines using gene reporter assays following their introduction upstream of the Renilla luciferase. Our results indicate that the LCR activity of the variants reflects their phylogenetic classification. As such, variants from the A2-sublineage were the most active in C33A cells when compared to the A1-sublineage prototype. Mutational analysis showed that the increased in LCR activity was the result of several weakly-acting variations, including the 79-bp deletion and C7732G. In HeLa cells, the lower transcriptional activity of the A2 variants was caused by A7879G which induces repression of the LCR in these cells. These results show that the combination of weakly-acting variations in the LCR significantly alters viral gene expression.

In the second study, we show that the mutation of important regulatory elements also significantly enhances the transcriptional activity of the LCR. The rare HPV33 LCR10 variant from the B lineage exhibits increased transcriptional activity compared to the prototype. Mutational analysis shows that the T7791C variation is responsible for this phenotype. The stimulating effect of this variation is caused by the disruption of a repressor

C/EBP $\beta$  site as determined in an in vitro DNA-binding assay and gene reporter assays. A second site necessary this time for the activation of the LCR was also identified.

Finally, in the third study, we show that C/EBP $\beta$  is a common regulator of the oncogenic HPV gene expression. Several C/EBP $\beta$  sites were found in the enhancer region of the LCR of these HPV types as determined by our prediction using both energy- and frequency-based matrices. Some of these sites were found to be similar to the repressor and activating sites in the HPV33 LCR. Taken together, our studies highlight the importance of characterizing individual variations in the LCR in order to improve our knowledge on the mechanisms involved in the regulation of viral gene expression. We also show that the activating and inhibitory functions of C/EBP $\beta$  are important for the modulation of viral transcription.

**Keywords** : HPV33, LCR, variant, polymorphism, phylogeny, luciferase, polarization fluorescence, prediction, C/EBP $\beta$ , energy-based matrix, frequency-based matrix.



# Table des matières

Résumé.....	i
Abstract.....	iii
Table des matières.....	v
Liste des tableaux.....	ix
Liste des figures.....	x
Liste des sigles.....	xiii
Liste des abréviations.....	xiv
Remerciements.....	xviii
Chapitre premier : revue de la littérature.....	1
1.1 Le virus du papillome humain : agent étiologique du cancer du col de l'utérus.....	1
1.1.1 Classification des papillomavirus.....	2
1.1.2 Épidémiologie de l'infection au virus du papillome humain.....	4
1.1.3 Histoire naturelle de l'infection au virus du papillome humain.....	7
1.1.4 Facteurs prédictifs de la progression de l'infection au virus du papillome humain.....	9
1.1.5 Pathogénèse du virus du papillome humain.....	14
1.1.6 Évasion de la surveillance immunitaire par le virus du papillome humain.....	16
1.1.7 Prévention de l'infection au virus du papillome humain.....	18
1.1.8 Traitements de l'infection au virus du papillome humain.....	19
1.2 Biologie des papillomavirus.....	24
1.2.1 Cycle répliatif des papillomavirus.....	24
1.2.2 Structure du virus.....	28
1.2.3 Organisation génomique.....	29
1.2.4 Protéines précoces.....	30
1.2.5 Protéines tardives.....	41
1.3 Réplication de l'ADN viral.....	42
1.4 Expression des gènes viraux.....	45
1.4.1 Organisation de la longue région de contrôle.....	46

1.4.2	Structure de la longue région de contrôle .....	46
1.4.3	Fonction de l'amplificateur .....	49
1.4.4	Détermination du tropisme des VPHs.....	50
1.4.5	Rôles des facteurs de transcription dans la régulation de l'expression génique ..51	
1.4.6	Autres facteurs de régulation de l'expression des gènes.....	57
1.4.7	Promoteur tardif .....	58
1.4.8	Transition de l'expression des gènes viraux précoces et tardifs .....	58
1.4.9	Étude de la liaison des facteurs de transcription .....	59
1.4.10	Étude des fonctions des facteurs de transcription .....	60
1.5	Variantes du virus du papillome humain .....	62
1.5.1	Caractérisation des variantes.....	62
1.5.2	Pathogénicité des variantes .....	63
1.5.3	Implications fonctionnelles des variations .....	64
1.6	Buts du projet de recherche.....	72
Chapitre deuxième : La combinaison de variations nucléotidiques spécifiques module l'activité transcriptionnelle des variantes de la longue région de contrôle du virus du papillome humain de type 33 .....		
		74
2.1	Mise en contexte de l'article .....	75
2.2	Contribution des auteurs .....	75
2.3	Texte de l'article .....	75
2.3.1	Abstract .....	75
2.3.2	Introduction.....	76
2.3.3	Materials and Methods.....	78
2.3.4	Results.....	80
2.3.5	Discussion .....	84
2.3.6	Bibliography.....	87
2.3.7	Figures and Tables .....	92
Chapitre troisième : Un polymorphisme rare augmente l'activité transcriptionnelle d'une variante de la longue région de contrôle du VPH33 en abolissant un des deux sites de liaison pour le facteur cellulaire C/EBP $\beta$ .....		
		106
3.1	Mise en contexte de l'article .....	107
3.2	Contribution des auteurs .....	107

3.3 Texte de l'article .....	108
3.3.1 Abstract .....	108
3.3.2 Introduction .....	109
3.3.3 Materials and Methods.....	110
3.3.4 Results .....	114
3.3.5 Discussion .....	120
3.3.6 Bibliography.....	122
3.3.7 Figures and Tables .....	128
Chapitre quatrième : Construction d'une matrice poids-position basée sur l'énergie de liaison d'un homodimère de C/EBP $\beta$ à l'ADN et son utilisation pour la prédiction de sites à haute affinité dans la longue région de contrôle de 175 types de VPH .....	144
4.1 Mise en contexte de l'article .....	145
4.2 Contribution des auteurs .....	145
4.3 Texte de l'article .....	145
4.3.1 Abstract .....	145
4.3.2 Introduction.....	146
4.3.3 Materials and Methods.....	148
4.3.4 Results .....	152
4.3.5 Discussion .....	158
4.3.6 Bibliography.....	160
4.3.7 Figures and Tables .....	164
Chapitre cinquième : Discussion.....	181
5.1 Modulation de l'expression génique du VPH33 par les variations dans le LCR.....	182
5.1.1 Le LCR est représentatif de la variabilité génomique du VPH33.....	182
5.1.2 Association entre les variantes non-prototypiques et le développement de cancer du col de l'utérus.....	184
5.1.3 Activité transcriptionnelle élevée des variantes non-prototypiques.....	186
5.1.4 Influence des variations nucléotidiques faibles sur l'activité transcriptionnelle du LCR.....	189
5.1.5 Influence des variations nucléotidiques fortes sur l'activité transcriptionnelle du LCR.....	191
5.2 Modulation de l'expression des gènes viraux par C/EBP $\beta$ .....	196

5.2.1 Bi-fonctionnalité de C/EBP $\beta$ au sein du même promoteur.....	196
5.2.2 La formation de l'amplificateur à la base de l'activation du LCR dans les cellules HeLa? .....	199
5.3. Conservation des sites C/EBP $\beta$ des VPHs muqueux .....	202
Chapitre sixième : Conclusions et perspectives .....	205
6.1 Transition précoce/tardive de l'expression génique virale .....	206
6.2 Effet de la co-ségrégation des variations ailleurs dans le génome viral .....	206
6.3 Détermination des voies de signalisation cellulaires régulant l'expression génique virale.....	207
6.4 Étude de l'amplificateur du VPH33 .....	208
Bibliographie.....	i
Annexe .....	i

# Liste des tableaux

## Chapitre premier

Tableau 1.1 Études sur les variantes du VPH16 .....	66
Tableau 1.1 (suite) Études sur les variantes du VPH16 .....	67
Tableau 1.1 (suite) Étude sur les variantes du VPH16.....	68
Tableau 1.2 Études sur les variantes du VPH18 .....	69
Tableau 1.3 Études sur les variantes des VPH31 et VPH33 .....	70
Tableau 1.4 Études sur les variantes des VPH35, VPH52 et VPH5 .....	71

## Chapitre deuxième

Table 2.1. Naturally-Occurring Polymorphisms in the Long Control Region of Human Papillomavirus Type 33 .....	104
Table 2.2. Nucleotide Variations affecting Transcriptional Activity of the Long Control Region of Human Papillomavirus Type 33.....	105

## Chapitre troisième

Table 3.1. Binding affinities ( $K_D$ ) of C/EBP $\beta$ for different sites as determined in competitive fluorescence polarization assays .....	143
--	-----

## Chapitre quatrième

Table 4.1. Effect of systematic permutations at each position of one half-site on the affinity of a C/EBP $\beta$ homodimer for its consensus binding sequence .....	176
Table 4.2. Binding predictor matrix definition .....	177
Table 4.3. Correspondence of measured and predicted $K_i$ of different DNA sequences for C/EBP $\beta$ .....	178
Table 4.4. Functions of known C/EBP $\beta$ sites .....	179
Table 4.5. Affinities ( $K_i$ ) of C/EBP $\beta$ homodimer for previously described binding sites .	180

# Liste des figures

## Chapitre premier

Figure 1.1 Analyse phylogénétique basée sur l'homologie de la séquence du cadre de lecture de L1 pour 170 types de VPH.....	3
Figure 1.2. Taux d'incidence du cancer du col de l'utérus et de mortalité dans plusieurs régions du Monde.....	5
Figure 1.3 Histoire naturelle de l'infection au VPH.....	8
Figure 1.4 Cycle réplcatif productif des papillomavirus.....	25
Figure 1.5 Modélisation de la capside du virus du papillome bovin de type 1.....	28
Figure 1.6 Organisation génomique du VPH16 et fonctions principales des cadres de lecture.....	29
Figure 1.7 Fonctions de l'oncoprotéine E6 des VPHs-HR.....	31
Figure 1.8 Fonctions de l'oncoprotéine E7 des VPHs-HR.....	33
Figure 1.9 Représentation schématique de l'origine de réplcation minimale du VPB1.....	43
Figure 1.10 Assemblage de l'hélicase de E1 sur l'origine de réplcation minimale.....	44
Figure 1.11 Structure de la longue région de contrôle du VPH33.....	47
Figure 1.12 Représentation schématique des différents membres de la famille C/EBP.....	55

## Chapitre deuxième

Figure 2.1. Phylogeny of HPV33 Long Control Region variants.....	92
Figure 2.2 Transcriptional activity of HPV33 LCR variants in C33A and HeLa cells.....	94
Figure 2.3. Effect of the 79-bp deletion and of the C7732G variation on the transcriptional activity of the prototype HPV33 LCR.....	96
Figure 2.4. Effect of the C7732T and A7874C variations on the transcriptional activity of the HPV33 LCR12 variant.....	98
Figure 2.5. The specific C7537A, C7732G and A7879G variations in combination with the 79 bp deletion are responsible for the enhanced transcriptional activity of HPV33 A2-sublineage variant LCRs.....	100
Figure 2.6. The 79bp deletion increases the activity of the prototypical LCR harbouring the C7537A, C7732G and A7879G variations in C33A cells, but induces a decrease in gene expression in HeLa cells.....	102

### Chapitre troisième

Figure 3.1. The enhanced transcriptional activity of the HPV33 LCR10 variant is attributable primarily to the T7791C variation .....	128
Figure 3.2. The T7791C variation stimulates the transcriptional activity of the HPV33 LCR to a similar extent as inactivation of the C/EBP $\beta$ binding site.....	128
Figure 3.3. The T7791C variation reduces the affinity of C/EBP $\beta$ for its binding site .....	130
Figure 3.4. Identification of two additional binding sites in the LCR of HPV33 with different affinities for C/EBP $\beta$ .....	134
Figure 3.5. Mutational analysis of the three C/EBP $\beta$ binding sites present in the HPV33 prototypical LCR.....	136
Figure 3.6. Effect of the C/EBP $\beta$ LAP and LIP isoforms on the transcriptional activity of the HPV33 prototypical LCR.....	138
Figure 3.7. Modulation of the HPV33 LCR activity by LAP and LIP is altered by C/EBP $\beta$ binding site mutations .....	140
Figure 3.8. Conservation of the minimal enhancer region in the LCR.....	142

### Chapitre quatrième

Figure 4.1. Sequence logo of the C/EBP $\beta$ consensus binding site as determined with the frequency- and energy-based position weight matrices .....	164
Figure 4.2. Correspondence of the measured and predicted binding affinities of C/EBP $\beta$ of C/EBP $\beta$ for different DNA sequences .....	165
Figure 4.3. Comparison of the sets of predicted C/EBP $\beta$ sites of frequency- and energy-based position weight matrixes .....	166
Figure 4.4. Conservation of C/EBP $\beta$ binding sites in the LCR of known HPV types.....	168
Figure 4.5. C/EBP $\beta$ binding sites in the LCR of HPV11, HPV16 and HPV18 show different degree of affinity for C/EBP $\beta$ .....	170
Figure 4.6. The HPV18 BS2 is important for activation of the LCR in HeLa and U2OS cells .....	172
Figure 4.7. Distribution of C/EBP $\beta$ binding site(s) in the enhancer region of the LCR of different HPV types.....	174

### Chapitre cinquième

Figure 5.1 Activité transcriptionnelle du LCR prototypique et des variantes dans des kératinocytes primaires .....	190
Figure 5.2 Influence des variations nucléotidiques sur l'activité transcriptionnelle du LCR du VPH33 dans des kératinocytes primaires.....	191

Figure 5.3 Effets de l'introduction des variations A7874C et A7879G sur l'activité transcriptionnelle du LCR prototypique .....	195
Figure 5.4 Effet de la forskolin sur l'activité transcriptionnelle du LCR prototypique et mutants.....	196
Figure 5.5 Effet d'inhibiteurs et d'activateurs de la voie dépendante de l'AMPc sur l'activité transcriptionnelle du LCR.....	190
Figure 5.6 Effets de la mutation des sites C/EBP $\beta$ l'activité transcriptionnelle du LCR prototypique dans des kératinocytes primaires humains.....	191
Figure 5.7 Modèle de la formation de l'amplificateur sur le LCR du VPH18 .....	195
Figure 5.8 Activité transcriptionnelle du LCR pour différents types de VPHs .....	196

## Annexe

Figure A1 Motifs de localisation cellulaire dans la région N-terminale de la protéine E1.....	ii
Figure A2 La protéine E1 des VPHs atypiques, dont celle du VPH33, est majoritairement nucléaire.....	iii
Figure A3. Développement et validation d'un essai de réplication de l'ADN par les protéines E1 et E2 du VPH33 .....	iv
Figure A4 Le transport nucléaire de la protéine E1 du VPH33 implique un motif NLS de type monopartite .....	vi
Figure A5. La protéine E1 du VPH33 n'interagit pas avec le complexe Cdk2/cyclines E/A in vivo .....	viii
Figure A6. La protéine E1 du VPH33 n'est pas exportée hors du noyau .....	ix



## Liste des sigles

ICTV: *International Committee on Taxonomy of Viruses*

## Liste des abréviations

AA : asiatique-américaine  
Af : africaine  
AMPc : adénosine monophosphate cyclique  
APC : complexe promoteur de l'anaphase (*anaphase-promoting complex*)  
ARNsi : petits ARNs interférents  
As : asiatique  
ATF : facteur de transcription activateur  
ATM : *Ataxia Telangiectasia Mutated*  
CBM : motif de liaison des cyclines  
CC : cancer du col de l'utérus  
CDP : *CCAAT displacement protein*  
CKII : caséine kinase II  
CIN2/3 : lésion intraépithéliale cervical de grades 2 et 3  
CIN3 : lésion intraépithéliale cervical de grade 3  
CMH-I : complexe majeur d'histocompatibilité de classe I  
CO : contraceptif oral  
Co : séquence nucléotidique codante optimisée (*codon-optimized*)  
CR : région conservée (*conserved region*)  
CRE : élément de réponse à l'AMPc  
CRM1 : *chromosome region maintenance 1*  
DBD : domaine de liaison à l'ADN (*DNA-binding domain*)  
DCF : dysplasie corticale focale  
E : européenne  
EGFR : récepteur du facteur de croissance épidermique  
EV : épidermodysplasie verruciforme  
FADD : protéine adaptatrice associée au domaine de mort FAS  
Fluc : *firefly* luciférase  
FT : facteur de transcription  
GFR : facteur de croissance épidermique  
GRE : élément de réponse aux glucocorticoïdes

GTF : facteur de transcription général  
HARSAH : homme ayant des relations sexuelles avec d'autres hommes  
HLA : antigène des leucocytes humains  
HSIL : lésion intraépithéliale squameuse de haut-grade  
Hsp : protéine de choc thermique  
IFN : interféron  
ISG : gène stimulé par l'interféron (*interferon-stimulated gene*)  
KRF-1 : *keratinocyte specific factor 1*  
LCR : longue région de contrôle  
LSIL : lésion intraépithéliale squameuse de bas-grade  
MVA : virus de la vaccine modifiée Ankara  
NA : nord-américaine  
ND : non déterminé  
NES : signal d'export nucléaire  
NHK : cellule épithéliale rénale humaine normale (*normal human kidney epithelial cells*)  
NLS : signal de localisation nucléaire  
OR : rapport des cotes (*odds ratio*)  
Ori : origine de réplication virale  
pAE : site de polyadénylation précoce  
pAL : site de polyadénylation tardif  
PCR : réaction en chaîne de la polymérase  
PDGF- $\beta$  : facteur de croissance dérivé des plaquettes beta  
PEF-1 : *papillomavirus enhancer binding factor 1*  
PIC: complexe de pré-initiation de la transcription  
PML : protéine de la leucémie promyélocytaire  
Pol : polymérase  
pRb : protéine du rétinoblastome  
PR : récepteur de la progestérone  
PT : prototype  
PV : papillomavirus  
Rluc : Renilla luciférase  
RPA : protéine de réplication A  
RR : risque relatif

SNP : polymorphisme nucléotidique (*single-nucleotide polymorphism*)  
SPML mutant de l'épissage (*splice site mutant*)  
SUMO : modificateur apparenté à l'ubiquitine (*Small Ubiquitin-like Modifier*)  
TAD : domaine de transactivation  
TAP-1 : transporteur associé au traitement d'antigène  
TBP : protéine de liaison à la boîte TATA  
TGF- $\beta$  : facteur de croissance transformant beta  
TNF : facteur de nécrose tumorale  
Topo I : topoisomérase I  
TRAIL : ligand inducteur de l'apoptose liée au facteur de nécrose tumorale  
TSS : site d'initiation de la transcription  
VIH : virus de l'immunodéficience humaine  
VLP : particule de type viral  
VPB : virus du papillome bovin  
VPH : virus du papillome humain  
VPH-HR : virus du papillome humain à haut risque  
VPH-BR : virus du papillome humain à bas risque  
WT : sauvage (*wild type*)

Je tiens à témoigner mon éternelle reconnaissance envers mes parents. Sans leur sacrifice et leur soutien constant, je ne serai pas la personne que je suis aujourd'hui. Cette thèse est dédiée à eux, mais également à ma sœur et à mon frère pour leur grande générosité et pour avoir toujours cru en moi.

## Remerciements

Je souhaite adresser mes remerciements les plus sincères à mon directeur de recherche, Jacques Archambault, pour m'avoir accueilli dans son laboratoire, pour avoir partagé sa passion des sciences et pour m'avoir poussé à donner le meilleur de moi-même. Je tiens également à remercier mon co-directeur de thèse, Dr François Coutlée, pour son implication dans ma formation, mais surtout pour m'avoir donné la pique des sciences alors que j'étais une étudiante à la maîtrise dans son laboratoire.

J'aimerais ensuite remercier les membres de mon jury, Dr Anne Gatignol, Dr Gerardo Ferbeyre, Dr André Tremblay et Dr Guy Lemay pour l'évaluation de ma thèse. Je remercie Dr Nathalie Grandvaux, Dr Michel Bouvier et Dr Éric Lécuyer pour avoir siégé en tant que membres de mon comité de thèse.

L'accomplissement de ce travail n'aurait pas été possible sans l'aide et collaboration de mes collègues de travail tant anciens qu'actuels. Des pensées particulières sont adressées à Amélie, Hélène, Fanny, Michaël, Monika, Catherine, Neera et David pour m'avoir assisté au cours de mon doctorat. Je remercie également mes amis à l'IRCM, spécialement Marie et Lorelei, pour m'avoir offert leur amitié et pour leur soutien constant.

Enfin, j'exprime ma gratitude à tous mes proches, famille et amis, qui m'ont toujours soutenu et encouragé au cours de la réalisation de ce doctorat. Je remercie mes grandes amies Geneviève, Myriam et Marielle pour leur support moral. Enfin, je tiens à remercier mes parents Susana et David, ma sœur Johana, mon frère David, *mi abuelita* et tous mes nièces et neveux pour leur amour et simplement pour être présent dans ma vie.

# **Chapitre premier : revue de la littérature**

Le rôle des virus du papillome humain (VPH) oncogéniques comme agent étiologique du cancer du col de l'utérus (CC) est très bien documenté, en particulier pour les types 16 et 18 (1). Un ensemble de facteurs de risque autant viraux qu'environnementaux contribue à la pathogénicité de ces VPHs. Un de ces éléments concerne la variabilité génomique du VPH qui dans plusieurs cas, altère les activités de transcription et de réplication du virus. L'infection par certaines variantes du VPH16, particulièrement par les isolats non-prototypiques, est fortement associée au développement de CC (revu dans (2)). Alors que la majorité des études sur la variabilité intratypique des VPHs portent sur les types 16 et 18, peu d'information est disponible concernant les variantes de types moins prévalents, tels que le VPH33, et surtout par rapport à leur implication dans l'oncogénicité des VPHs. L'objectif premier de cette thèse est de combler en partie cette lacune.

Ce chapitre est une revue de la littérature sur les papillomavirus (PVs), principalement sur les VPHs oncogéniques. Nous verrons comment l'histoire naturelle de l'infection au VPH et l'épidémiologie du CC sont dictées par le cycle de vie des VPHs et les protéines virales qui y sont impliquées. Un regard particulier sera porté sur les variantes du VPH et sur comment les polymorphismes du génome viral peuvent altérer le processus de carcinogénèse.

## **1.1 Le virus du papillome humain : agent étiologique du cancer du col de l'utérus**

Les PVs de la famille *Papillomaviridae* sont un groupe hétérogène constitué de virus infectant un large spectre d'hôtes, dont l'être humain. Plus de 170 types de VPHs ont été identifiés, dont certains causent des lésions bénignes et d'autres sont plutôt associés aux cancers anogénitaux (3). Les prochaines sections portent sur le rôle des VPHs oncogéniques dans le développement de CC.

### 1.1.1 Classification des papillomavirus

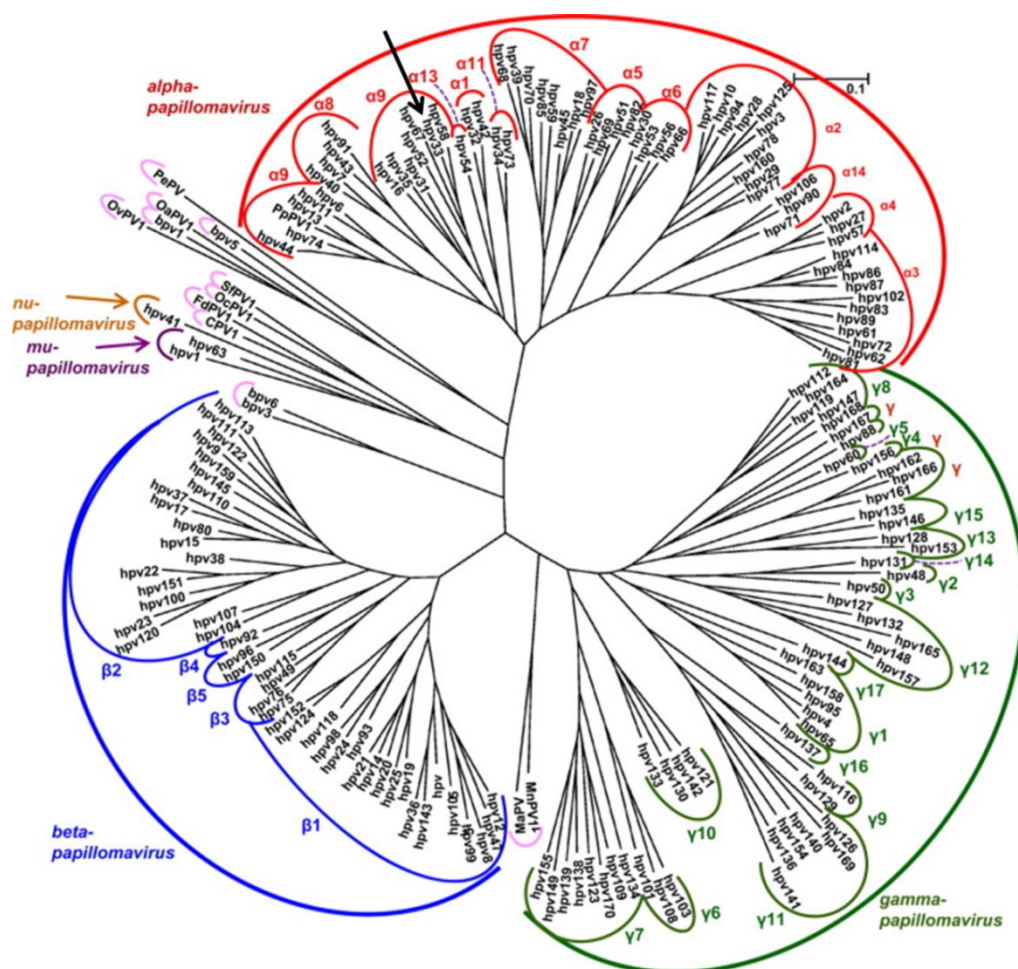
Le système de classification des PVs élaboré par le comité international de taxonomie des virus (ICTV pour *International Committee on Taxonomy of Viruses*) se base sur l'homologie des séquences nucléotidiques du cadre de lecture L1, un des gènes les mieux conservés parmi les PVs (revu dans (4) et (3)). Un nouveau type est défini quand la séquence de L1 révèle moins de 90% d'homologie avec les types connus. Une variante d'un type est établie quand cette homologie diffère de moins de 2% avec le type de référence, ou le prototype, qui est souvent le premier isolat identifié. Les variantes ont été traditionnellement nommées selon des termes géographiques faisant ainsi allusion à la prévalence d'une variante dans une région particulière du Monde. Cette classification est assez juste pour le VPH16 et 18, les plus fréquents pathogènes causant le CC, chez lesquels certains isolats sont plus fréquemment détectés en Asie ou en Amérique par exemple, mais l'est moins pour les autres types tels que le VPH31 et 33. Récemment, il a été proposé de regrouper les variantes en différentes lignées selon l'homologie de la séquence nucléotidique du génome viral entier (5). La section 1.5 de cette thèse est spécialement dédiée à l'importance des variantes dans le développement de CC.

Afin d'uniformiser la classification des PVs avec le système de phylogénie actuelle, les différents types ont été regroupés en espèces distinctes désignées par un numéro. Les différentes espèces infectant l'être humain sont réparties dans 5 genres : alpha, beta, gamma, mu et nu-papillomavirus. Les VPHs regroupés dans ces genres sont biologiquement distincts de par leur tropisme, leur pathogénicité et leur organisation génomique (6). Les VPHs ayant une importance clinique particulière sont couramment classés selon leur tropisme ou leur potentiel oncogénique.

Ainsi, les VPHs à tropisme cutané infectent les épithéliums de la peau et font généralement partie des beta- et gamma-papillomavirus. Les VPHs à tropisme muqueux faisant partie du genre alpha-papillomavirus infectent les muqueuses anogénitales et oropharyngées. Ces derniers sont classés en deux groupes distincts selon leur capacité à induire le CC : les VPHs à haut risque (VPHs-HR) et à bas risque (VPHs-BR) qui sont associés au CC et aux lésions bénignes, respectivement (1, 7, 8). Les VPHs-HR comprennent entre autres les types 16 et 18. Ces types sont responsables à eux seuls de près



de 70% des cancers utérins (9). Les VPHs-BR comprend notamment les types 6 et 11 qui sont associés au développement de verrues génitales et de papillomes (8).



**Figure 1.1** Analyse phylogénétique basée sur l’homologie de la séquence du cadre de lecture de L1 pour 170 types de VPH. La flèche noire représente la position du VPH33 qui fait donc partie de l’espèce 9 du genre alpha. D’après de Villiers (2013) ((3)).

## **1.1.2 Épidémiologie de l'infection au virus du papillome humain**

La prévalence globale de l'infection au VPH est estimée à 10%, mais varie selon les populations et les groupes d'âge étudiés (10). La prévalence de l'infection au VPH est plus élevée chez les femmes âgées entre 20-24 ans avec un taux de 44.8% (11). La prévalence diminue avec l'âge à l'exception des femmes de plus de 55 ans où une recrudescence de l'infection est observée (12, 13). La forte prévalence de l'infection au VPH chez les jeunes femmes serait associée à l'âge au début de l'activité sexuelle chez les jeunes femmes et possiblement à la réactivation de l'infection due à un affaiblissement du système immunitaire ou à une nouvelle exposition chez les femmes plus âgées (12, 13). La prévalence de l'infection au VPH est davantage élevée chez les femmes infectées par le virus de l'immunodéficience humaine (VIH) (14). De plus, la présence de lésions au niveau du col de l'utérus est plus élevée chez ces femmes avec 38.3% des femmes VIH-positives présentant des lésions contre 16.2% des femmes VIH-négatives (15).

### **1.1.2.1 Cancer du col de l'utérus**

Les taux d'incidence du CC ont grandement diminué depuis l'implantation de programmes de dépistage des anomalies au niveau du col de l'utérus (16). Le taux d'incidence de ce cancer dans l'ère pré-dépistage était de plus de 25 cas par 100 000 femmes, mais a significativement diminué au Canada depuis l'introduction du test Pap (17). Effectivement, le taux d'incidence était de 14.2 par 100 000 femmes en 1979 et a diminué à 7 cas par 100 000 femmes au Canada (6 sur 100 000 femmes au Québec) (18, 19). En 2012, le CC comptait pour 1.5% de tous les nouveaux de cancers chez les femmes contrairement à l'ère pré-dépistage où le CC était le cancer féminin le plus prévalent. Cette baisse dans l'incidence du CC démontre le succès d'un programme de dépistage des lésions précancéreuses du col de l'utérus. D'ailleurs, dans les pays où le dépistage n'est pas encore implanté, les taux d'incidence et de mortalité du CC en 2012 était de 15.7 et de 8.3 par 100 000 femmes, ce qui est très semblable aux chiffres dans les pays développés dans l'ère pré-dépistage (voir figure 1.2 pour les taux d'incidence et de mortalité selon les régions du Monde) (20).

L'introduction de la vaccination immunisant contre le VPH16 et VPH18 chez les jeunes filles contribuerait à diminuer davantage l'incidence du CC. Certes, il est encore prématuré pour déterminer si la diminution de la prévalence de l'infection au VPH se traduira réellement en une baisse des cas de CC. Toutefois, il semblerait que l'immunisation contre les types 16 et 18 corrèle avec une diminution de la prévalence des lésions de haut-grade chez les femmes ayant reçu au moins une dose de vaccin entre 2008 et 2012 (21).

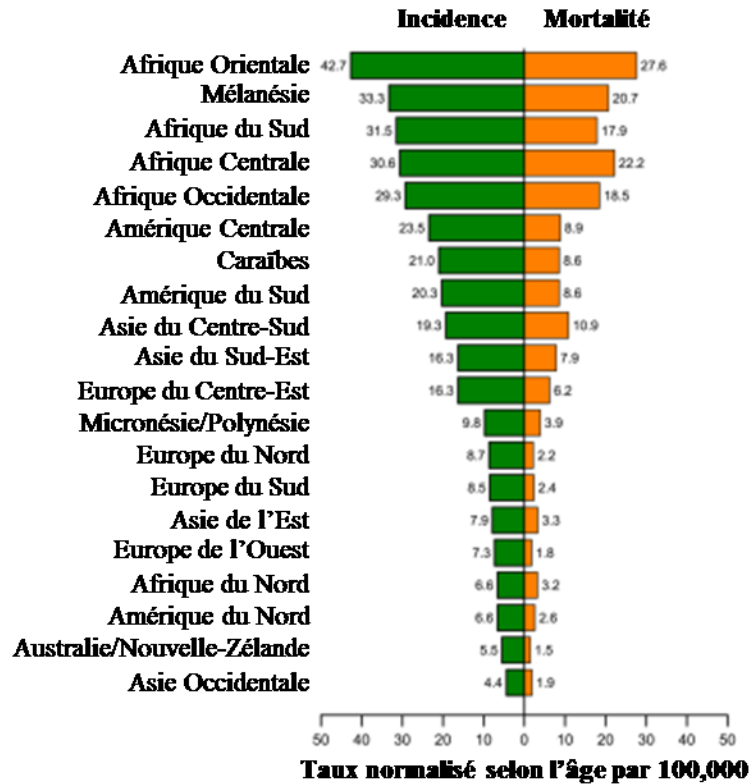


Figure 1.2. Taux d'incidence du cancer du col de l'utérus et de mortalité dans plusieurs régions du Monde. D'après Torre et al. (2015) (20).

### **1.1.2.2 Autres cancers anogénitaux**

Les taux d'incidence du cancer de la vulve et du vagin sont assez faibles et se situent autour de 1.4 et 0.4 cas pour 100 000 femmes au Canada, respectivement (22). Ces deux cancers sont associés à l'infection au VPH, bien que cette association soit moins forte que celle entre les VPHs et le CC. Le VPH est détecté dans 25-40% des cas de cancer de la vulve et de 50 à 100% des cas de cancer du vagin selon les différentes études (23, 24). Le VPH16 (72.5%) est le type le plus prévalent suivi du VPH33 (6.5%). Fait intéressant, 30% des femmes avec un cancer vaginal ont développé au préalable des lésions au niveau du col de l'utérus (25).

Le taux d'incidence du cancer de l'anus au Canada est d'environ 1.0 et 1.1 cas par 100 000 personnes chez les hommes et les femmes, respectivement (22). Certaines populations-cibles sont plus vulnérables telles que les femmes ayant un historique de lésions ou de CC et les personnes immunosupprimées. Chez les hommes VIH-positifs ayant des relations sexuelles avec d'autres hommes (HARSAHs), le taux d'incidence est aussi élevé que 131 par 100 000 personnes (26). Certaines études ont suggéré la possibilité de la propagation du VPH d'un site à une autre région distincte par auto-inoculation à partir d'une infection préexistante (27, 28). Cette explication est basée sur le fait que les mêmes types de VPHs sont retrouvés dans les différents sites d'infection (27, 29). Un facteur de risque de cancer de l'anus est certainement l'immunosuppression (30). Une étude a montré que 95% des HARSAHs VIH-positifs étaient infectés par le VPH (31). Près de la moitié de ces individus présentaient des lésions anales de haut-grade au moment du recrutement. Finalement, le taux d'incidence du cancer du pénis au Canada est de 0.6 cas par 100 000 hommes (22). Le cancer du pénis est plutôt rare et affecte souvent les hommes âgés (26). La présence de VPH, en majorité le VPH16 et VPH18, est détectée dans approximativement 50% de ces cancers (32).

### **1.1.2.3 Cancers des voies aérodigestives**

Les cancers des voies aérodigestives sont plus fréquents chez les hommes que chez les femmes. Au Canada, les taux d'incidence des cancers des voies aérodigestives sont de 3.9 et 0.7 cas par 100 000 personnes pour les hommes et les femmes, respectivement (22).

Le facteur de risque le plus important pour le développement de ces cancers est le tabagisme. Toutefois, certaines études ont montré la présence du VPH dans approximativement 26% des cas, principalement dans les cancers des amygdales (33-35), mais ce nombre augmente à près de 70% dans une étude québécoise (36). Le VPH16 est retrouvé dans 87% des cancers oro-pharyngés (37). Le VPH a également été détecté dans quelques cas de cancers de l'œsophage. Dans une étude, près de 17% des cas étudiés contenaient de l'ADN des VPHs du type 6, 11, 16, 18 et 30 (38).

#### **1.1.2.4 Dysplasie corticale focale**

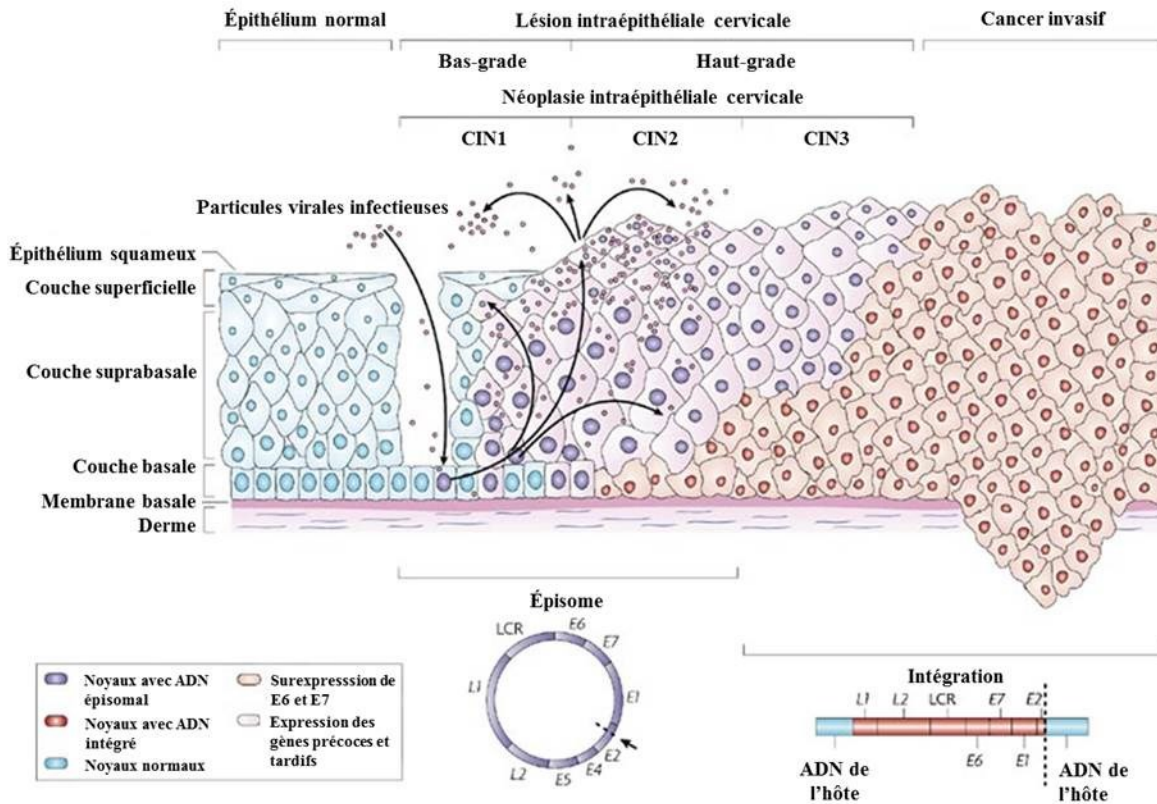
La dysplasie corticale focale (DCF) est une malformation du cortex cérébral qui survient lors du développement fœtal et est une cause commune d'épilepsie chez l'enfant (39). Étonnamment, de l'ARNm et des protéines du VPH16 ont été détectés dans tous les cas de DCF du type IIB (n=50) et non dans les spécimens corticaux contrôles (n=36) dans une étude (40). Dans une autre, de l'ADN du VPH ainsi que du cytomégalo virus, du virus de l'herpès humain de type 6 et du virus de l'herpès simplex de type 1 ont été détectés dans des spécimens de DCF (41). Ces résultats pourraient suggérer un rôle du VPH dans le développement d'anomalies au niveau du cortex, bien que certaines études montrent le contraire (42-44).

#### **1.1.3 Histoire naturelle de l'infection au virus du papillome humain**

Le développement de CC est un processus graduel qui commence souvent par l'apparition de lésions bénignes ou précancéreuses pouvant évoluer jusqu'à un état invasif (45, 46). L'infection au VPH est acquise lors des relations sexuelles avec un partenaire infecté. La plupart des infections sont éliminées assez rapidement après une période de 6 mois pour la moitié des femmes et après plusieurs années après l'acquisition pour la grande majorité (47, 48).

Les changements cellulaires qui surviennent pendant le développement de CC sont catégorisés selon la sévérité des lésions au niveau du col de l'utérus. Les lésions de bas-grade (CIN1 ou LSIL) peuvent progresser vers des lésions de haut-grade (CIN2/3 ou HSIL) dont le CIN3 est le véritable précurseur de CC (49). Dans certains cas, l'apparition de

lésions de haut-grade peut survenir sans le passage par une lésion de bas-grade (50). L'intégration du génome viral corrèle souvent avec les stades avancés de l'infection (voir section 1.1.5). Un schéma montrant l'évolution de l'histoire naturelle est représenté à la figure 1.3.



**Figure 1.3 Histoire naturelle de l'infection au VPH.** Le développement de lésions ou de néoplasies au niveau du col de l'utérus est un processus en continu où les lésions de bas-grade (LSIL ou CIN1) progressent vers un haut-grade (HSIL ou CIN2/3) et dont certaines progresseront vers l'état invasif. Le génome viral s'intègre souvent dans l'ADN de l'hôte au cours de la progression de l'infection de manière que la plupart des CC contiennent de l'ADN viral intégré. Ainsi, l'ADN viral est fréquemment retrouvé sous la forme épisomale dans les phases précoces et sous la forme intégrée, ou un mélange des deux, dans les phases tardives de la progression vers le cancer. Figure adaptée de Woodman, Collins et Young (2007) (46).

Les différents stades de néoplasie intraépithéliale du col de l'utérus (CIN de grade 1, 2 ou 3) ou lésion intraépithéliale squameuse de bas-grade (LSIL) ou de haut-grade (HSIL) sont diagnostiqués par des examens cytologiques et histologiques et gradé selon la sévérité (46). Le taux de progression de CIN1 à CIN3 est estimé à 1% sur une période d'un an. Ce taux augmente à 16% et 25% pour la progression des CIN2 à CIN3 après 2 et 5 ans, respectivement (51). La régression des lésions CIN2 chez les jeunes femmes est approximativement de 70% (52). Chez les femmes plus âgées, le taux est de 30-50% sur une période de 2 ans (53, 54). Certaines lésions CIN3 peuvent progresser vers la phase invasive si elles ne sont pas traitées (55), mais ce processus peut prendre des années (56). Il est difficile de déterminer avec certitude quelles lésions progresseront ou régresseront, d'où le besoin d'identifier des marqueurs de progression.

#### **1.1.4 Facteurs prédictifs de la progression de l'infection au virus du papillome humain**

L'infection au VPH est une cause nécessaire pour le développement de CC, mais d'autres déterminants ou facteurs du virus, de l'hôte et de l'environnement sont également requis pour le développement de cancers associés au VPH.

##### **1.1.4.1 Facteurs viraux**

###### *1.1.4.1.1 Type et variante de VPH*

La majorité des CC et une partie des autres cancers anogénitaux et oro-pharyngés sont causés par les VPHs-HR, notamment par les types 16 et 18 (1, 7, 31). L'oncogénicité de ces virus est en partie déterminée par les oncoprotéines E6 et E7 qui induisent l'immortalisation cellulaire et la prolifération incontrôlée en prévenant l'activité suppressive tumorale de p53 et pRb (voir section 1.2.4). Ainsi les femmes infectées par le VPH16 avec une cytologie normale ont une probabilité de 26.7% de progresser vers des lésions de haut-grade ou vers le CC sur une période de 12 ans si l'infection persiste (57). Pour les femmes infectées par les types 18, 31 et 33, le risque est estimé à 19.1, 14.3 et 14.9%, respectivement.

Plusieurs études ont démontré l'importance des variantes de VPH dans le développement de CC (voir section 1.5 de cette thèse). Ainsi, l'infection par certaines variantes du VPH16 et VPH18 corrèle avec un plus grand risque de progression vers des lésions de haut-grade ou le CC (2). L'infection par les variantes non-Européennes du VPH16 est plus souvent persistante et plus souvent associée aux lésions de haut-grade et au CC (2, 58).

#### *1.1.4.1.2 Persistance de l'infection et intégration*

L'établissement d'une infection au VPH est un prérequis pour le développement d'un CC (59). L'expression continue des oncogènes E6 et E7 est nécessaire au maintien du génome du VPH. Plusieurs études ont montré une association entre la persistance de l'infection au VPH16 et un risque de développer des lésions au niveau du col de l'utérus (risque relatif (RR) entre 9.13 et 11.67 selon les études) qui augmente avec la durée de l'infection (60-64). Ainsi, les femmes infectées par le VPH16 avaient 8.9, 23.8 et 47.4% plus de risque de développer des lésions de haut-grade après 3,5 et 12 ans de suivi, respectivement (45).

L'intégration du VPH dans le génome de l'hôte peut induire la dérégulation de l'expression des gènes précoces viraux. Une grande majorité des cancers contiennent des formes intégrées du VPH (65-67). Contrairement à ce qui était pensé à l'origine, la détection des formes intégrées du VPH n'est pas exclusive aux lésions cancéreuses. Ainsi, dans quelques cas, le génome intégré du VPH est autant décelé dans les lésions bénignes et de haut-grade ainsi que dans les tissus à apparence normale, et ce, souvent en présence d'épisome (68-72). Il semblerait tout de même que l'intégration soit un événement qui corrèle avec le degré de sévérité des lésions et même avec un pronostic négatif (73-76). De ce fait, certaines études ont montré la prédominance des formes intégrées dans les lésions de haut-grade et dans le CC (70, 77, 78).

#### *1.1.4.1.3 Nombre de types et charge virale*

La co-infection avec plusieurs types de VPH est souvent observée dans les lésions au niveau du col de l'utérus. Le rôle de l'infection multiple sur le processus de



carcinogénèse du VPH est encore inconnu. La prévalence de l'infection multiple varie de 9 à 78% selon les populations étudiées, le type d'étude effectuée (transversale ou longitudinale), de l'âge des participantes et même du statut de séropositivité au VIH (79-84). Le risque de développer un HSIL (RR=91.7-424.0) ou un CC (RR=31.8) est habituellement plus élevé pour les femmes infectées par plusieurs types de VPH (surtout lorsque le VPH16 est présent) que pour celles infectées par un seul type (RR=18.6-19.9) (80, 84), bien que certaines études n'ont pas trouvé de lien entre l'infection multiple et le CC (83, 85, 86).

Certaines études ont montré qu'une charge virale élevée est associée à un risque accru de développer un HSIL et un CC, surtout lors de l'infection par le VPH16 (56, 87-92), mais également lors de l'infection par les types 31, 18 et 45 (93, 94), que l'infection soit unique ou multiple. Ainsi, les femmes ayant une charge virale élevée du VPH16 étaient jusqu'à 68 fois plus à risque de développer un HSIL comparativement aux femmes avec une charge virale plus faible (88, 92). Par contre, le risque de développer un CC avec une charge virale élevée est beaucoup plus bas avec les autres types analysés. Ainsi, avoir une charge virale élevée pour les types 31, 18 et 45 (ces deux derniers ayant été quantifiés simultanément) augmente le risque de cancer de 3.2 et 2.6 fois, respectivement (94).

Il est possible de quantifier la charge virale selon le statut du génome du VPH en mesurant le nombre de copies des gènes E6 et E2 par PCR en temps réel. Lors de l'intégration, le cadre de lecture de E2 subit une rupture de manière à permettre la linéarisation du génome viral dans l'ADN de l'hôte, alors que le gène E6 reste intact. Le calcul du ratio E6/E2 permet donc de déterminer le statut du génome du VPH. Un ratio élevé (plus de 2) indiquerait une prédominance de la forme intégrée du VPH tandis qu'un ratio plus faible indiquerait la présence de forme épisomale. Ainsi, il semblerait que parmi les deux formes génomiques, c'est la charge virale épisomale du VPH16 et VPH33 qui est associée à un plus grand risque de développer un HSIL chez les femmes et chez les HARSAHs (72, 95, 96). Par contre ceci n'a pas été observé pour le VPH18 et chez les femmes VIH-positives (87, 97). Il faut noter l'absence de standardisation du prélèvement des échantillons cliniques (frottis vs lavage cervicovaginal) et de la mesure de la charge virale par PCR en temps réel (rapportée en nombre de copies virales/ cellules ou µg d'ADN humain ou valeur Ct). D'ailleurs, ce manque de normalisation est souvent rapporté dans la

littérature où plusieurs études ont montré la grande variabilité de la quantité d'ADN entre les échantillons récoltés au sein d'une même étude (89, 94, 98).

#### **1.1.4.2 Facteurs de l'hôte**

##### *1.1.4.2.1 Tabagisme et pratique sexuelle*

Des nombreuses études démontrent que le tabagisme est un risque important pour le développement un HSIL ou un CC autant chez les anciennes fumeuses que les fumeuses actuelles, bien que la cessation de la cigarette réduise le risque (99-107).

L'âge du début des relations sexuelles, le nombre de partenaires sexuels et les pratiques sexuelles sont aussi des facteurs de risque pour le CC (108-111). Un nombre élevé de partenaires sexuels est associé avec une plus grande probabilité d'acquérir une infection au VPH, dont le risque augmente avec le nombre de partenaires (112). L'âge à la première relation sexuelle aurait un impact sur la présence de HSIL, les femmes ayant eu des relations précoces (avant l'âge de 16 ans), ayant plus de risque de développer un CIN3. Le risque de développer un cancer de l'anus est plus grand chez les hommes ou les femmes pratiquant le sexe anal, mais particulièrement chez les hommes infectés au VIH (31, 113, 114).

##### *1.1.4.2.2 Facteurs génétiques*

La variation naturelle génétique peut rendre certains individus plus susceptibles que d'autres à l'infection au VPH. Les molécules d'antigène des leucocytes humains (HLA) joue un rôle primordial dans la réponse immune contre les pathogènes. Certaines études ont observé qu'avoir un HLA du haplotype DRB1\*1301 ou DRB1\*0501 était protecteur (115, 116), tandis que les femmes porteuses de DQB1\*0602 et DRB1\*1501 développaient plus souvent une infection au VPH16 persistante ou un CC (117, 118). La prédisposition ou la résistance au développement de CC pour un allèle particulier dépendrait de la population étudiée (119, 120).

Certaines variations dans d'autres gènes ou promoteurs, tels que *IL10* et *MTHFR*, confèrent un risque accru de néoplasies ou de CC (121, 122). D'autres études ont montré

que des polymorphismes (SNP) au niveau des gènes de la GST (GSTM1 et GSTT1), de MDM2 (SNP309 T/G), Fas (SNP-670 A/G) et de *TP53* augmentent le risque de CC chez certaines populations spécifiques (123-133). Finalement, l'épidermodysplasie verruciforme (EV) est une maladie génétique autosomique récessive du système immunitaire qui rend sensible aux infections cutanées des VPHs de type 5 et 8. Des mutations dans les gènes *EVER1* et *EVER2* sont retrouvées chez plusieurs individus atteints d'EV (134-136).

### **1.1.4.3 Facteurs environnementaux**

#### *1.1.4.3.1 Immunosuppression et co-infection*

Les personnes infectées par le VIH ou avec greffe d'organes sont plus à risque non seulement de développer une infection persistante au VPH, mais aussi de développer un CC ou autres cancers anogénitaux (137-141). La co-infection de type bactérienne ou virale peut également influencer le développement du CC. Ainsi, les infections par *Chlamydia trachomatis* et par le virus de l'herpès simplex de type 2 augmentent le risque de CC dans certaines études (142-145), mais chez d'autres non (146-148).

#### *1.1.4.3.2 Utilisation des contraceptifs oraux*

Plusieurs études ont montré que l'utilisation prolongée de contraceptifs oraux (CO) sur une période de plus de 10 ans est associée à un plus grand risque de développer des lésions CIN-3 et le CC (107, 149). L'utilisation de CO n'est pas associée au développement de lésions de bas-grade et semblerait même avoir un effet protecteur pour la progression des lésions de bas-grade (107, 150). Ces résultats suggèrent donc un rôle des estrogènes et de la progestérone dans le développement du CC dans les étapes avancées du processus de carcinogénèse et non au stade initial de l'infection au VPH. Les grossesses multiples sont également un facteur de risque de progression (107, 151).

La zone de transformation du col de l'utérus est le site préférentiel de l'infection au VPH. Cette région est particulièrement sensible aux hormones féminines. Des souris transgéniques chez lesquelles les gènes précoces du VPH16 sont exprimés sous le contrôle du promoteur de la kératine 14 humaine ont été exposées à différentes doses d'estrogènes

(152). La localisation et le nombre de tumeurs chez les souris traitées différaient selon la dose d'hormone administrée. Ainsi, les souris traitées avec des doses de 0.10 ou 0.25 mg aux 60 jours développaient plus fréquemment un CC et autres cancers génitaux. À des faibles doses (0.05mg aux 60 jours), les souris ont développé des cancers uniquement dans la zone de transformation du col de l'utérus. Dans les cellules HeLa, de faibles doses d'estrogène (0.1 mM) stimulent la prolifération des cellules et l'expression des gènes viraux et contribueraient donc à la carcinogénèse des VPHs (153). Par contre, à de fortes doses d'estrogène ou de tamoxifène (5 mM), la prolifération cellulaire est inhibée et les cellules montrent un phénotype apoptotique. Le tamoxifène est un antagoniste des récepteurs d'estrogène dans les tissus mammaires qui est administré aux femmes atteintes de cancer du sein (154). À noter que le risque de développer un cancer de l'endomètre augmente pour ces dernières (155).

### **1.1.5 Pathogénèse du virus du papillome humain**

La progression de l'infection vers le CC est un phénomène plutôt rare et résulte de la mise en place d'un environnement cellulaire favorisant la carcinogénèse, tels que l'établissement d'une infection persistante, l'induction de la surexpression des oncogènes E6 et E7, l'intégration du génome viral et les changements épigénétiques (revus dans (156-159)).

#### **1.1.5.1 Persistance de l'infection au virus du papillome humain**

Une faible proportion des individus infectés par le VPH développent une infection persistante (160). La persistance de l'infection par un VPH-HR est le facteur de risque le plus important pour la progression vers le cancer invasif (1, 63). Une infection persistante est reconnue chez une personne infectée quand l'ADN viral est détecté dans au moins deux spécimens consécutifs pendant un intervalle d'au moins 1 an. Toutefois, certains pensent que le virus pourrait être capable d'établir un état de latence où il est indétectable par les méthodes de détection et peut se réactiver après une longue période (158, 161).

### **1.1.5.2 Surexpression des oncogènes E6 et E7**

La fonction première de la protéine E6 est d'interférer avec la protéine suppresseur de tumeur p53 (162). Cette dernière est une composante essentielle de la régulation de l'expression des gènes induits en réponse à un dommage à l'ADN ou à un stress cellulaire. La protéine E6 en formant un complexe avec l'ubiquitine ligase E6AP interagit avec p53 et induit sa dégradation par le protéasome. La réduction des niveaux protéiques de p53 permet de maintenir les cellules infectées dans un état prolifératif permissif à la réplication du génome viral. Quant à la protéine E7, sa fonction principale est d'interagir avec un autre suppresseur de tumeur pRb qui est important pour la transition du cycle cellulaire de G<sub>1</sub> à la phase S (163). Ensemble, ces deux oncoprotéines peuvent promouvoir l'immortalisation et la transformation des cellules infectées en dérégulant le cycle cellulaire. Ceci est d'autant plus accentué lorsque les oncoprotéines sont surexprimées. La surexpression des oncogènes peut être induite lors de l'intégration du génome viral dans l'ADN de l'hôte ou par la présence de variations dans la longue région de contrôle (LCR) qui contient les séquences promotrices et régulatrices de la transcription des gènes viraux (159).

### **1.1.5.3 Intégration du génome viral**

L'intégration est un déterminant important pour la progression vers le CC quoiqu'elle n'est pas essentielle; dans plusieurs cas, l'ADN viral est détecté autant sous la forme intégrée qu'épisomale et parfois même seulement sous la forme circulaire chez certaines femmes ((164-167). L'intégration du génome viral entraîne souvent la destruction du cadre de lecture de E2. La perte de ce régulateur transcriptionnel résulte en l'expression incontrôlée de E6 et E7 (65). L'instabilité génomique induite par ces oncogènes peut faciliter l'intégration du génome dans l'ADN de l'hôte. Il n'existe aucune préférence pour un site d'intégration particulier dans l'ADN de l'hôte, mais quelques études montrent que l'intégration se produit fréquemment dans les sites fragiles communs qui, de par leur nature, sont plus sensibles aux ruptures (168, 169).

#### **1.1.5.4 Changements épigénétiques**

Les changements épigénétiques contribuent également à la dérégulation de l'expression des oncogènes E6 et E7, ce qui expliquerait la présence uniquement d'épisomes viraux chez certains CC (166). La méthylation de dinucléotides CpG dans certains sites de liaison pour la protéine virale E2 (E2BS) dans le LCR peut effectivement altérer l'affinité de l'interaction E2-ADN. La protéine E2 virale agit principalement comme un répresseur de l'activité transcriptionnelle virale des VPHs (revu dans la section 1.2.4 de cette thèse). Trois sites pour E2 à proximité du promoteur (E2BS1, -2 et 3-) sont impliqués généralement dans la répression de l'activité du LCR, alors que le site distal E2BS4 induit plutôt l'activation du LCR (170). La méthylation de sites E2BS1 et -2 au niveau des CpG retrouvé en deux copies pour le VPH16 et VPH18 est souvent associée à une augmentation de l'activité transcriptionnelle virale causée par une inhibition de la liaison de E2 sur l'ADN viral (171-175).

Une autre façon de moduler l'expression génique est de modifier les queues des histones associées à l'ADN viral. La présence d'histones H3 et H4 acétylés sur le LCR corrèle avec une activité transcriptionnelle active de l'épisome viral (176, 177). Les niveaux d'acétylation sont accrus lors de la progression de l'infection par le VPH vers le cancer tandis que les marqueurs H3K9me2 et H3K27me2 associés à la répression du LCR sont perdus (178). Ces résultats suggèrent donc la présence d'une signature de changements épigénétiques spécifiques qui pourraient être utilisés comme biomarqueurs de la progression maligne de l'infection par le VPH.

#### **1.1.6 Évasion de la surveillance immunitaire par le virus du papillome humain**

Les VPHs ont développé plusieurs mécanismes pour échapper au système immunitaire afin de persister dans les cellules infectées.

### 1.1.6.1 Immunité innée

Les cytokines telles que le facteur de croissance transformant  $\beta$  (TGF- $\beta$ ), le facteur de nécrose tumorale (TNF) et les interférons sont sécrétées par les kératinocytes (179, 180). Brièvement, TGF- $\beta$  et TNF inhibent non seulement la prolifération des cellules infectées par le VPH, mais aussi l'expression des gènes E6 et E7 qui sont nécessaires pour le maintien de l'état de transformation dans des cellules immortalisées par le VPH16 (181, 182). La réponse aux interférons (IFN) induite lors de la reconnaissance de motifs moléculaires associés aux pathogènes, souvent des acides nucléiques, active la transcription des gènes de plusieurs cytokines telles que les IFN- $\alpha$  et  $\beta$  (183). Ces IFN sont sécrétés et lient leurs récepteurs ce qui active la voie de signalisation JAK-STAT et la transcription des gènes stimulés par l'interféron possédant des activités antivirales (180). Les oncoprotéines E6 et E7 agissent directement sur l'expression des gènes régulés par les interférons ou ISG (*interferon-stimulated genes*) ainsi que sur l'expression des interférons eux-mêmes (184). La protéine E6 diminue également les niveaux nucléaires de STAT-1, alors que les niveaux d'ARNm restent inchangés. De plus, E6 diminue la liaison de STAT-1 sur des éléments de réponse aux IFN tel que déterminé dans un essai de retard sur gel. La protéine E7 de son côté affecte également la réponse antivirale de l'IFN- $\alpha$  (185). Fait intéressant, l'expression de E7 du VPH6 ou du VPH1 inhibe également la réponse médiée par l'IFN- $\alpha$ , ce qui montre que ce mécanisme échappatoire viral de la réponse immune n'est pas restreint aux types de VPH oncogéniques (185).

### 1.1.6.2 Immunité adaptative

Les lymphocytes T CD4<sup>+</sup> peuvent se commettre en deux lignées différentes, Th1 et Th2, sécrétant des cytokines différentes telles que l'IFN- $\gamma$ , TNF et IL-2, pour les Th1, et IL-4, IL-5, IL-10 et IL-13, pour les Th2 (186). Les lymphocytes T auxiliaires sont notamment nécessaires pour induire la prolifération des cellules B impliquées dans la réponse humorale contre la protéine de la capsid virale L1 et ce processus est exploité pour la confection d'un vaccin prophylactique contre le VPH (187). Les cellules CD8<sup>+</sup> sont les effecteurs de la réponse cytotoxique et sécrètent donc des cytotoxines contre les cellules infectées. Les lymphocytes CD8<sup>+</sup> sont activés lorsqu'un antigène viral est présenté par le complexe majeur d'histocompatibilité de classe I (CMH-I) (188). L'induction de cette

activité cytotoxique est souvent due aux protéines virales E6 et E7 (189). La réponse immune médiée par les lymphocytes T est un processus important pour l'élimination de l'infection. Des hauts niveaux de cellules T CD4<sup>+</sup> et CD8<sup>+</sup> corrélerent avec une meilleure régression des lésions (190, 191). La progression de l'infection vers des stades plus sévères est habituellement reliée à des taux de CD4<sup>+</sup> et CD8<sup>+</sup> plus faibles (192) et corrèle avec un changement de profil des cellules Th1 vers un phénotype de Th2 (193, 194). Le taux de survie sur 5 ans est nettement plus élevé chez les femmes avec un CC ayant des niveaux élevés de CD4<sup>+</sup> et CD8<sup>+</sup> (195). Ces résultats montrent l'importance de présenter une réponse immune adaptée adéquate.

Le cycle viral du VPH est étroitement lié au processus de différenciation des kératinocytes (voir section 1.2.1 de cette thèse). L'adaptation du cycle viral au programme de différenciation cellulaire contribue à la prévention de l'activation de l'immunité adaptée. Ainsi, le VPH infecte les cellules épithéliales où la présence de cellules du système immunitaire est habituellement plus faible. L'expression des gènes viraux varie le long des différentes strates de l'épithélium. Les protéines tardives qui seraient les plus immunogènes sont donc exprimées dans les couches supérieures de l'épiderme où la surveillance immune par les cellules présentatrices d'antigène est plus faible. Les virions produits à la fin du cycle sont relâchés suite à la desquamation cellulaire normale et non par la lyse des cellules, ce qui restreint l'induction d'une réponse inflammatoire. Il semblerait aussi que lors de l'infection au VPH, l'expression du CMH-I soit altérée ce qui affecterait la présentation des antigènes par les cellules de Langerhans dans certains spécimens de CC (196). Ceci pourrait être causé par l'altération de l'expression du transporteur associé au traitement d'antigène (197).

### **1.1.7 Prévention de l'infection au virus du papillome humain**

La prévention de l'infection au VPH était plutôt difficile auparavant et ne s'effectuait qu'en limitant les contacts sexuels. Avec le développement de la vaccination, il est maintenant possible de s'immuniser contre certains types de VPH. Toutefois, une des méthodes les plus efficaces dans la prévention secondaire de la progression vers le stade invasif reste le dépistage des lésions intraépithéliales du col de l'utérus.



### **1.1.7.1 Dépistage des lésions précancéreuses et du cancer du col de l'utérus**

L'utilisation répandue de l'analyse de frottis cytologiques ou test Pap comme stratégie de dépistage des dysplasies précancéreuses a résulté en une nette diminution de l'incidence de CC (16, 198). Le test Pap peut s'effectuer par étalement sur lame d'un échantillon récolté avec une cytobrosse ou en milieu liquide (199). Cette dernière méthode peut facilement être couplée à un test de détection des VPHs-HR, soit par le test d'amplification de signal *Hybrid Capture 2* ou par réaction de polymérase en chaîne (PCR), si les résultats montrent la présence d'anomalies (200).

### **1.1.7.2 Vaccination**

Des vaccins prophylactiques immunisant contre les VPHs-HR de type 16 et qui sont faites de la protéine de la capsid L1 ont été produits. Gardasil® par Merck est un vaccin quadrivalent contre le VPH16 et VPH18 ainsi que contre les VPHs-BR, VPH6 et VPH11, qui causent des condylomes génitaux. Cervarix® de Glaxo Smith Kline immunise contre les types 16 et 18 seulement. L'administration du vaccin Gardasil® aux jeunes filles et garçons de 9 à 17 ans ainsi qu'aux femmes immunocompromises entre 18 et 26 ans est recommandée (Santé et Services Sociaux Québec, Gouvernement du Québec, HPV Vaccination Program, [<http://sante.gouv.qc.ca/en/conseils-et-prevention/vaccin-contre-les-infections-par-les-virus-du-papillome-humain-vph/>], page visitée le 26 juillet 2016). Le vaccin est plus efficace chez les jeunes filles n'ayant jamais été exposées à l'infection au VPH. Il est à noter que ces vaccins immunisent seulement contre deux VPHs-HR. Or, plus d'une trentaine d'autres types, tel que VPH33, sont connus pour causer le CC. Ces vaccins procurent une protection croisée minime contre certains de ces types (201). Un nouveau vaccin nonavalent est disponible et offre une protection élargie contre 9 types oncogéniques (202)

### **1.1.8 Traitements de l'infection au virus du papillome humain**

Une minorité des infections au VPH progresseront vers un état malin chez certains individus. Les traitements actuels consistent à l'ablation des lésions précancéreuses causées par le VPH et le traitement par radiothérapie et/ou chimiothérapie des lésions cancéreuses (203). Ces traitements sont de nature invasive. D'autres alternatives conçues pour inhiber le

cycle viral ou stimuler une réponse immune contre le VPH ont été développées et seront le sujet des prochaines sections.

#### **1.1.8.1 Traitements traditionnels**

Les thérapies actuelles consistent à traiter les lésions bénignes telles que les condylomes avec de l'acide trichloroacétique, la podophyllotoxine et l'imiquimod. Le traitement physique des lésions à un stage plus avancé de l'infection se fait par cryothérapie et chirurgie où les tissus atteints sont excisés. Dans des cas radicaux, l'hystérectomie est requise (203).

Les autres thérapies administrées actuellement aux femmes avec un cancer à un stage précoce combinent la chimiothérapie à la radiation. La combinaison de ces deux approches améliore la survie globale et la survie sans progression de la maladie. La chimio-radiation avec l'utilisation de cisplatine augmente de 22-68% le taux de survie après 5 ans (204). La cisplatine est un composé qui interagit avec l'ADN et induit l'activation des voies de signalisation de l'apoptose faisant intervenir plusieurs composantes de la réponse à un dommage à l'ADN telles que ATR et p53 (205). Pour le traitement des femmes avec un CC à un stage avancé, l'ajout de bévacicumab, un anti-angiogénique, est également associé à une meilleure survie (206).

#### **1.1.8.2 Traitements aux interférons**

L'exposition à de faibles doses d'IFN dans des kératinocytes immortalisés par le VPH16 stimule l'expression des gènes viraux précoces et résulte en une meilleure réplication virale (207). L'inhibition de la voie de signalisation JAK-STAT par l'intermédiaire des oncoprotéines E6 et E7 diminue la réponse induite par l'IFN (208-210). Toutefois, une exposition à de fortes doses d'IFN- $\beta$  inhibe l'expression des gènes viraux et de la réplication virale (210, 211). Ainsi, l'effet des IFN dépend de la concentration utilisée, mais également du contexte cellulaire. Effectivement, l'IFN- $\gamma$  induit des réponses différentes dans les cellules HeLa, CaSki, chez lesquelles l'expression de E6 et E7 est réduite, et dans les SiHa où la transcription des oncogènes est davantage favorisée (212).

### **1.1.8.3 Traitements aux acides ribonucléiques interférents**

Les petits ARNs interférents (ARNsi) permettent de cibler directement les séquences des VPHs, notamment E6 et E7, et peuvent être utilisés à grande échelle pour identifier de nouvelles cibles thérapeutiques (213, 214). Le traitement aux ARNsi est efficace pour réduire la taille des tumeurs en particulier lors de l'administration conjointe de cisplatine dans un modèle de xénogreffe de souris avec des cellules SiHa et CaSki (215). Certains groupes ont utilisé des ARNsi pour provoquer l'inhibition de E6 et E7. Par contre, il est encore incertain si c'est la sénescence des cellules ou l'apoptose qui est induite alors que les deux phénotypes ont été observés (216-219). Il faut noter que les séquences nucléotidiques de E6 et E7 sont différentes entre types et variantes de VPHs et que la conception d'un ARNsi unique ciblant E6 ou E7 pour tous les VPHs serait une tâche difficile. Le choix de la région génique à cibler est donc un élément déterminant important pour une efficacité optimale des ARNsi.

### **1.1.8.4 Traitements aux petites molécules inhibitrices**

Les protéines E1 et E2 sont des cibles intéressantes pour le développement d'un antiviral puisqu'elles sont nécessaires à la réplication du génome viral (voir les sections 1.2.4 et 1.3). L'acide biphenylsulfonacétique, une petite molécule inhibitrice de l'activité ATPase de l'hélicase E1 du VPH6 et VPH18, a été identifiée lors d'un criblage à haut débit (220). L'inhibition de la réplication par cette molécule implique la tyrosine à la position 486 de E1 du VPH6 (Y492 chez le E1 du VPH18) située à proximité du site de liaison de l'ATP et bloque donc la liaison de l'ATP par un mécanisme allostérique (221).

Le recrutement de l'hélicase E1 est favorisé par son interaction avec la protéine E2 qui lie des séquences spécifiques de part et d'autre de l'origine de réplication virale ou ori (revue dans (222) et (223)). La protéine E2 a également des fonctions dans la ségrégation et la partition de l'épisome viral lors de la mitose et dans la régulation de l'expression des gènes viraux (224). L'interaction E1-E2 se fait au niveau de la partie C-terminale de E1, le domaine ATPase, et le domaine de transactivation (TAD) de E2 et peut être inhibée par des dérivés d'indanédione pour le VPH6 et 11 (225-228). Récemment, notre laboratoire a caractérisé une nouvelle interaction entre la protéine cellulaire p80/UAF1 et la protéine E1 du VPH31 qui dépend des 40 premiers acides aminés de E1 (229). La mutation de résidus

clés dans cette région diminue la réplication de l'ADN viral de 70% (230, 231). La surexpression d'un court peptide dérivé de cette région (N40) est également suffisante pour inhiber la réplication de l'ADN viral. D'autres molécules ciblant cette fois-ci l'interaction E6-E6AP (ubiquitine ligase E3 associée à la protéine E6) ont été également développées pour prévenir la dégradation de p53 par ce complexe (214, 232, 233).

### **1.1.8.5 Immunothérapies**

Les protéines virales E6 et E7 sont essentielles à la progression de l'infection au VPH vers le cancer invasif. Du fait qu'elles sont constitutivement exprimées lors du cycle viral, les protéines E6 et E7 sont des bonnes cibles pour l'induction d'une réponse immune. Plusieurs vaccins thérapeutiques sont présentement en développement dont certains montrent un potentiel réel pour le traitement des néoplasies.

#### *1.1.8.5.1 Vaccins à base d'ADN et de peptides*

Le vaccin à base d'ADN consiste à délivrer un plasmide codant pour une protéine virale non-pathogène, mais pouvant induire une bonne réponse immunitaire. L'efficacité d'un tel vaccin dépend entre autres du système de délivrance de l'ADN. Le vaccin p18ConE6E7 qui code pour une fusion de la protéine E6/E7 du VPH18 induit une forte réponse par les cellules T cytotoxiques chez des souris et des singes rhésus. Dans une autre étude, des femmes infectées par le VPH16 et présentant des lésions de haut-grade ont été immunisées avec le vaccin Sig-E7 (detox)-HSP70 qui consiste en une fusion protéine de E7 du VPH16 et la protéine HSP70 (234). Malheureusement, une faible réponse contre E7 a été induite chez la moitié des patientes injectées alors que la régression n'a été observée que chez 3 des 9 femmes évaluées.

Les vaccins à base de peptides consistent à injecter un ensemble de peptides restreint à un antigène des leucocytes humains (HLA) particulier. Dans un essai clinique, des femmes présentant des néoplasies intraépithéliales de la vulve de haut-grade ont été immunisées avec un vaccin contenant un mélange de peptides de E6 et E7 du VPH16 et un adjuvant de Freund incomplet. La régression des lésions n'a été observée que chez 5 des femmes des 20 vaccinées (235). Par contre, cette régression corrèle avec une meilleure

réponse immune des lymphocytes T. Dans une autre étude en phase II où un vaccin à base de peptides de E6/E7 du VPH16 a été injecté à des femmes infectées par le VPH16 et présentant des lésions de haut-grade, aucune infiltration des lymphocytes T au niveau des lésions n'a été observée bien qu'une réponse immune n'ait été détecté par ELISPOT (236).

#### *1.1.8.5.2 Protéine et virus recombinant*

Un vaccin à base de protéines recombinantes contient plusieurs épitopes d'un antigène. Pour en améliorer l'immunogénicité, les protéines E6 et E7 sont fusionnées avec des protéines de la capsid virale ou avec des protéines bactériennes de choc thermique (Hsp). Le vaccin HPV16 L1E7 CVLP contient des protéines chimères L1-E7 assemblées en VLP chimère. Alors qu'une hausse du taux d'anticorps contre L1 et E7 a été mesurée suite à l'injection du vaccin chez des femmes infectées par le VPH16 et présentant des lésions de haut-grade, la régression des lésions n'a été observée que pour 39% des patientes contre 25% dans la groupe placebo (237). Le vaccin SGN-00101 qui contient des protéines Hsp de *Mycobacterium bovis* fusionnées à la protéine E7 du VPH16 n'a pas permis d'obtenir de meilleurs taux de régression alors que seulement 22-35% des femmes ont éliminé leurs lésions malgré une induction plus forte de la réponse immune par les lymphocytes CD4<sup>+</sup> (238, 239).

Finalement, le vaccin MVA E2 utilisant le virus de la vaccine modifiée Ankara (MVA) comme vecteur codant la protéine E2 du virus du papillome bovin (VPB) a été injecté au niveau de l'utérus de 36 femmes présentant des lésions précancéreuses. Après 6 semaines de traitement, 94% des femmes ont montré une régression complète des lésions et dans certains cas, une élimination de l'infection due à l'induction d'une réponse immune cytotoxique et humorale (240). En essai clinique de phase 2, la régression complète de lésions de haut-grade a été observée chez 56% des femmes (241), alors qu'en phase 3, l'administration de ce vaccin corrèle avec une régression complète des lésions du col de l'utérus, de la vulve et de l'anus chez 89% des personnes (242).

## **1.2 Biologie des papillomavirus**

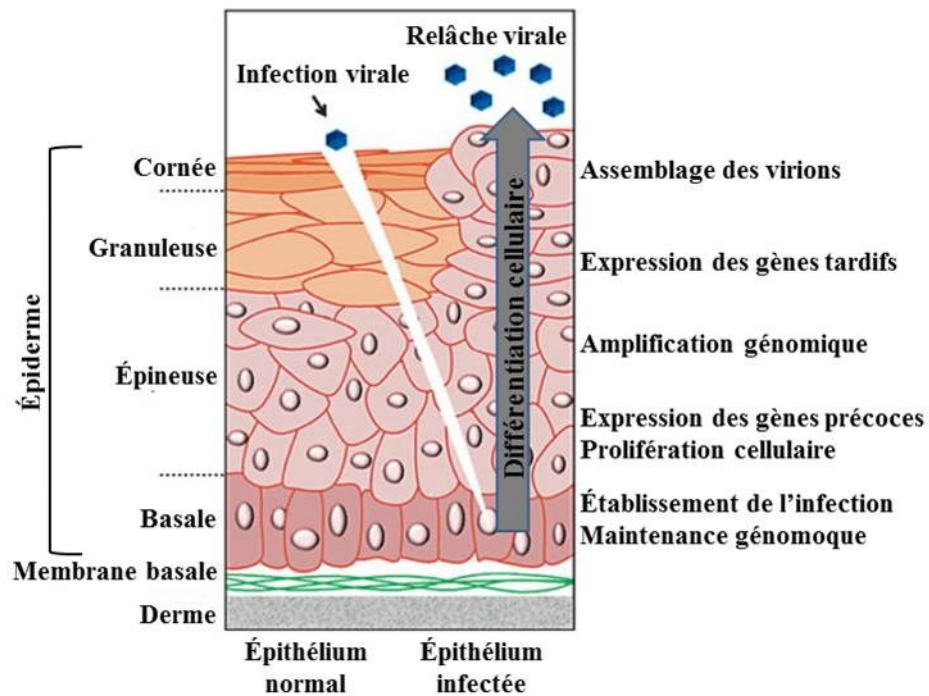
Les prochaines sections concernent surtout le cycle viral, les fonctions des protéines codées et leurs rôles dans la transcription, réplication et oncogénicité des VPHs.

### **1.2.1 Cycle réplcatif des papillomavirus**

Le cycle viral est étroitement lié à la différenciation cellulaire des kératinocytes et dépend de l'efficacité des virus à maintenir leur génome dans les cellules infectées. Le cycle viral présenté à la figure 1.4 peut être divisé en plusieurs phases : l'attachement, l'entrée et le transport viral, la réplication de l'ADN et l'assemblage et relâche des virions.

#### **1.2.1.1 Attachement, entrée et transport intracellulaire**

La plupart des études sur l'étape d'initiation de l'amplification du génome viral sont basées sur le VPB1 (243-245). Lors du processus normal de différenciation cellulaire (en absence d'infection), les cellules basales ayant un potentiel prolifératif illimité sont les seules à se diviser. Une des cellules-filles continuera à se diviser tandis que l'autre migrera vers les couches suprabasales de l'épithélium et se commettra à la différenciation. Cette cellule perd son potentiel de division et acquiert de nombreux changements, autant morphologiques que biochimiques (246). À un stade terminal de leur changement, les kératinocytes entreprennent le processus d'apoptose et sont ultimement desquamées.



**Figure 1.4 Cycle réplcatif productif des papillomavirus.** Le VPH infecte les cellules de la couche basale via une microlésion de l'épithélium. L'infection y est alors établie et maintenue par la réplcation stable de l'épisome viral. L'expression des gènes précoces E6 et E7 favorise l'état prolifératif des cellules infectées qui lors de la migration dans les couches suprabasales (épineuse et granuleuse) perturbe le processus de différenciation cellulaire. Le génome viral est alors amplifié en un grand nombre de copies. L'expression des gènes tardifs est également induite, ce qui permet la production des protéines de la capsidie et l'assemblage subséquent de virions. Les particules virales sont par la suite relâchées par la desquamation normale de l'épithélium (cornée). Selon Lehoux, D'Abramo et Archambault (2009) (157).

L'infection initiale des cellules basales surviendrait suite à l'entrée des particules virales via probablement des microlésions de l'épithélium. Le mécanisme exact de l'attachement et de l'entrée du virus est controversé, mais se ferait par l'intervention des récepteurs de sulfate d'héparine protéoglycanes et possiblement par des interactions avec les composantes de la matrice extracellulaire telle que la laminine 5 (247-249). La capsidie

virale change de configuration et est transférée vers un récepteur différent qui serait l'alpha-6-intégrine (250, 251). Le virus est alors internalisé par endocytose de vésicules de clathrine (252-254) ou de cavéole (255, 256). Le démantèlement de la capsid virale se produit à la surface de la cellule ou dans l'endosome précoce 6 heures après l'entrée du virus. La protéine L2 est alors clivée par la furine, une proprotéine convertase (251, 257). Le clivage de L2 est essentiel pour l'infectivité des PVs et induit l'échappement du complexe L2-génome du compartiment endosomal. Le génome viral est alors relocalisé dans le noyau cellulaire au niveau d'une structure appelée domaine nucléaire 10 (ND10) ou corps nucléaires de protéine de la leucémie promyélocytaire (PML) pour l'initiation de la transcription virale (258).

#### **1.2.1.2 Phase de réplication de maintenance**

Le VPH infecte initialement les cellules de la couche basale avec un potentiel de prolifération et donc de progression dans le cycle cellulaire (259). Le génome viral est par la suite rapidement amplifié afin de permettre l'établissement de l'ADN viral en plusieurs copies, entre 10-200 épisomes lors de la division cellulaire (260). Cette amplification initiale nécessiterait l'intervention des protéines virales E1 et E2. Le génome viral est ensuite maintenu en un nombre de copies constant (261, 262). Lors de la mitose, l'épisome viral est réparti également dans les deux cellules-filles par l'intermédiaire de la protéine virale E2 et la protéine cellulaire Brd4 (225, 263). L'épisome du VPB1 est attaché aux chromosomes mitotiques (264) tandis que le génome des VPHs de type 11, 16 et 18 est attaché aux les fuseaux mitotiques (265).

Le génome viral serait répliqué de manière similaire à l'ADN de la cellule-hôte soit approximativement une fois pendant la phase S de cycle cellulaire (266, 267). Toutefois, plusieurs groupes ont suggéré que l'ADN viral est répliqué plutôt de manière aléatoire i.e. que seulement quelques génomes viraux sont répliqués une ou plusieurs fois durant la phase (264, 268). Ces deux modes ne sont pas contradictoires puisqu'il a été montré que le génome du VPH16 est répliqué de ces deux manières selon le type de cellule utilisé (266).



### **1.2.1.3 Phase de réplication amplificatrice ou réplication végétative**

La réplication végétative a lieu dans les couches suprabasales de l'épithélium où les cellules sont en processus de différenciation terminale (269). Les cellules non-infectées sortent normalement du cycle cellulaire lorsqu'elles quittent la couche basale et ne permettent plus la réplication de l'ADN. Puisque l'amplification du génome viral nécessite des conditions cellulaires permettant la prolifération cellulaire et donc la réplication de l'ADN, le virus doit nécessairement empêcher l'induction des signaux d'arrêt du cycle cellulaire et d'apoptose, entre autres. Ceci est orchestré par les oncoprotéines E6 et E7 dont les fonctions (voir section 1.2.4 de cette thèse) stimulent la progression du cycle cellulaire et favorisent donc par extension la réplication de l'ADN. Ainsi, les cellules infectées demeurent actives et progressent vers la phase S lors de la migration vers les couches suprabasales de l'épithélium (270, 271). À ce moment, le promoteur viral tardif est activé, ce qui induit une expression accrue des gènes nécessaires à l'amplification génomique tels que E1, E2, E4 et E5 (272-274). L'expression de la protéine E4 dans les couches supérieures de l'épithélium corrèle avec l'amplification végétative du génome viral (275, 276). Le génome viral est amplifié en des centaines ou milliers de copies par cellule par le mode de réplication du type cercle roulant (277).

### **1.2.1.4 Assemblage et relâche des virions**

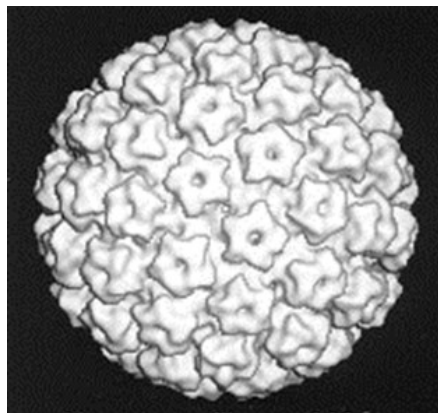
Finalement, lors de la synthèse des virions, les protéines de la capsid L1 et L2 sont produites. L'utilisation de technologies permettant la production de particules virales synthétiques telles que les particules de type virus ou VLP, les pseudovirions et les quasivirions (voir la revue de (278)) a amélioré la compréhension de cette étape du cycle viral. La protéine L1 est capable de s'auto-assembler en capsomère en absence de L2 (279). Par contre, L2 est nécessaire à l'empaquetage du génome viral ainsi qu'à l'infectivité des virions (280-282). La protéine L2 est la première à être exprimée et à être relocalisée dans les noyaux des kératinocytes différenciés. Alors que L2 se localise directement dans les PML, L1 se trouve dispersé dans le noyau. L1 s'assemble préalablement en capsomères ou pentamères dans le cytoplasme selon le modèle du bras invasif. Brièvement, ce modèle consiste en la formation de ponts disulfures entre pentamères. L'extrémité en C-terminal ou

bras d'un pentamère s'étend jusqu'à la protéine adjacente. Ainsi, chaque pentamère reçoit 5 bras et en donne 5 aux pentamères voisins (282-284).

Les capsomères de L1 sont par la suite relocalisés dans les PML où L2 y est intégré (285). Une copie du génome viral est condensée par des histones cellulaires et incorporée dans la capsid virale. Finalement, les cellules de la couche cornée s'allongent et s'aplatissent et leur noyau se dissout. La protéine E1<sup>E4</sup> interagirait avec les kératines de cette couche pour faciliter la relâche des virions (286, 287). Les nouvelles particules virales vraisemblablement non-cytolytiques sont relâchées par la desquamation des cellules (288). Finalement, il est important de mentionner que dans une majorité des CC, l'ADN des VPHs-HR est intégré dans le génome de l'hôte, ce qui a pour conséquence l'avortement du cycle viral productif où la synthèse des virions n'est pas complétée (67, 70, 289).

### 1.2.2 Structure du virus

Les PVs sont un groupe de virus à ADN non-enveloppés. La capsid de type icosaédrique, d'un diamètre d'environ 55 nM (600 Å) est formée de 72 pentamères de la protéine majeure L1 en association avec la protéine mineure L2 (voir figure 1.5) (290). Chaque capsomère est constitué de 5 molécules de L1 et d'une protéine L2 (291).



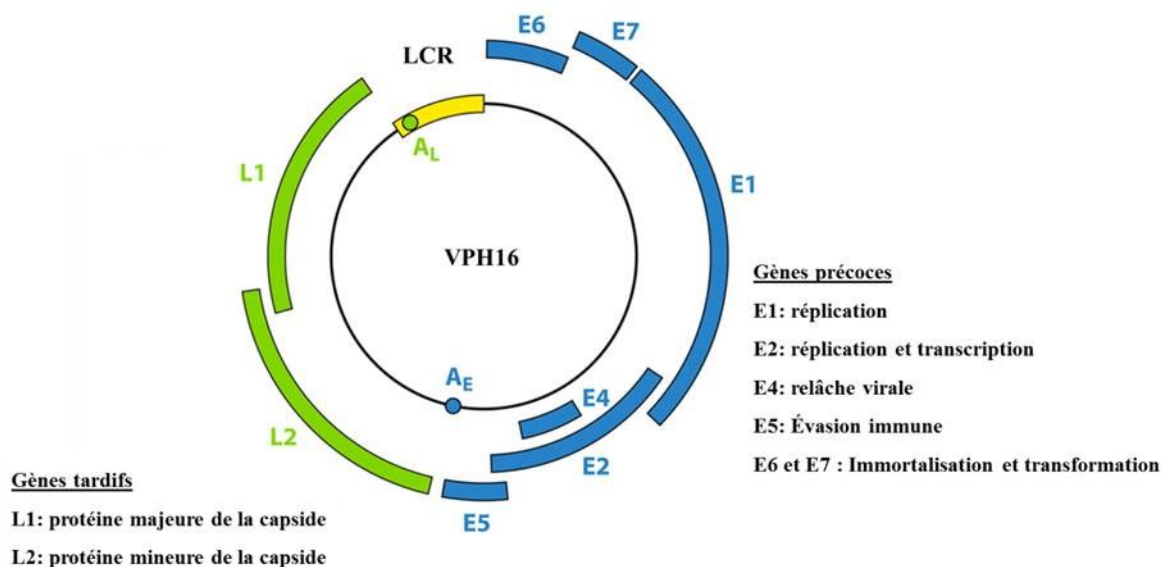
**Figure 1.5 Modélisation de la capsid du virus du papillome bovin de type 1.**

Reconstruction tridimensionnelle d'images obtenues par cryo-microscopie électronique. D'après Baker et al. (1991) (290).

La stabilité de la structure de la capside est assurée par l'interaction entre les différents pentamères selon le modèle du bras invasif, semblable au mode d'interaction observé chez les polyomavirus (292). L'extrémité C-terminale de L1 d'un pentamère donné contacterait le bras d'un monomère d'un capsomère voisin par des ponts disulfure entre les cystéines 175 et 428 formé lors de la maturation de la capside chez le VPH16 (282-284).

### 1.2.3 Organisation génomique

Le génome des PVs consiste en de l'ADN double-brin circulaire d'une taille approximative de 8 kb. Le génome viral est condensé dans la capside par l'interaction avec des histones cellulaires sous la forme d'un minichromosome (293). Tel que mentionné précédemment, l'ADN viral de certains VPHs peut être linéarisé de sorte qu'il s'intègre dans le génome de l'hôte (voir section 1.1.5 de cette thèse).



**Figure 1.6 Organisation génomique du VPH16 et fonctions principales des cadres de lecture.** Le génome est divisible en trois régions : la précoce qui porte les gènes de E6, E7, E1, E2, E4 et E8, la tardive qui porte les gènes L1 et L2 et le LCR, une région non-codante essentielle à la modulation de la transcription et réplication virale. AE et AL sont les sites de polyadénylation précoce et tardive, respectivement. D'après Stanley (2012) (180).

L'épisome viral comporte trois régions distinctes : la région précoce, la région tardive et le LCR tel que présenté à la figure 1.6. La région précoce est constituée des cadres de lecture des gènes E6, E7, E1, E2, E4 et E5. Le cadre de lecture pour le gène E8, donnant la protéine fusion E8<sup>E2C</sup>, est également retrouvé dans cette région pour VPB1, VPH31, VPH16 et chez quelques PVs du lapin seulement (294-297). Les protéines codées par ces gènes permettent en général les phases précoces du cycle viral sont impliquées dans la régulation de la réplication, de la transcription et de la ségrégation virale ainsi que dans la transformation cellulaire (revu dans (298)). La région tardive porte les cadres de lecture des gènes L1 et L2 qui codent pour les protéines structurales majeures et mineures de la capsid, respectivement. Le LCR est une région non-codante de moins de 1 kb. Cette région contient l'ori, le promoteur précoce ainsi que plusieurs sites de liaison pour divers facteurs de transcription cellulaire et pour la protéine virale E2. Cette région est importante pour la régulation de la transcription induite par l'ARN polymérase de type II à partir du promoteur précoce et tardif. Le rôle du LCR dans la régulation de l'expression génique des PVs sera le sujet de la section 1.4.

### **1.2.4 Protéines précoces**

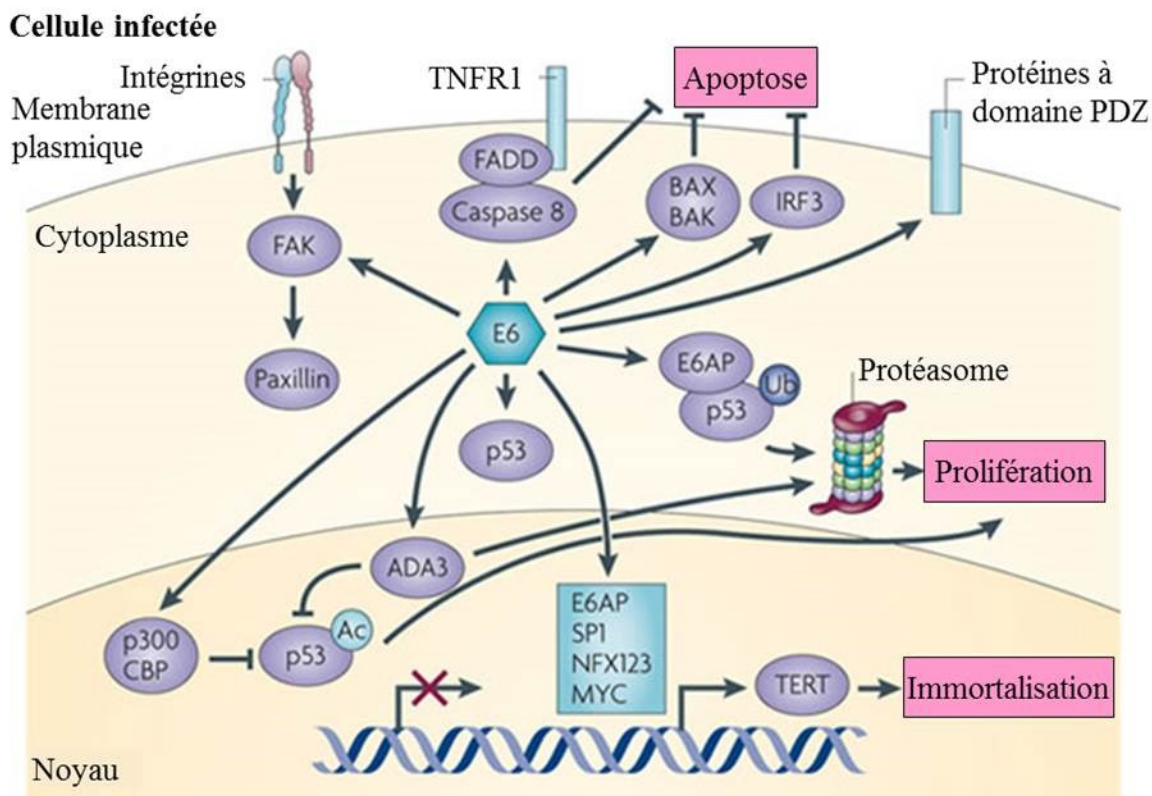
La région précoce du génome des PVs porte les cadres de lectures des gènes précoces E6, E7, E1, E2, E4, E8 et E5. Les sections suivantes portent sur les fonctions des protéines codées par ces gènes dans la réplication et la régulation de l'expression des gènes et dans la transformation cellulaire.

#### **1.2.4.1 La protéine E6**

E6, qui est exprimée à partir du promoteur précoce, est une protéine riche en cystéine et contient deux motifs doigts de zinc. Ces deux domaines connectés par une hélice alpha forment une poche basique qui contacte E6AP (299). Chez les VPHs-HR, un motif de liaison PDZ (PSD-95/DLG/ZO-1) (xT/SxV) est situé dans l'extrémité C-terminale et cible un grand nombre de protéines contenant ce motif. E6 en soit ne possède aucune activité enzymatique et malgré sa petite taille (150 acides aminés pour E6 du VPH16), elle

est capable d'interagir avec une multitude de facteurs (revu dans (162)). Un des partenaires d'interaction les plus importants est certainement E6AP (300).

E6 est nécessaire pour le maintien de l'épisome dans les kératinocytes infectés (301, 302) et en collaboration avec E7, assure une réplication stable de l'ADN (302). La protéine E6, par elle-même, montre une activité d'immortalisation des cellules de souris C127 lorsque l'oncogène ras est activé (303, 304), bien que E7 soit nécessaire pour l'immortalisation de kératinocytes primaires humains (305, 306). Toutefois, la fonction principale de E6 des VPHs-HR est d'inhiber les signaux d'arrêt du cycle cellulaire et d'apoptose induits par E7. Les fonctions de E6 sont résumées dans la figure 1.7.



**Figure 1.7 Fonctions de l'oncoprotéine E6 des VPHs-HR.** L'association de E6 avec ses partenaires d'interaction, notamment p53 et les protéines à domaine PDZ, induit entre autres l'arrêt des signaux pro-apoptotiques, la prolifération et l'immortalisation cellulaire. D'après Moody et Laimins (2010) (307).

Durant un stress cellulaire ou un dommage à l'ADN, le suppresseur tumoral p53 initie les voies de signalisation pour la réparation de l'ADN et peut induire l'arrêt du cycle cellulaire et l'apoptose. La fonction la mieux caractérisée de E6 est d'inhiber les activités de p53 (162, 308). E6 interagit avec p53 et recrute l'ubiquitine ligase E6AP permettant la stabilisation de E6 (309), mais surtout l'ubiquitination de p53 et sa dégradation par le protéasome (162, 310, 311). E6 peut également inactiver p53 en inhibant son interaction avec ces cibles transcriptionnelles et en interagissant avec des régulateurs post-traductionnels tels que l'acétyltransférase CBP/p300 (312-314) et l'enzyme de conjugaison SUMO (modificateur apparenté à l'ubiquitine) Ubc9 (315)

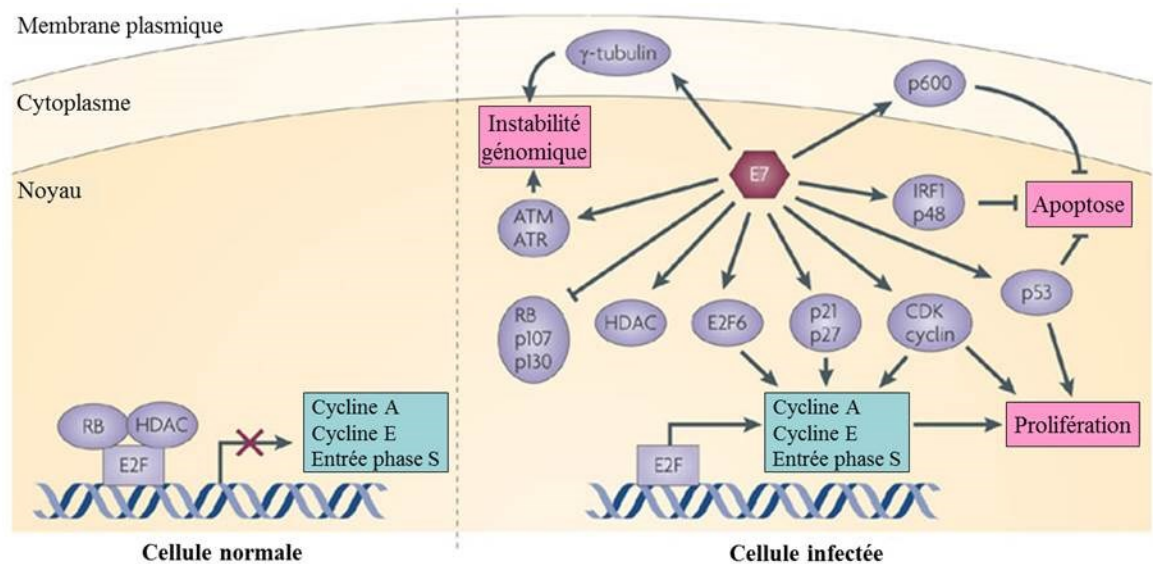
D'autres facteurs cellulaires impliqués dans l'apoptose sont ciblés par E6, notamment les protéines pro-apoptotiques Bax et Bak (316-320) et la protéine anti-apoptotique Bcl-2 (321). De plus, l'oncoprotéine peut inhiber l'apoptose par Fas et TRAIL en réduisant les niveaux des protéines adaptatrices ou effectrices telles que FADD et la procaspase 8 (322-324). Ainsi, E6 peut favoriser la prolifération cellulaire induite par E7 en inhibant l'apoptose. Par ailleurs, les cellules épithéliales sont caractérisées par une polarisation bien marquée qui est maintenue par plusieurs protéines, les complexes Par et Scribble par exemple, dont plusieurs contiennent des domaines PDZ (325). Ces régulateurs sont ciblés par E6 pour la dégradation par le protéasome, ce qui résulte en la perte de l'adhésion cellulaire et de la polarité (revu dans (326)). L'homologue humain de la protéine disque large hDlg faisant partie d'un complexe contenant notamment Scribble est important pour la formation des jonctions cellulaires et pour le maintien de la polarité apicobasale des cellules épithéliales (327-329). Alors que la protéine E6 induit la dégradation de hDlg in vitro (330), les niveaux de cette protéine et la localisation intracellulaire sont inchangées dans un modèle de culture organotypique en radeau de cellules épithéliales où le génome du VPH31 est maintenu (331).

Finalement, E6 favorise l'immortalisation des cellules infectées en induisant l'expression de hTERT et l'activité de télomérase. L'activation du promoteur de hTERT dépend de l'interaction E6-E6AP et est liée à la dégradation de l'élément répresseur de la télomérase transcriptase NFX1 (332-335). Il est à noter qu'une forme tronquée en C-terminal de E6, E6\*, est produite par épissage. E6\* est habituellement un inhibiteur des fonctions de la protéine pleine longueur, mais peut induire la dégradation de certaines

protéines cellulaires telles que Akt et MAGI-1/2 en l'absence de E6 pleine longueur (336, 337).

### 1.2.4.2 La protéine E7

La protéine E7 est une petite protéine d'environ 100 acides aminés. Elle contient en N-terminal deux régions conservées, CR1 et CR2 (163). Cette dernière contient un motif d'interaction LxCxE. Une troisième région conservée (CR3) en C-terminal contient un domaine de liaison au zinc doté de deux motifs CxxC. La protéine E7 peut se dimériser pour former dans certains cas des oligomères (338, 339). La protéine E7 peut être phosphorylée par la caséine kinase II (CKII) et peut subir l'ubiquitination (340, 341). Les fonctions de E7 sont résumées à la figure 1.8.



**Figure 1.8 Fonctions de l'oncoprotéine E7 des VPHs-HR.** L'association de E7 avec ses partenaires d'interaction, notamment avec les protéines de la famille pRb et les protéines à domaine PDZ, induit entre autres la dérégulation du cycle cellulaire, la prolifération cellulaire et l'instabilité génomique. D'après Moody et Laimins (2010) (307).

La protéine E7 favorise la prolifération cellulaire en maintenant les cellules en différenciation dans un état propice à la synthèse de l'ADN (revu dans (163)). L'expression de E7 du VPH16 est suffisante pour induire la transformation de fibroblastes du rat, des cellules rénales primaires de rat en collaboration avec l'oncogène *ras* (342). La région C-terminale participe à la liaison avec le suppresseur tumoral pRb (protéine du rétinoblastome) et autres protéines apparentées telles que p107 et p130 (343). Ces protéines sont impliquées dans l'induction de la transition G1/S en régulant l'expression des gènes cibles de E2F et l'inhibition de Cdk2 (revu dans (344) et (345)). pRb et les autres protéines de poche interagissent avec les membres de la famille E2F pour inhiber l'entrée en phase S, mais aussi pour réguler l'apoptose et la différenciation cellulaire.

L'activité de E2F est régulée selon l'état de phosphorylation de pRb par les complexes Cdk4/6-cycline D et Cdk2-cycline E en fin de G<sub>1</sub> (344, 346). pRb sous sa forme hypophosphorylée régule négativement le cycle cellulaire en interagissant avec E2F. Sous la forme phosphorylée, pRb se dissocie de E2F, permettant l'activation des gènes impliqués dans l'initiation et la progression de la phase S. La protéine E7 lie pRb sous la forme hypophosphorylée et donc cause la dissolution du complexe pRb/E2F pour inhiber le point de contrôle en G<sub>1</sub>/S. pRb est ensuite ciblée pour la dégradation protéasomale par l'intermédiaire de l'ubiquitine ligase cullin 2 (341, 347). E2F est relâché et peut donc activer la transcription des gènes nécessaires à la progression en phase S.

La protéine E7 des VPHs-HR lie E2F avec une meilleure affinité que celles des VPHs-BR (348). E7 s'associe avec E2F6, une composante du complexe répresseur polycomb nécessaire au maintien de la répression de certains gènes de différenciation (349). La levée de cette répression permet donc le découplage de la différenciation de la prolifération cellulaire (350). E7 interagit également avec les déacétylases d'histone (HDAC) de classe I, des co-répresseurs de la transcription et inhibe leur interaction avec le promoteur de E2F2, et ce, indépendamment de sa liaison avec pRb (348). Ceci contribue à augmenter l'expression de E2F et donc de son activité de transcription dans des kératinocytes différenciés et influencerait la progression de la phase S (351).

E7 contribue à la dérégulation du cycle cellulaire en interagissant avec des inhibiteurs de CKI p21<sup>CIP1</sup> (352) et p27<sup>KIP1</sup> (353), des facteurs important pour l'induction de l'arrêt du cycle cellulaire en G<sub>1</sub> en réponse à un dommage à l'ADN. p21 et de p27 inhibent



les activités du complexe Cdk2-cycline E/A (revus dans (354) et (355), respectivement), ce qui favorise la progression du cycle cellulaire. E7 augmente aussi l'activité de CDC25A, une phosphatase essentielle à l'activation du complexe Cdk2-cycline E (356), favorisant ainsi l'entrée et le maintien actif de la phase S. E7 induit l'instabilité génomique en subvertissant les points de contrôle en G<sub>1</sub>/S et en phase M, en induisant des anomalies au niveau des centrosomes et en causant une ségrégation anormale des chromosomes (357-359). Cette aneuploïdie est également induite par la dégradation de pRb qui résulte en l'association de E7 avec la tubuline et l'activation de Cdk2 et des réponses aux dommages à l'ADN (360, 361). La protéine E7 est capable de s'associer avec une des composantes de cette réponse, la kinase ATM (*Ataxia Telangiectasia Mutated*) et de phosphoryler CHK2, BRCA1 et NBS1 (362). Finalement, le ciblage de pRb par la protéine E7 cause l'activation de suppresseur tumoral p53 et donc de l'apoptose (363). L'expression de l'oncoprotéine E6 permet d'inhiber cet effet indésirable pour le VPH.

#### **1.2.4.3 La protéine E1**

Le cadre de lecture E1 est le mieux conservé chez les PVs dû au fait de la fonction essentielle de la protéine dans le cycle viral. La protéine E1 est un membre de la superfamille III (SF3) des hélicases et est nécessaire à l'orchestration de la réplication de l'ADN viral. La protéine E1, d'une taille d'environ 600-650 acides aminés, est composée de trois segments soit la région N-terminale, le domaine de liaison à l'ADN (DBD) et le domaine en C-terminal. Ces trois régions sont nécessaires pour la régulation des fonctions de E1 in vivo, pour la liaison de E1 sur l'ori et pour le déroulement de l'ADN viral, respectivement (364-368) (revu dans (222)). La protéine E1 subit plusieurs modifications post-traductionnelles telles que l'ubiquitination par le complexe APC (*anaphase-promoting complex*), la sumoylation par Ubc9 et la phosphorylation entre autre par Cdk2 (369-371). Le domaine C-terminal est responsable de l'activité ATPase/hélicase de la protéine E1. Ce domaine est sous-divisé en trois régions fonctionnelles : le domaine d'oligomérisation, le module de liaison à l'ATP et le module en C-terminal. Ces régions sont impliquées dans la multimérisation de E1, l'hydrolyse de l'ATP, le maintien de la forme oligomérique de E1 et dans l'interaction avec E2 (367, 372-376). Le DBD de la protéine E1 reconnaît des

séquences spécifiques dans l'ori, les sites E1BS par l'intermédiaire de la protéine E2 (377, 378).

La région N-terminale de E1 contient plusieurs motifs nécessaires au transport nucléo-cytoplasmique de la protéine. Un de ces motifs, la séquence de localisation nucléaire (NLS) de type bipartite [consensus : (K/R)(K/R)-X<sub>(10-12)</sub>-(K/R)<sub>3-5</sub>] (379) est responsable de l'importation nucléaire de E1 par les importines cellulaires alpha (380). Plusieurs résidus à proximité du NLS peuvent être phosphorylés par les MAP kinases ERK1 et JNK2, ce qui stimule l'importation nucléaire de la protéine E1 (381-383). Une séquence d'export nucléaire (NES) [consensus :  $\Phi$ -X<sub>(2-3)</sub>- $\Phi$ -X<sub>(2-3)</sub>- $\Phi$ -X- $\Phi$ , où  $\Phi$  représente un acide aminé hydrophobe] est retrouvée entre les deux modules du NLS et est impliquée dans l'export de E1 vers le cytoplasme par l'exportine cellulaire Crm1 (*Chromosome Region Maintenance 1*) (222). La mutation de ce motif ou l'usage de leptomycine B, un inhibiteur de Crm1, induit l'accumulation nucléaire de la protéine E1 (384, 385). Un motif de liaison aux cyclines (CBM) [consensus : RxL] à proximité du NES lie les cyclines E/A et permet le recrutement de la kinase Cdk2 (385, 386).

La fonction du complexe Cdk2-cycline E /A est de réguler le transport nucléo-cytoplasmique de E1 par la phosphorylation de résidus-clés [consensus : S/T-P], possiblement en synchronisation avec le cycle cellulaire (384, 385, 387-389). Pour le VPH31, l'export nucléaire de E1 est inhibé par la phosphorylation des sérines aux positions 92 et 106, ce qui affecte le maintien du génome dans des kératinocytes primaires (385). La protéine E1 du VPH11 contient une sérine additionnelle (S89) dont la phosphorylation inhibe également l'export nucléaire de E1, comme S93 et S107 (384). À noter que la caséine kinase 2 (CKII) phosphoryle la protéine E1 du VPB1, mais aucun rôle de cette kinase dans la phosphorylation de E1 des VPHs n'a été observé (390, 391). L'accumulation nucléaire de E1 induit l'arrêt du cycle cellulaire en phase S et active une réponse aux dommages à l'ADN par la voie ATM (392, 393). L'induction d'une telle réponse est nécessaire pour l'amplification génomique dans les cellules différenciées (362, 394). Ainsi, la régulation du transport nucléo-cytoplasmique de E1 permet la réplication de l'ADN viral tout en prévenant l'effet antiprolifératif de l'accumulation de E1 dans le noyau. La régulation du transport nucléo-cytoplasmique de E1 est un mécanisme qui semble conservé chez les PVs du fait que le NLS, le NES, le CBM et les résidus phosphorylés par Cdk2

sont retrouvés chez la plupart des PVs. Le VPH33, qui est une exception à cette règle, ne possède pas de NES et CBM apparents, bien que le motif NLS soit présent. Des études en cours dans notre laboratoire indiquent que la protéine E1 du VPH33 n'est effectivement pas exportée du noyau, ce qui suggère d'une part que la protéine crée moins de dommages à l'ADN et de l'autre, de l'existence d'un mécanisme alternatif, autre que l'export nucléaire, pour contrer la voie ATM. Certains résultats préliminaires sont présentés en annexe de cette thèse.

Finalement, la réplication de l'ADN chez les PVs nécessite les protéines virales E1 et E2 et une multitude de facteurs cellulaires. Ainsi, la réplication de l'ADN viral dépend de l'interaction de E1 avec diverses protéines cellulaires du réplisome comme la polymérase ADN primase alpha (Pol  $\alpha$ -prim), la protéine de réplication A (RPA) et la topoisomérase I (Topo I) (395-397). E1 s'associe également avec la protéine cytoplasmique p80/UAF1 qui est relocalisé dans les foyers nucléaires en présence de E1 et E2 (229, 230). La mutation de certains résidus nécessaire pour interagir avec p80/UAF1 réduit la réplication transitoire de l'ADN (229, 230). E1 interagit également avec l'histone H1 et le complexe de remodelage de la chromatine SWI/SNF possiblement pour induire la dérépression de l'ADN viral (398, 399). E1 interagit aussi avec les chaperonnes HSP40 et HSP70, ce qui favorise la liaison de E1 sur l'origine et stimulerait la réplication (400). À noter que la protéine E1 est clivée par les caspases 3 et 7 qui sont activées lors de la différenciation cellulaire (401). Le clivage du N-terminal de E1 est requise pour l'amplification génomique dans des cellules différenciées.

#### **1.2.4.4 La protéine E2**

La protéine E2 d'environ 360 acides aminés est composée d'un TAD en N-terminal et d'un DBD permettant la dimérisation en C-terminal. Ces deux domaines sont connectés par une région flexible dite charnière. Par l'intermédiaire du motif de localisation nucléaire (NLS), la protéine E2 est retrouvée majoritairement dans le noyau et, en présence de E1, elle est relocalisée dans les foci nucléaires, les sites de réplication de l'ADN viral. La protéine E2 subit plusieurs modifications post-traductionnelles telles que la phosphorylation, l'acétylation et la sumoylation de résidus clés contenus dans la région charnière et le DBD. Ces modifications régulent notamment la stabilité, la localisation

cellulaire et donc les fonctions de E2 (402-405). La protéine E2 est normalement dégradée par le protéasome assez rapidement. Son état de phosphorylation et son association avec divers facteurs tels que Brd4 (406, 407) et même avec la protéine virale E1 stabilise et prévient donc sa dégradation (408).

Une des fonctions principales de E2 est de moduler l'expression des gènes viraux. La protéine lie spécifiquement le motif 5'-ACC(N)<sub>6</sub>GGC-3' dont le nombre et position sont conservés dans le LCR des VPHs muqueux (409). Ces sites permettent l'activation ou la répression de la transcription selon la concentration de E2, de l'occupation des sites, du contexte cellulaire et de la présence de sites pour des facteurs de transcription cellulaire. Ainsi, le site de E2 le plus éloigné du promoteur (E2BS4) active la transcription tandis que les sites à proximité du promoteur (E2BS1, -2 et -3) agissent plutôt comme répresseur (410-412). Chez les VPHs muqueux, la protéine E2 est plutôt un répresseur de la transcription et limite donc l'expression des oncogènes E6 et E7. La protéine E2 interagit directement avec les oncoprotéines E6 et E7 et inhibe leurs activités oncogéniques (413, 414).

La protéine E2 joue également un rôle dans la réplication du génome viral. Ainsi, E2 est impliquée dans l'initiation de la réplication de l'ADN viral en recrutant E1 sur l'ori. Suite à la formation du double hexamère de E1, la protéine E2 est alors exclue du complexe. Une autre fonction de la protéine E2 dans le cycle viral est de maintenir le génome viral lors de la mitose. Ainsi, la protéine E2 lie simultanément l'épisome viral et la chromatine par l'intermédiaire de la protéine cellulaire Brd4 (225, 415, 416). En plus de son rôle dans la ségrégation, la protéine Brd4 est aussi nécessaire à une réplication virale optimale étant donné qu'elle co-localise avec E1 et E2 dans les foci de réplication (417). Les protéines E2 des différents PVs lient les chromosomes mitotiques sur des sites distincts. Ainsi, E2 de VPB1 et VPH1 lie les bras des chromosomes tandis que les VPHs muqueux interagissent plutôt avec les régions péracentromériques des chromosomes (418, 419). E2 favorise également l'empaquetage du génome viral dans les capsides, possiblement par son interaction avec la protéine de la capside mineure L2 (258, 420). L'expression ectopique de E2 induit l'inhibition de la croissance des cellules et l'apoptose par des mécanismes dépendants de p53 ou non (421, 422).

Finalement, les différents types de PV peuvent produire plusieurs isoformes de E2 par épissage alternatif tels que E8<sup>E2</sup>, E8<sup>E2C</sup>, E1<sup>E2</sup>, E1M<sup>E2</sup>, E9<sup>E2</sup> et E2-TR (223, 295, 296, 423-425). Ces isoformes qui possèdent le DBD en C-terminal agissent comme des répresseurs de l'expression des gènes et de la réplication virale en compétitionnant pour les E2BS ou en dimérisant avec la protéine pleine longueur (426). L'isoforme E8<sup>E2</sup> module les niveaux de réplication virale et est nécessaire au maintien du génome du VPH31 seulement (296).

#### **1.2.4.5 La protéine E1<sup>E4</sup>**

Le transcrit E1<sup>E4</sup> principal est produit à partir du cadre de lecture de E4 à l'intérieur de celui de E2. La protéine E1<sup>E4</sup> est un produit d'épissage qui contient les cinq premiers acides aminés de la protéine E1 joints à la séquence de E4 (427). Des sous-produits de E1<sup>E4</sup>, tels que les polypeptides E1<sup>E4S/L</sup>, E6<sup>E4</sup> et E2<sup>E4</sup>, sont produits par l'utilisation de sites initiateurs alternatifs ou par clivage protéolytique, mais leurs rôles restent inconnus (428-430). La protéine E4 comporte 3 régions fonctionnelles : une région N-terminale contenant un motif riche en leucine (LLxLL), important pour l'association aux kératines, une région centrale riche en proline nécessaire pour induire l'arrêt du cycle cellulaire en G<sub>2</sub> et une région C-terminale impliquée dans la multimérisation (161, 431-433). L'expression de E4 précède celle des protéines tardives L1 et L2 et est impliqué dans les phases tardives du cycle viral au niveau du contrôle de l'amplification génomique et de la synthèse des virions (434) (272, 435). E1<sup>E4</sup> permet de réorganiser le réseau de kératine de la cellule infectée, ce qui faciliterait la relâche de virions et possiblement la transmission de l'infection (436).

Les niveaux de E1<sup>E4</sup> augmentent alors que le promoteur tardif est activé et corrèle avec la réplication végétative de l'ADN viral (161). L'accumulation de E4 en des structures filamenteuses dans le cytoplasme des cellules infectées a lieu dans les couches médianes et suprabasales de l'épithélium (429, 436-438). E1<sup>E4</sup> induit la phosphorylation et l'ubiquitination de la kératine menant possiblement à sa dégradation protéasomale (439). Parallèlement, E4 subit plusieurs modifications comme la phosphorylation par ERK, un membre de la famille des MAP kinases (440) facilitant ainsi sa multimérisation et son association avec la kératine. L'expression de E1<sup>E4</sup> dans des cellules en prolifération

résulte en l'arrêt du cycle en G<sub>2</sub> accompagné de la séquestration du complexe Cdk1-cycline B via son CBM (441-443). Cet arrêt en G<sub>2</sub> inhiberait les fonctions prolifératives des oncoprotéines E6 et E7 dans les couches supérieures de l'épithélium et favoriserait donc l'amplification végétative de l'ADN (161). La protéine E4 pourrait également induire l'apoptose en s'associant avec les mitochondries, ce qui fragiliserait les cellules infectées et favoriserait la sortie des virions (444).

#### **1.2.4.6 La protéine E5**

La protéine E5 d'environ 40 à 85 acides aminés est constituée de trois domaines hydrophobiques. Elle est associée aux membranes du réticulum endoplasmique et de l'appareil de golgi (445). La protéine E5 possède des activités de transformation, mais contrairement à E6 et E7, son rôle dans le cycle viral est moins bien défini. Chez VPB1, E5 est principalement exprimé dans les cellules basales de l'épithélium, mais au cours de la différenciation, elle est plutôt exprimée dans les cellules des couches supérieures (446). La protéine E5 transforme des cellules en culture murines NIH3T3 et des fibroblastes non-différenciés et serait donc à elle seule l'unique responsable de la production de papillomes par les fibropapillomavirus (445, 447-450). La protéine E5 permet la stimulation de l'activité kinase du récepteur de PDGF- $\beta$  résultant en l'induction de la mitose (451, 452). Toutefois, E5 ne serait pas essentiel à la production de virions dans un modèle de xénogreffe de souris (453).

La protéine E5 des VPHs-HR montre également une activité de transformation, mais moindre que les oncogènes E6 et E7. E5 stimule l'activité du récepteur du facteur de croissance épidermique (EGFR) (454). L'activité de transformation de E5 est également accrue par l'expression de EGF et de E7, ce qui induit la croissance indépendante de l'ancrage et la prolifération cellulaire (455-457). E5 stimule les activités d'immortalisation de E6 et E7 (458) et induit des néoplasies plus sévères au niveau du col de l'utérus chez des souris (459). La protéine E5 réduit l'expression de protéines cellulaires nécessaires à la présentation d'antigènes à la surface des cellules tels que le CMH-1 et CD1d comme moyen d'évasion de la réponse immune (460, 461). Finalement, l'expression de E5 serait requise pour permettre l'amplification génomique du VPH16 et 31 permettant aux cellules différenciées de rester dans un état propice à leur prolifération (273, 274).

#### **1.2.4.7 La protéine E8<sup>E2</sup>**

La protéine E8<sup>E2</sup>C est exprimée à partir d'un produit d'épissage du transcrit de E2. E8<sup>E2</sup>C est constituée des 12 premiers acides aminés de E8 fusionnés au domaine de liaison et de dimérisation en C-terminal de la protéine E2 (462). E8<sup>E2</sup>C est un répresseur de l'expression des gènes précoces qui compétitionne avec la protéine E2 pour les sites E2BS (295, 296, 463, 464). E8<sup>E2</sup>C interagit notamment avec le complexe NCoR/HDAC3 et le recrute au LCR pour induire la répression de la transcription (465, 466). Cette protéine limite également la réplication transitoire du génome viral du VPH31 dans des cellules non-différenciées et différenciées, indiquant ainsi que E8<sup>E2</sup>C joue un rôle autant dans les phases précoces que tardives du cycle viral (467-469). L'expression de E8<sup>E2</sup>C est nécessaire au maintien à long terme du génome viral du VPH31, mais pas pour le génome du VPH16 (295, 296).

### **1.2.5 Protéines tardives**

La région tardive du génome des PVs porte les cadres de lectures des gènes tardifs L1 et L2 et dont leur expression est sous le contrôle du promoteur tardif viral.

#### **1.2.5.1 La protéine L1**

Le cadre de lecture L1 code pour la protéine majeure de la capsidie d'environ 55 kDa (470). Le gène L1 est l'un des mieux conservés parmi les PVs et l'homologie de sa séquence nucléotidique est à l'origine de la base de la classification des PVs en différents types (4). L'entrée et l'attachement viral nécessitent que la protéine L1 interagisse avec les récepteurs de sulfate d'héparine protéoglycanes et possiblement avec les intégrines (247, 250). Dans un cycle viral productif, la protéine L1 est exprimée dans les phases tardives sous le contrôle du promoteur tardif. Les protéines L1 s'assemblent en 72 capsomères pentamériques dans lesquels la protéine L2 y est intégrée pour former la capsidie virale (voir section 1.2.2) (258). La protéine L1 a la capacité de s'auto-assembler en des particules de type VLP présentant une immunogénicité similaire à une protéine virale native (471-473). Les épitopes les plus immunogènes sont retrouvés sur la protéine L1 et donc la grande

majorité des anticorps neutralisants produits lors de la réponse immunitaire cible ces sites (474-477). La production des vaccins prophylactiques Cervarix<sup>TM</sup> et Gardasil<sup>®</sup> est basée sur l'utilisation de cette technologie.

### **1.2.5.2 La protéine L2**

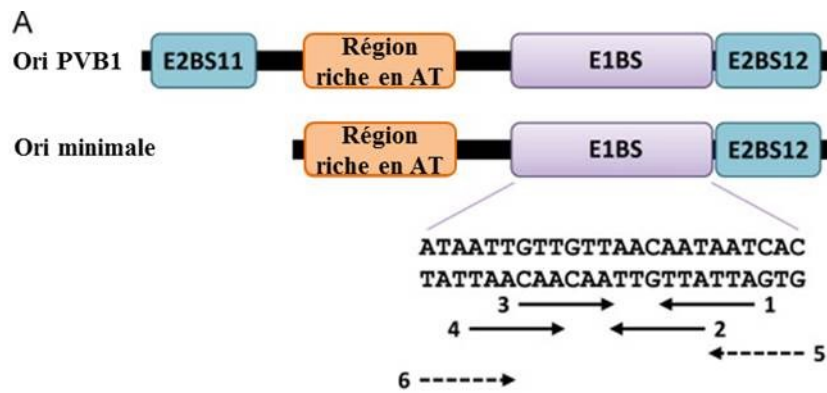
Le gène L2 code pour la protéine mineure de la capsidite d'approximativement 500 acides aminés. Tel que mentionné précédemment, la capsidite virale est constituée principalement de la protéine L1 assemblée en capsomère. La protéine L2 y est intégrée, en environ une copie par capsomère (478). Son expression a lieu lors de la phase tardive du cycle viral et est détectée dans les cellules différenciées dans les couches supérieures de l'épithélium. Contrairement à L1, la protéine L2 n'a pas la capacité de s'auto-assembler en VLP, mais peut y être intégré quand L1 est également produite.

La protéine L2 n'est généralement pas exposée à la surface du virus, à l'exception de quelques acides aminés qui peuvent être ciblés par des anticorps neutralisants (479-481)). Lors de son attachement aux récepteurs de sulfate d'héparine protéoglycanes, la capsidite subit un changement de conformation permettant le clivage de L2 par la furine (251), ce qui expose un épitope assez immunogène (482). L'infectivité des virus serait dépendante de l'intégrité d'une liaison disulfure intramoléculaire entre deux cystéines situées dans cet épitope (483). Finalement, la protéine L2 serait requise pour une meilleure efficacité de l'encapsidation du génome viral possiblement en interagissant avec l'ADN et en relocalisant certaines protéines telles que E2 dans les PML (285, 484, 485).

## **1.3 Réplication de l'ADN viral**

L'origine de réplication viral est constituée d'un segment riche en AT et d'une région de 18 nucléotides qui portent quatre E1BS (E1BS1-4) et trois E2BS (222, 486-488). Les sites pour E1 comportent la séquence consensus 5'-ATTGTT-3' et sont organisés en paires inversées nécessaires pour la dimérisation de E1 sur l'ori (Fig. 1.9) (489, 490). Deux sites additionnels (E1BS5 et E1BS6) sont retrouvés à l'extérieur de la région de 18 nucléotides et seraient impliqués dans l'assemblage du double-trimère, un intermédiaire du double-hexamère chez le VPB1 (489, 490).

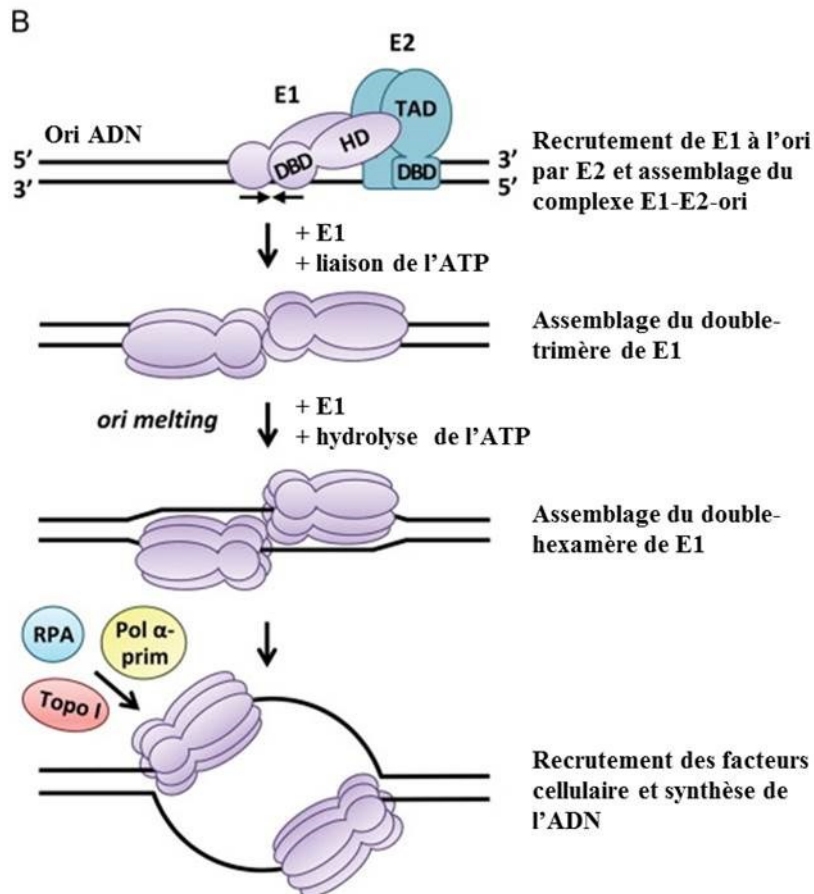




**Figure 1.9 Représentation schématique de l'origine de réplication minimale du VPB1.** Schéma représentant l'origine de réplication du PVB1 constituée de six sites de liaison pour E1 flanqués en 5' par une région riche en AT et en 3' de deux sites pour E2 (E2BS11 et E2BS12).

L'assemblage du double-hexamère présenté à la figure 1.10 nécessite en premier lieu la liaison de E1 à l'origine, ce qui est favorisé par l'interaction avec E2 dont trois de ses sites flanquent l'ori (491-495). Par la suite, un dimère de E1, suivi du double-trimère de E1 sont formés alors que E2 est exclu du complexe (493, 496, 497). Ce complexe est suffisant pour ouvrir l'ADN pour la formation de la double-hélicase sous sa forme active. Le replisome cellulaire constitué entre autres des protéines RPA, Topo I, Hsp40, Hsp70 et l'ADN polymérase primase alpha, est recruté à l'ori pour initier la réplication bidirectionnelle du génome viral (395-397). Chaque hexamère encercle un ADN simple brin pour le déroulement du génome viral en avant de la fourche de réplication (498, 499).

Récemment, la structure de l'hélicase E1 de VPB1 liée sur de l'ADN simple brin a été résolue et montre que l'hélicase forme un anneau autour de l'ADN (500). L'orientation 3' vers 5' de l'hélicase serait produite par le mécanisme de « l'escalier roulant en spirale » où un résidu de phosphate de sucre se trouvant en 5' dans le canal est contacté par la structure en épingle à cheveux d'un monomère de E1 et est alors redirigé vers le 3' suite à l'hydrolyse de l'ATP (222). Le mécanisme est dynamique puisque le nucléotide suivant est alors chargé dans le canal de l'hélicase. Les interactions entre monomères de E1 seraient suffisantes pour induire les changements de conformation de l'hélicase (501).



**Figure 1.10 Assemblage de l'hélicase de E1 sur l'origine de réplication minimale.** Représentation schématique de l'assemblage du double-hexamère de E1. Les protéines E1 et E2 sont recrutées à l'ori pour la formation successive du complexe E1-E2-ori, du double-trimère et du double-hexamère de E1. Les facteurs cellulaires impliqués dans la réplication de l'ADN viral (l'ADN polymérase  $\alpha$ , RPA et Topo1 notamment) sont ensuite recrutés. Selon Bergvall, Melendy et Archambault (2013) (222).

L'ADN viral est maintenu en un nombre de copies constant lors de la phase de maintien du cycle viral et serait obtenu par deux mécanismes. Le premier mode, dit ordonné, permet la réplication de l'ADN une fois seulement au cours de la phase S tel qu'observé pour le virus Epstein-Barr (502, 503). Le deuxième mode, dit aléatoire, permet

la réplication de quelques épisomes seulement plusieurs fois par phase S, indiquant que ce type de réplication serait indépendant de la protéine E1 (268, 504). Bien que le mode exact de réplication de l'ADN soit encore controversé, une étude indique que ces deux approches ne sont pas incompatibles et seraient dictés par la cellule hôte (266, 504). Le génome du VPH16 est répliqué une fois par phase S dans les cellules W12, mais d'une manière aléatoire dans les cellules NIKS telle qu'observée pour le VPH31 (266). La surexpression de E1 dans les cellules W12 changerait la réplication vers un mode aléatoire.

## 1.4 Expression des gènes viraux

L'expression des gènes des PVs dépend de la machinerie transcriptionnelle de l'hôte dont l'holoenzyme est composé de l'ARN polymérase II, des facteurs de transcription généraux (GTF pour *general transcription factor*) TFIIA, TFIIB, TFIID, TFIIIE, TFIIF et TFIIH et du complexe médiateur (coactivateur). L'assemblage du complexe de préinitiation de la transcription requiert la liaison de TBP (*TATA-binding protein*), une sous-unité de TFIID, et le recrutement séquentiel des autres GTF et du complexe médiateur (revu dans (505)). La présence de sites de facteur de transcription sur les régions promotrices permet de réguler l'expression des gènes en augmentant l'affinité des GTF pour le promoteur et en stimulant l'activité de l'ARN polymérase.

La transcription des gènes des PVs par les promoteurs précoces et tardifs se fait sur le brin positif de l'ADN viral. Le promoteur précoce (p97, p105 et p97, pour les VPHs 16,18 et 31, respectivement) est retrouvé dans le segment 3' du LCR tandis que le promoteur tardif (p670, p811 et p742, pour les VPHs 16,18 et 31, respectivement) est retrouvé dans le cadre de lecture de E7 (506). Quelques promoteurs alternatifs ont été identifiés pour le VPH16 et 18 (428, 462, 507). Les ARNm produit par les différents promoteurs sont polycistroniques et subissent l'épissage à partir de divers sites (508). Les ARNm précoces et tardifs sont polyadénylés au site de polyadénylation précoce (pAE) situé en amont du gène L2 et au site de polyadénylation tardif (pAL) situé dans le LCR pour les transcrits précoces et tardifs, respectivement (508, 509). La plupart des gènes précoces sont transcrits à partir du promoteur précoce viral et sont exprimés le long du cycle viral. Toutefois, leur expression est souvent réduite dans les phases tardives. La protéine E4 est

une exception puisqu'elle est principalement exprimée tardivement lors de l'infection. Les gènes tardifs sont plutôt transcrits à partir du promoteur tardif. L'expression des gènes, autant par le promoteur précoce et tardif, est régulée par une panoplie de FT viraux et cellulaires avec des sites de liaison dans le LCR.

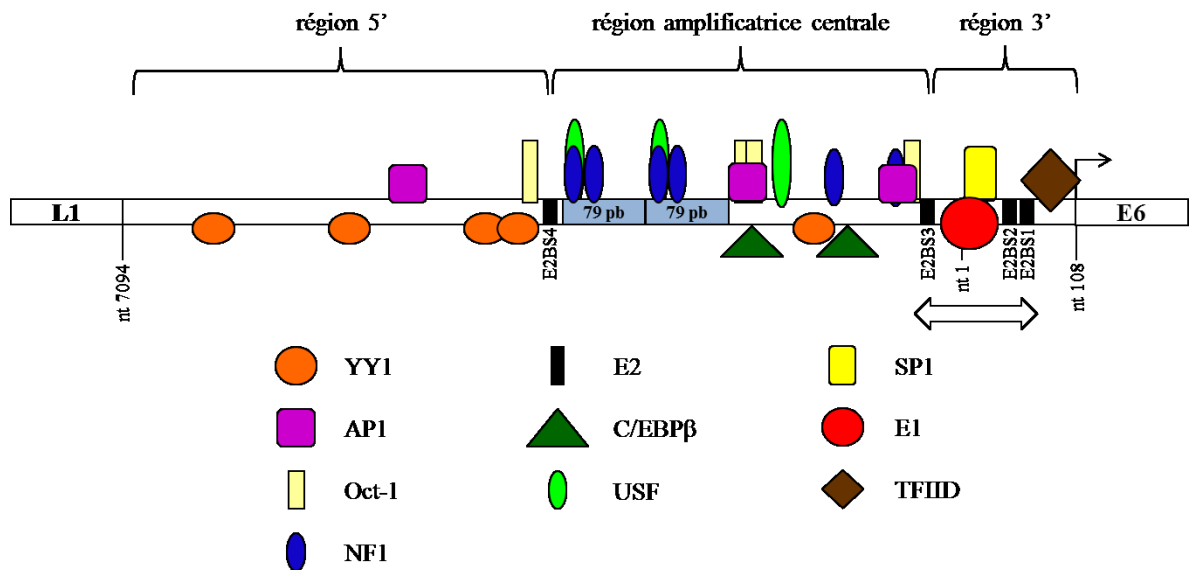
### **1.4.1 Organisation de la longue région de contrôle**

La chromatine est empaquetée par les nucléosomes qui sont constitués d'octamère de 2 copies des histones H2A, H2B, H3 et H4 (revu dans (510)). Les nucléosomes régulent l'expression des gènes en contrôlant l'accès des FT et de la machinerie de transcription sur l'ADN. Les queues des histones subissent plusieurs modifications covalentes telles que l'acétylation chez les histones H3 et H4 (AcH3 et AcH4, respectivement) et la diméthylation de lysines spécifiques (2meH3K4) qui permettent le relâchement de l'ADN et donc un meilleur accès des FT. Le génome des VPHs est également associé aux nucléosomes (176, 177, 398, 511). Deux de ces nucléosomes sont positionnés dans la région centrale du LCR et chevauchent l'ori et le promoteur précoce du VPH16 et module l'accès à l'ADN viral, la régulation de la transcription et de la réplication virale (511). La présence de nucléosomes sur l'ADN viral varie selon l'état de différenciation des cellules infectées (176). Alors, que les histones AcH3, AcH4 et 2meH3K4 sont plus fortement associés au promoteur précoce dans des cellules non-différenciées (un indicatif que la transcription est active), il y a une plus grande présence de ces mêmes histones sur le promoteur tardif dans les cellules différenciées, ce qui corrèle avec l'induction de la transcription des gènes tardifs (176).

### **1.4.2 Structure de la longue région de contrôle**

Le LCR est un segment non-codant du génome viral qui contient le promoteur précoce et plusieurs séquences régulatrices nécessaires à la modulation de l'expression des gènes précoces des PVs (revu dans (409) et (512)). Un schéma du LCR du VPH33 est présenté à la figure 1.11. D'une longueur approximative de 800 pb, le LCR est compris entre les cadres de lecture des gènes L1 et E6 en 5' et en 3', respectivement. Il existe peu d'homologie entre la séquence nucléotidique du LCR des différents PVs sinon la présence

de quatre E2BS, de sites pour des FT tels que AP1, SP1 et YY1, d'une ori et du promoteur précoce. Le LCR est divisible en trois régions distinctes: la région en 3' à proximité du



gène de E6, la région centrale et la région en 5' à proximité du gène L1.

**Figure 1.11 Structure de la longue région de contrôle du VPH33.** Le LCR du VPH33 est une région non-codante de 924 pb qui est flanquée par les cadres de lectures de L1 et E6 à ses extrémités 5' et 3', respectivement. Le LCR est divisible en trois régions fonctionnellement distinctes : la région 5', la région amplificatrice centrale et la région 3'. La présence de quatre sites de liaison pour E2 (E2BS1-4) et d'un site pour la liaison de l'hélicase de E1 sont des caractéristiques conservées chez les VPHs. Le LCR du VPH33 porte également une parfaite duplication de 79 pb. Les sites potentiels pour les facteurs cellulaires YY1, AP1, Oct-1, NF1, C/EBPβ, USF, Sp1 et pour TFIID sont indiqués. L'origine de réplication est indiquée par une flèche.

La région 3' du LCR est délimitée par E2BS3 et par le codon initiateur de E6. Cette région contient le site d'initiation de la transcription (TSS) en aval de la boîte TATA (consensus : TATAAA). L'initiation de la transcription nécessite également la liaison

d'activateurs tels que Sp1 et AP1, dont leurs sites de liaison sont à proximité de la boîte TATA. La région 3' contient une région dite silencieuse où un site de liaison pour la protéine de déplacement CCAAT ou CDP/Cut (*CCAAT displacement protein* aussi référée à *human Cut protein* (Cut)) est retrouvé chez le VPH16 (513, 514). Ce facteur est spécifiquement exprimé dans les cellules basales de l'épithélium et non dans les cellules différenciées, suggérant son implication dans le couplage de la transcription des gènes viraux durant les différentes étapes du cycle viral (515).

Les sites E2BS1, E2BS2 et Sp1 flanquent la boîte TATA. L'occupation de ces sites par leur facteur respectif permet donc de contrôler l'activation de la transcription et de la réplication virale. La liaison de Sp1 et de E2 est mutuellement exclusive (516). À une haute concentration de E2, Sp1 est déplacé, ce qui active la réplication du génome viral et inhibe la transcription (517-519). Ainsi, la protéine E2 est responsable de réguler négativement l'activité transcriptionnelle du LCR en compétitionnant pour la liaison à l'ADN avec le complexe d'initiation de la transcription (518-520). La répression induite par E2 est levée par la mutation des sites E2 à proximité du promoteur (521). La région 3' porte également l'ori qui consiste en plusieurs sites de liaison pour la protéine E1 flanqués de deux sites de liaison de E2 (E2BS1 et E2BS2) de chaque côté et de séquences régulatrices adjacentes (voir section 1.3 de cette thèse).

La région centrale du LCR, ou région amplificatrice, est comprise entre les sites E2BS3 et E2BS4. La fonction première de ce segment est d'induire l'activation de la transcription dans les cellules épithéliales (522). Cette région contribue donc à la détermination du tropisme cellulaire des VPHs (523-525). Ainsi, cette région active la transcription chez le VPH18 dans des lignées cellulaires de CC telles que les cellules HeLa et SiHa et non dans des cellules d'origine non-épithéliale (526). La région centrale contient plusieurs sites de liaison pour des activateurs tels que NF1, AP1 et Oct1 (409, 527). La région amplificatrice minimale du VPH16 et 18 contient des sites pour les facteurs encore inconnus PEF-1 (*papillomavirus enhancer binding factor 1*) et KRF-1 (*keratinocyte specific factor 1*), respectivement (528-530). La région centrale contient également des sites de liaison pour les récepteurs des glucocorticoïdes (GRE) et de la progestérone (PR) qui induit l'activation du LCR du VPH16, ce qui contribuerait à la transformation des cellules

infectées et corrèlerait avec un risque plus élevé de développer le CC chez les femmes ayant eu plusieurs grossesses (531-534)

Finalement, la région 5', délimitée par le codon de terminaison de L1 et par E2BS4, contient le site de polyadénylation tardif. Cette région est impliquée dans l'activation de la transcription chez le VPH31, dans une moindre mesure que la région centrale, de par le site E2BS4 activateur (410-412), mais aurait une activité inhibitrice chez le VPH6 (535, 536). La région 5' contribue également à la modulation de la réplication virale, car la délétion de cette région dans le VPH31 augmente la réplication transitoire et stable du génome viral (537). Cette région contient également plusieurs motifs d'attachement à la matrice qui permettent l'adhésion des épisomes viraux à la matrice nucléaire dans les cellules infectées (538, 539). Ces motifs sont importants pour l'activation de l'expression génique qui serait favorisée par le rapprochement du génome viral aux éléments nécessaires à la transcription. La matrice nucléaire consiste en du matériel insoluble de type fibrogranulaire qui contiendrait entre autre de la lamine, des ribonucléoprotéines, des polymérases et certains FT (540).

### **1.4.3 Fonction de l'amplificateur**

Telle que mentionné précédemment, plusieurs activateurs de la transcription lient la région centrale du LCR. Le recrutement de ces éléments favorise la formation de l'amplificateur (*enhanceosome*), un complexe nucléoprotéique d'ordre supérieur, et permet l'organisation des FT en un agencement tridimensionnel qui stimule l'assemblage du complexe de pré-initiation de la transcription (PIC) (541). L'amplificateur permet la modification physique de l'ADN viral et favorise un recrutement séquentiel et optimal de la machinerie transcriptionnelle sur le promoteur de façon spécifique en réponse à un stimulus donné (542, 543). Ainsi, la fonction de l'amplificateur est de permettre l'activation coopérative et synergique de l'expression des gènes précoces par plusieurs éléments activateurs. L'amplificateur du VPH18 est constitué entre autre de la protéine architecturale HMG-I(Y) (*high-mobility-group*), du co-activateur CBP/p300 et plusieurs autres FT (543). La présence d'un site AP1 est essentielle pour l'assemblage de l'amplificateur au niveau de la région amplificatrice du LCR (543). L'amplificateur du VPH18 semblerait être

spécifique au type cellulaire, ce qui expliquerait en partie le fait que ce LCR soit particulièrement actif dans les cellules HeLa (541, 544, 545). La conservation du site de liaison pour AP1 suggère que la formation de l'amplificateur est un processus important pour l'activation du LCR chez les VPHs.

#### **1.4.4 Détermination du tropisme des VPHs**

La phylogénétique des PVs basée sur l'homologie de la séquence nucléotidique reflète assez bien leur pathogénicité. De ce fait, les VPHs-HR et -BR sont classés dans le groupe des alpha-papillomavirus. Les PVs appartenant à ce genre présentent certaines différences au niveau de leur tropisme. Ainsi, les VPHs-HR et le VPH6 infectent préférentiellement les muqueuses génitales tandis que le VPH11 est souvent retrouvé dans la région oropharyngée (546-548). Certains VPHs du groupe alpha tels que le VPH7 infectent plutôt des sites cutanés (549). Le mécanisme à la base de la diversité du tropisme des VPHs au sein du même groupe n'a pas encore été élucidé, mais des différences dans la régulation de l'expression génique virale pourraient, en partie, contribuer à fournir cette spécificité (550-552). De ce fait, les VPHs muqueux sont habituellement plus actifs dans des cellules des muqueuses telles que les lignées de CC que dans les lignées cutanées telles que les cellules HaCat (550, 553-555). La région centrale du LCR semblerait être impliquée dans la détermination du tropisme des VPHs, bien que le mécanisme exact derrière ce phénomène soit encore inconnu (409). Il est possible qu'un facteur de transcription exprimé dans les cellules épithéliales induise cette spécificité, mais aucune étude ne peut démontrer sans équivoque un rôle pour un facteur en particulier excepté peut-être le facteur KRF-1 ou PEF-1 dont leur liaison n'a été montrée que sur le LCR du VPH18 et VPH16 (409, 528-530). D'un autre côté, la spécificité pourrait être plutôt fournie par plusieurs FT ubiquitaires (556). Effectivement, l'amplificateur qui se formerait seulement dans les cellules compétentes pour le cycle viral pourrait participer à la spécificité de l'activité transcriptionnelle des VPHs (541, 543). Une troisième hypothèse est qu'un cofacteur spécifique encore inconnu serait recruté au LCR par l'interaction avec une autre protéine cellulaire ou même après la formation de l'amplificateur.



### **1.4.5 Rôles des facteurs de transcription dans la régulation de l'expression génique**

Les FT sont classifiés en familles distinctes selon la similarité du domaine de liaison à l'ADN. Ainsi, il existe 4 grandes classes de FT qui sont les suivantes : les facteurs contenant un homéodomaine, les protéines contenant un doigt de zinc, les protéines contenant une glissière de leucine et les facteurs de types hélice-spirale-hélice (557, 558). Les FT lient de courtes séquences nucléotidiques spécifiques formant des homo- et des hétérodimères entre membres de la même famille ou de classe différente.

Les VPHs affichent des activités transcriptionnelles différentes qui varient selon le contexte cellulaire, ce qui reflète probablement les interactomes distincts de chacun des types (545, 559-563). La multitude de combinaisons possibles permet de diversifier la régulation de l'expression génique de ces VPHs. Plusieurs facteurs de transcription interagissent avec le LCR du VPH tels que Sp1, AP1, C/EBP, NF1, Oct-1 et YY1. Les prochaines sections se veulent une revue des publications les plus significatives sur le rôle de ces facteurs dans la modulation de l'expression des gènes viraux.

#### **1.4.5.1 La protéine E2**

Une des fonctions de la protéine virale E2 est de moduler l'expression des gènes précoces viraux. En fonction de sa concentration et de l'occupation de ses sites, la protéine E2 induit la répression ou l'activation de la transcription des gènes des PVs (voir section 1.2.4 de cette thèse). La protéine reconnaît et lie la séquence consensus 5'-ACC(N)<sub>6</sub>GGT-3', où les nucléotides espaceurs (N) sont préférablement riches en AT (223, 564). Le nombre de sites (habituellement de quatre) ainsi que leur localisation dans le LCR sont assez conservés chez les VPHs muqueux du groupe alpha. Le nombre de E2BS dans le génome des PVs du groupe beta est assez variable pouvant aller jusqu'à une occurrence de 17 sites chez le VPB1, tel que montré par des essais d'empreintes de l'ADN et de retard sur gel, et dont plusieurs sont atypiques et ne correspondent pas au consensus (223, 565). Le site distal du promoteur (E2BS4) active la transcription tandis que les sites à proximité du promoteur (E2BS3, -2 et -1) agissent plutôt comme répresseur (170, 410-412, 517). En général, plusieurs sites pour des FT cellulaires chevauchent les sites de E2 et peuvent donc

moduler l'activité de E2 selon leur occupation (409, 566). À noter que E2BS1, E2BS2 et E2BS3 flanquent les sites de liaison de E1 et sont requis pour la réplication optimale de l'ADN virale et pour la production de particules virales (487, 567, 568).

#### **1.4.5.2 Sp1 et AP1**

Un site Sp1 a été initialement identifié dans la région promotrice du VPH16 et VPH18 par des essais de retard sur gel (518, 569, 570). L'analyse de la séquence de différents LCR indique que ce site est hautement conservé chez les VPHs (409). Sp1 est un important activateur de l'expression génique des VPHs et sa liaison sur le LCR est modulée par l'occupation de sites pour le répresseur YY1 et pour la protéine virale E2 (517, 521, 571). Les niveaux de Sp1 sont plus élevés dans les cellules épithéliales que dans les fibroblastes, indiquant que Sp1 joue un rôle dans la spécificité de la transcription (572). AP1 est impliqué dans la régulation de la prolifération, la transformation cellulaire et l'apoptose (revu dans (573)). AP1 est composé d'homodimère ou d'hétérodimère des membres de la famille Jun (Jun, JunB et JunD), Fos (FosB, Fra-1 et Fra-2), et ATF (*activating transcription factor*) (ATF $\alpha$ , ATF-2 et ATF-3) (574). AP1 reconnaît la séquence consensus 5'-TGA(G/C)TCA-3' (575). L'expression différentielle des membres des familles Jun et Fos dans différents tissus et par leur partenaires d'interaction permet la formation d'une multitude de dimères avec des fonctions distinctes (541, 574, 576, 577). AP1 est un activateur important de l'expression génique des PVs et pourrait être impliqué dans la spécificité de l'activation de la transcription virale dans les cellules épithéliales (527, 578-581). Plusieurs sites de liaison pour AP1 sont présents dans la région centrale et dans le 3' du LCR, dont deux sont hautement conservés chez les VPHs autant par rapport à leurs positions qu'à leurs fonctions sur la modulation de l'expression des gènes viraux tels que montré par des essais de retard sur gel ou d'empreinte de l'ADN et par des essais fonctionnels (582-587). Toutefois, la stimulation de la région centrale du LCR chez le VPH11 ne dépendrait pas de ce facteur contrairement aux VPHs-HR (588). À noter que l'interaction de AP1 sur le LCR serait à l'origine de la formation de l'amplificateur chez le VPH18 (541).

### **1.4.5.3 NF1 et Oct-1/2**

NF1 est impliqué dans le développement des tissus et dans la différenciation cellulaire (revu dans (589)). Alors que NF1 active et réprime la transcription cellulaire, il semblerait que ce facteur n'agisse que comme un activateur de l'expression des PVs (409). Des essais d'empreintes de l'ADN montrent que NF1 lie plusieurs motifs 5-TTGGC-3 contenus dans la région centrale du LCR du VPH16 (590). La régulation de la transcription virale par NF1 semblerait être un mécanisme commun chez les PVs alors que la présence de sites de liaison est une caractéristique conservée chez la plupart des types (578, 588, 590-592). Le facteur Oct-1 est impliqué dans la répression de la transcription du VPH18 dont plusieurs sites sont trouvés dans la région centrale du LCR (593, 594). Un site pour le facteur KRF-1 qui chevauche un de ces éléments induit l'activation du LCR du VPH18 par compétition pour la liaison à l'ADN avec Oct-1 (530). Chez le VPH16, Oct-1 activerait plutôt l'expression des gènes par son interaction avec NF1 et dépendrait de la présence de Oct-2, un autre membre de la famille POU (595-597).

### **1.4.5.4 YY1**

YY1 appartient à la famille des protéines à doigt de zinc. Ce facteur est un régulateur important de l'expression d'un grand nombre de gènes impliqués dans la prolifération et la différenciation cellulaire (revu dans (598)) et de la transcription chez le VIH et le virus de l'herpès simplex de type 1 (599, 600). YY1 est également un élément important pour la régulation de l'expression des gènes précoces des VPHs en agissant autant comme un activateur que comme un répresseur transcriptionnel. Le LCR des VPHs de type 16 et 18 contiendrait une dizaine de sites YY1 dont quelques-uns ont bien caractérisés dans des essais de retard sur gel (571, 601, 602). Certains de ces sites chevauchent des sites AP1 et Sp1 et induiraient la répression de l'activité transcriptionnelle par compétition pour la liaison à l'ADN (571, 601, 603-605). Chez le VPH31, YY1 est important pour induire l'activité de la région centrale amplificatrice (536).

YY1 peut exercer une double fonction sur l'activité promotrice chez certains PVs (602, 606). Ainsi, pour le VPH18, YY1 induit la répression du LCR dans les cellules HepG2 (602, 607, 608), mais, en interagissant avec C/EBP $\beta$ , peut également induire l'activation du LCR dans les cellules HeLa (609). La surexpression de YY1 pourrait

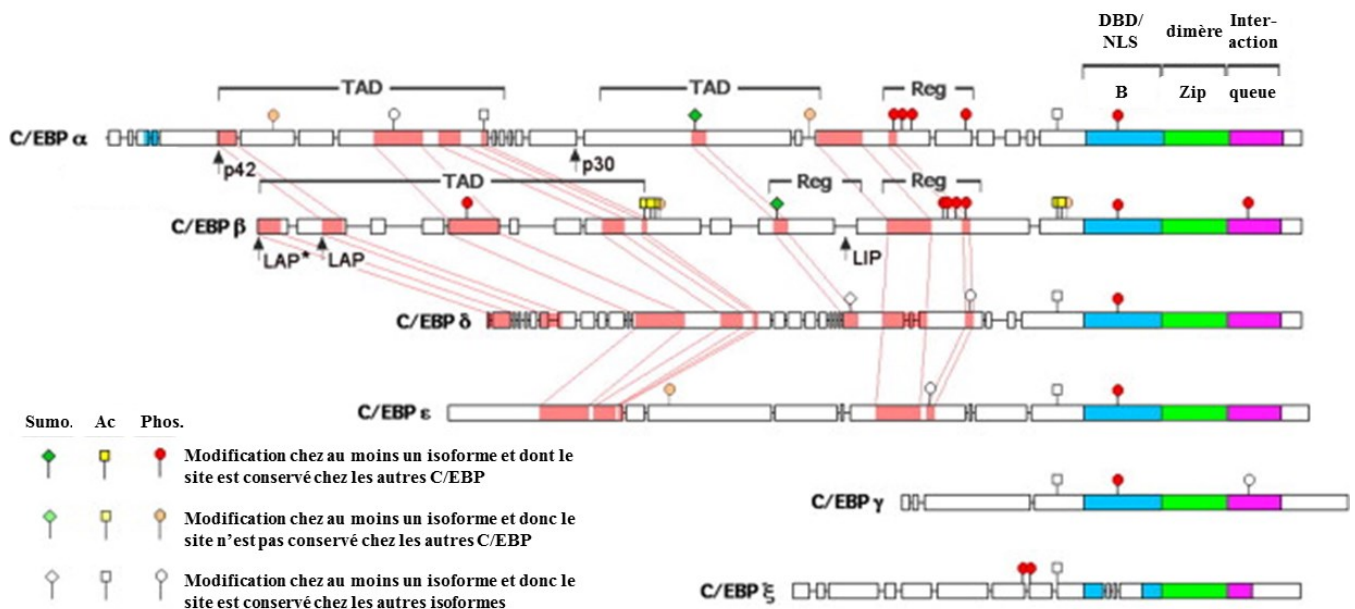
contribuer à la progression du cancer du sein ou de la prostate (610, 611). Dans les cas de CC, les niveaux de YY1 sont plus élevés que dans les tissus normaux (612, 613). Plusieurs variantes de VPH contenant des polymorphismes ou des délétions dans des sites YY1 ont été identifiées (603, 605, 614, 615). La mutation de ces sites résulte en l'augmentation de l'expression des oncogènes E6 et E7 et de la réplication de l'ADN (603, 605, 614, 615). Ainsi, la dérégulation de la répression induite par YY1 serait associée à un plus grand risque de développer le CC. Finalement, il semblerait qu'YY1 inhibe la réplication de l'ADN viral en interagissant avec la protéine virale E2 dans un système acellulaire de réplication d'un plasmide contenant l'origine de réplication du VPH18 (616).

#### 1.4.5.5 C/EBP $\beta$

Le facteur de transcription C/EBP $\beta$  (*CCAAT/Enhancer-binding Protein  $\beta$* ) ou *nuclear factor of interleukin 6* (NF-IL6) qui appartient à la famille des protéines contenant une glissière de leucine régule une multitude de gènes impliqués dans l'immunité, la réponse inflammatoire et la différenciation cellulaire notamment (revu dans (617) et (618)). L'expression de C/EBP $\beta$  est régulée par la voie de signalisation dépendante des AMPc via des sites Cre (élément de réponse à l'AMPc) situés dans le promoteur (619-622). Fait intéressant, un site C/EBP $\beta$  est également retrouvé dans le promoteur du gène de CEBPB, indiquant que ce FT peut autoréguler son expression (623, 624). C/EBP $\beta$  reconnaît préférentiellement la séquence consensus 5'-ATTGCGCAAT-3' (617, 618). Les nucléotides aux positions +/-3 et +/-4 suivi de la position +/-2 dans un demi-site (ATTGC) sont les plus importants pour la liaison de C/EBP $\beta$ , indiquant qu'une certaine dégénérescence des autres positions est permise (618, 625). C/EBP $\beta$  peut interagir avec d'autres membres de sa famille (C/EBP $\alpha$  par exemple) et de la famille des bZIP (ATF et CREB) ainsi qu'avec d'autres éléments régulateurs tels que NF- $\kappa$ B, GR, c-Myb, CBP/p300 et HDAC1 (revu dans (618),(617) et (626)).

Trois isoformes de C/EBP $\beta$  peuvent être obtenus par l'utilisation de sites initiateurs de la traduction alternatifs (627). Deux de ces isoformes, LAP\* (38 kDa) et LAP (35 kDa) ou communément appelé LAP pour *liver-enriched activator proteins*, sont des activateurs de la transcription. Le troisième isoforme LIP (20kDa) pour *liver-enriched inhibitory protein* est un répresseur de l'expression des gènes. Les isoformes LAP et LIP contiennent

un DBD et un domaine de dimérisation (bZIP) en C-terminal. La différence majeure entre ces deux isoformes est que LIP est dépourvu du TAD et de certaines séquences régulatrices présentes en N-terminal de LAP (voir Fig. 1.12). Bien que la présence du DBD soit commune aux deux isoformes, l'affinité de LIP pour certaines séquences d'ADN est meilleure (627). Finalement, la phosphorylation de certains résidus de C/EBP $\beta$  par ERK2 et PKC altère son affinité de liaison à l'ADN et module l'activité de transactivation (618, 625).



**Figure 1.12 Représentation schématique des différents membres de la famille C/EBP.** Les 6 membres de la famille C/EBP ( $\alpha$ ,  $\beta$ ,  $\delta$ ,  $\epsilon$ ,  $\gamma$  et  $\zeta$ ) sont représentés selon l'homologie du domaine de transactivation (TAD), des régions régulatrices (Reg), de la région basique (B) et du domaine de dimérisation de type glissière de leucine (Zip). Les modifications post-traductionnelles sont également indiquées (Sumo, Ac et Phos. étant sumoylation, acétylation et phosphorylation, respectivement). Les flèches représentent les sites d'initiation de la traduction alternatifs donnant lieu à la production de différents isoformes. D'après Tsukada et al. (2011) (618).

Les fonctions activatrices et inhibitrices de C/EBP $\beta$  sont déterminées en partie par l'expression relative des différents isoformes (ratio LAP/LIP) (563, 628). Alors que l'isoforme LAP est habituellement le mieux exprimé dans plusieurs types cellulaires (626), la surexpression de LIP par rapport à LAP favoriserait la tumorigenèse (629). Effectivement, l'expression de LIP est élevée dans les cancers du sein, de l'endomètre, des ovaires et colorectal (630-634). L'utilisation d'inhibiteurs de mTOR contribuerait à réduire les niveaux de LIP et à rétablir la balance LAP/LIP (629).

Le facteur C/EBP $\beta$  a été originellement identifié comme un répresseur de l'expression des gènes du VPH11 et 16 (584, 635-638). Certains sites C/EBP $\beta$  superposent des éléments pour AP1 et NF1, suggérant que C/EBP $\beta$  induit la répression par la compétition pour la liaison à l'ADN. Néanmoins, plusieurs études montrent que C/EBP $\beta$  pourrait avoir une double fonction régulatrice au sein de la même séquence promotrice dépendant du contexte cellulaire (562, 609). Ainsi, la surexpression de C/EBP $\beta$  résulte en la transactivation du LCR du VPH11 dans les C33A, mais en la répression dans les kératinocytes primaires (562, 637). Cette répression serait dépendante d'un site pour YY1 à proximité du promoteur. Dans un système de culture organotypique en radeau de cellules épithéliales, la mutation des sites C/EBP dans le LCR du VPH11 résulterait en l'augmentation de l'expression des gènes viraux dans des cellules prolifératives (639). Dans le cas du VPH18, son activité transcriptionnelle élevée dans les cellules HeLa dépend d'une région dite de transition (*switch region*) où la liaison d'un complexe constitué de YY1 et C/EBP $\beta$  est observée (602, 609). Dans les cellules HepG2, chez lesquelles le LCR du VPH18 n'est pas actif, le complexe YY1-C/EBP $\beta$  n'est pas détecté sur l'ADN. La surexpression de C/EBP $\beta$  induit la dissociation du complexe dans les cellules HeLa, causant la répression du LCR du VPH18, mais active le LCR dans les cellules HepG2 suite à la formation du complexe (606). Ensemble, ces résultats suggèrent que les partenaires d'interaction de C/EBP $\beta$  peuvent également moduler les fonctions de ce dernier.

Dans des cellules non-différenciées, LIP est la forme la mieux exprimée, ce qui corrèle avec l'état de répression observé dans les cellules non-différenciée chez la plupart des VPHs étudiés (628). Lors de la différenciation cellulaire, les niveaux de LAP restent stables, mais ceux de LIP diminuent significativement, ce qui permettrait l'activation du promoteur tardif. Une étude récente suggère que l'expression des oncogènes E6 et E7 du

VPH16 peut altérer l'activité transcriptionnelle de C/EBP et pourrait donc moduler l'expression des gènes-cibles de ce facteur (640). Chez les VPHs cutanés, l'oncoprotéine E7 ciblerait C/EBP $\beta$  et inhiberait sa capacité à induire l'expression de CCL20, un chimio-attractant important pour la migration des cellules de Langerhans dans l'épithélium (641). Les niveaux de CCL20 sont appauvris chez les individus atteints d'EV, indiquant que C/EBP $\beta$  serait impliqué dans la carcinogénèse des VPHs cutanés.

#### **1.4.6 Autres facteurs de régulation de l'expression des gènes**

Plusieurs VPHs contiennent des éléments de réponse aux glucocorticoïdes (GRE). La fonction des GRE est encore méconnue. Chez le VPH18, la mutation d'un élément GRE situé à proximité du promoteur E6/E7 induit l'activité transcriptionnelle (587, 642). Toutefois, la mutation du motif GRE n'a aucun effet sur l'activité transcriptionnelle du VPH31 dans un système de culture organotypique en radeau de cellules épithéliales (639).

Quelques études indiquent que les cytokines peuvent moduler l'expression des gènes des VPHs. Une première étude a montré que TNF- $\alpha$  et IL-1 sont des répresseurs de l'activité transcriptionnelle du LCR du VPH16, mais aucun effet n'a été observé pour IL-6 et IFN- $\gamma$  (643). L'expression des cytokines TGF- $\beta$ 1 et TGF- $\beta$ 2, IFN- $\alpha$  et IFN- $\gamma$  a un effet répresseur sur l'activité transcriptionnelle du LCR du VPH16 (644). Un criblage à haut débit basé sur la transfection stable de kératinocytes avec un rapporteur dont l'expression est sous le contrôle du LCR du VPH16 a permis d'identifier IL-4, IL-13, TGF- $\beta$ 1, TGF- $\beta$ 2, TGF- $\beta$ 3, TNF- $\alpha$ , IFN- $\alpha$  et IFN- $\beta$  comme des inhibiteurs de l'expression génique (645). Chez les VPHs cutanés de type 5 et 8, il a été observé que TGF- $\beta$ 1, TGF- $\beta$ 2 et TGF- $\beta$ 3 inhibaient l'activité transcriptionnelle chez le premier type seulement alors que IFN- $\alpha$  et IFN- $\beta$  inhibent la transcription chez les deux types (554). Ainsi, la production de cytokines contribuerait à la défense de l'hôte contre les infections par le VPH en diminuant l'expression des oncogènes E6 et E7. Toutefois, les VPHs ont développé plusieurs mécanismes pour prévenir l'activation du système immunitaire de l'hôte (discuté à la section 1.1.6).

Finalement, la méthylation des CpG dans les sites de liaison pour E2 du LCR est fréquemment observée (646). L'hyperméthylation du LCR semble corrélée avec une

expression accrue des oncogènes (171, 647-651), ce qui pourrait expliquer le fait que la méthylation du génome viral est associée à un risque plus élevé de développer un CC (648).

#### **1.4.7 Promoteur tardif**

La plupart des transcrits tardifs sont produits à partir du promoteur situé dans le cadre de lecture de E7 (p742 chez VPH31 et p670 chez le VPH16) et sont exprimés lors de la différenciation cellulaire (652-655). Ce promoteur génère plusieurs transcrits à partir de plusieurs sites d'initiation de la transcription et par l'épissage de E1 (177, 656-658). Des fonctions dans la régulation de la transcription tardive sont attribuables à C/EBP $\beta$ , Oct-1 et YY1 (176, 628, 659). Le promoteur tardif est également réprimé par CDP chez le VPH6 et 16 (515, 660). Des éléments régulateurs contenus dans la région amplificatrice du LCR contribuent aussi à la régulation du promoteur tardif, alors que leurs délétions réduisent l'activité du promoteur tardif (176). Tel que mentionné précédemment, l'expression de E4 précède celles des gènes tardifs, indiquant un rôle de E4 dans les phases tardives du cycle viral (161). Effectivement, la protéine virale E1<sup>E4</sup> contribuerait à activer la transcription tardive et la réplication du génome du VPH31 dans des cellules différenciées (661). La protéine virale E6 peut également induire l'activation du promoteur tardif du VPH8 en interagissant avec CBP/p300 (662).

#### **1.4.8 Transition de l'expression des gènes viraux précoces et tardifs**

Le promoteur précoce viral est plus actif dans les cellules prolifératives alors que le promoteur tardif est habituellement réprimé. Lors de la différenciation cellulaire, une transition survient dans la transcription induisant donc une augmentation de la production des transcrits à partir du promoteur tardif et une répression de l'expression des oncogènes E6 et E7 (652, 656, 657). Le mécanisme exact régissant la transition entre l'expression des gènes précoces et tardifs est encore inconnu, mais serait probablement le résultat du recrutement différentiel des FT durant les différentes étapes du cycle viral. L'utilisation d'un système de culture organotypique en radeau de cellules épithéliales permet d'étudier non seulement les fonctions du promoteur précoce et tardif, mais également la transition précoce-tardive de l'expression des gènes viraux.



Lors de la différenciation des kératinocytes primaires, une réorganisation de la composition des FT sur le LCR est observée (663). Les résultats les plus frappants concernent C/EBP $\beta$ . Ainsi dans les kératinocytes non-différenciés, C/EBP $\beta$  est fortement associé au LCR du VPH31, indiquant que ce facteur régule l'expression génique précoce du VPH31 (663). Dans une autre étude, C/EBP $\beta$  a été davantage retrouvé sur le promoteur précoce lors de la différenciation alors que les niveaux sur le promoteur tardif sont inchangés (176). Ceci indiquerait que ce facteur aurait une double fonction dans les cellules différenciées : réprimer le promoteur précoce et activer la transcription tardive. Dans une autre étude, utilisant cette fois-ci des micropuces d'ADN et de protéines, une trentaine de facteurs qui lient le LCR du VPH16 dans des cellules W12 ont été identifiés (664). Parmi ces facteurs, les niveaux d'AP1, NF1 et YY1, c-Myb et NFAT sur le LCR ont fluctué selon l'état de différenciation cellulaire, suggérant qu'ils contribuent à la régulation de la transition précoce-tardive de l'expression génique du VPH16.

Finalement, l'amplification virale étant également couplée à la différenciation cellulaire, l'augmentation du nombre de copies du génome pourrait également contribuer à l'enrichissement des transcrits tardifs (272, 654, 655, 665). Des éléments régulateurs d'ARN présents dans certaines régions codantes du génome des VPHs joueraient un rôle dans la régulation post-transcriptionnelle de la transition précoce-tardive de l'expression des gènes viraux en contrôlant l'épissage alternatif des ARNm (revu dans (666)). Ces éléments sont nécessaires pour la production de transcrits précoces et pour l'inhibition de la transcription tardive dans les cellules non-différenciées (508, 667-670).

#### **1.4.9 Étude de la liaison des facteurs de transcription**

L'analyse par empreinte de la déoxyribonucléase I (ADNase I), les essais de retard sur gel et l'immunoprécipitation de la chromatine sont des techniques classiques pour étudier la liaison de FT sur l'ADN des PVs (176, 409, 514, 592, 602, 635, 636, 664, 671). L'analyse computationnelle des séquences régulatrices du VPH permet de prédire assez fidèlement les sites de liaison des FT grâce à l'utilisation de bases de données telles que Transfac® (672) et JASPAR (673). La comparaison simultanée de plusieurs séquences permet d'identifier quels sont les sites les mieux conservés et qui contribueraient à la régulation générale de l'expression des gènes viraux. Ainsi, la conservation des sites pour

les facteurs AP1, Oct-1, NF1, Sp1 et YY1 chez plusieurs VPHs muqueux indique une fonction préservée dans la modulation de la transcription virale (409). L'identification d'un motif conservé chez plusieurs VPHs pourrait fournir un point de départ pour l'étude de l'interaction avec des nouvelles protéines.

Les prédictions de sites pour un FT donné sont habituellement basées sur l'utilisation d'une matrice de pondération des positions (ou *position weight matrix*) où la variabilité pour une position dans la séquence nucléotidique est caractérisée. Les séquences logos pour les TF peuvent être obtenues à partir des larges banques de données telles que Transfac et JASPAR (672, 673) et dont les interactions TF-ADN, qui ont servi pour ces analyses, ont été validés notamment par des études de ChIP-seq et DNase-seq. Ce type de matrice permet d'identifier les séquences les plus fréquemment retrouvées. Cette matrice part du principe que chaque position contribue indépendamment à la liaison globale d'un TF. Toutefois, certaines études montrent qu'il existe une certaine interdépendance des positions i.e. que l'effet d'une base sur l'affinité serait influencé par l'identité d'un nucléotide ailleurs dans la séquence (674, 675). La contribution d'une base donnée dans la liaison totale peut être facilement quantifiée dans des études de polarisation de la fluorescence (voir chapitre 4 de cette thèse). Les changements d'énergie pour chaque substitution peuvent par la suite être utilisés pour la construction d'une matrice énergétique qui, cette fois-ci, permettra d'identifier les sites avec une haute affinité pour un TF donné.

#### **1.4.10 Étude des fonctions des facteurs de transcription**

La présence d'un motif de liaison pour un facteur de transcription spécifique n'indique pas systématiquement que ce facteur joue un rôle dans la régulation de l'expression génique. Il est donc nécessaire de coupler les études d'identification des sites de liaison de facteurs cellulaire avec des essais fonctionnels.

##### **1.4.10.1 Essai de rapporteurs de gènes**

L'étude des effets fonctionnels des différents FT sur l'expression génique se base sur l'utilisation de cultures cellulaires primaires ou de lignées immortalisées. Pour étudier

l'activité transcriptionnelle basale du LCR, plusieurs utilisent des essais de rapporteur de gène luciférase ou de la chloramphénicol acétyltransférase couplés à la mutagenèse des éléments régulateurs (507, 545, 550, 554, 569, 593, 665, 676-678). Les effets de l'expression exogène des FT (515, 545, 679), de la progestérone et de l'estrogène (680), des corticostéroïdes (561) et des cytokines (554, 645) sur l'activité du LCR peuvent également être étudiés (515, 679).

Nous avons conçu un essai pour l'étude de l'activité transcriptionnelle du LCR du VPH33 dans les cellules C33A, HeLa et U2OS. Cet essai est basé sur la transfection d'une quantité croissante d'un plasmide exprimant la Renilla luciférase (Rluc) sous le contrôle du LCR du VPH33 et d'une quantité constante d'un autre plasmide exprimant cette fois-ci la *firefly* luciférase (Fluc), toujours sous le contrôle du LCR, qui sert comme contrôle interne pour l'efficacité de transfection (voir chapitres 2,3 et 4 pour les protocoles complets). Ainsi, l'augmentation du signal Rluc corrèle avec l'activité du LCR. Avec le signal de la Fluc, il est possible de normaliser les valeurs en calculant le ratio Rluc/Fluc, ce qui permet de comparer facilement l'activité transcriptionnelle d'un type avec un autre et ce, dans plusieurs types cellulaires. Notre essai est très sensible, ce qui est particulièrement utile pour l'étude des effets fonctionnels des polymorphismes dans le LCR dont certains ont un effet assez modeste sur l'expression des gènes viraux (voir chapitre 2 de cette thèse). Le désavantage principal est que seul l'activité du promoteur précoce à l'état d'équilibre (*steady-state*) peut être mesurée dans notre essai, ce qui ne permet pas de reconstituer le cycle viral complet. Par contre, notre système serait facilement adaptable pour l'étude de la transition de l'expression génique virale dans un système organotypique en radeau de cellules épithéliales.

#### **1.4.10.2 Système de culture organotypique en radeau de cellules épithéliales**

Le développement de culture organotypique en radeau de cellules épithéliales permet de récapituler assez bien le cycle viral lié à la différenciation cellulaire (681). Cette technique est utilisée dans l'étude de l'amplification virale, pour l'analyse des fonctions des FT dans la modulation des promoteurs viraux, mais aussi dans la transition précoce/tardive de l'expression des gènes viraux et ce, selon l'état de différenciation cellulaire (527, 654, 655). Brièvement, les kératinocytes sont cultivés sur une matrice de collagène et de

fibroblastes qui servent de support dans un milieu permissif à la prolifération. Les cellules sont ensuite placées dans une interface air-liquide. Les kératinocytes se différencient en stratification de manière similaire à un véritable épithélium (271, 682). Plusieurs lignées cellulaires immortalisées peuvent être utilisées notamment celles dérivées de lésion intraépithéliale, les cellules W12 et CIN612, qui contiennent l'épisode du VPH16 et VPH31, respectivement (277, 655, 657, 683-685). Des cellules primaires préalablement transfectées avec un plasmide portant le génome du VPH à l'étude ou infectées par un rétrovirus recombinant peuvent être utilisées, mais du fait de l'inconsistance du bagage génétique, plusieurs variations dans les résultats sont observées chez les cellules de différents donneurs (605, 686-690). Les kératinocytes peuvent également être cultivés en une monocouche ou dans un milieu de suspension et leur différenciation induite par l'utilisation d'un milieu de culture riche en calcium ou l'ajout de méthylcellulose, respectivement (561, 664, 665, 691, 692). Les cellules NIKS, une lignée de kératinocytes spontanément immortalisées sont souvent utilisées pour l'étude de l'amplification virale et de la transcription tardive chez ce type (266, 274).

## **1.5 Variantes du virus du papillome humain**

Les VPHs ont co-évolué avec les différentes populations humaines (693). Les VPHs accumulent des mutations dans leur génome de manière similaire à leurs hôtes puisque la réplication de l'ADN viral dépend de la machinerie répliquative cellulaire. La recombinaison entre différents virus serait possible quoique extrêmement rare (694). Néanmoins, l'accumulation de polymorphismes dans le temps serait à la base de la grande diversité des PVs.

### **1.5.1 Caractérisation des variantes**

La classification conventionnelle des différents isolats de VPH est basée sur l'homologie du cadre de lecture de L1, un des gènes les mieux conservés des PVs (voir section 1.1.1 de cette thèse). La plupart des études sur la variabilité intra-typique des VPHs se basent sur la comparaison des séquences d'ADN du LCR et/ou des cadres de lecture de

E6 et E7 (72, 695-701). Les variantes du VPH16 et VPH18 ont été originellement classifiées en des termes «géo-ethniques» selon leurs appartenances à certaines populations-cibles et régions (par exemple la variante européenne, africaine et asiatique) (699, 702). Toutefois, l'étude des variantes autres que celles des types 16 et 18 montre l'absence d'une distribution géographique préférentielle pour les variantes des types 31 et 33 (5, 703). Ceci, combiné à l'identification continue de nombreuses variantes dans différentes cohortes, a incité les taxonomistes à réviser la classification des variantes. Le groupe de Robert D. Burk a été le premier à proposer une classification des variantes en des termes alphanumériques qui se base sur l'homologie de la séquence nucléotidique de génomes entiers (702).

Les variations qui définissent une lignée sont retrouvées chez toutes les populations étudiées. Par contre, la présence de polymorphismes uniques dans chaque étude est un indicatif de l'évolution constante des PVs selon un processus de fixation de la lignée (*lineage fixation*) où un groupe de polymorphismes se fixe et devient une caractéristique propre à une lignée donnée (703, 704). Par exemple, les variantes du VPH33 sont classées dans 3 lignées majeures: A (divisée en 3 sous-lignées A1, A2 et A3), B et C. Cette dernière est rare et est retrouvée seulement en Afrique (5). Nous avons récemment associés C7732G, une caractéristique de la sous-lignée A2, à un plus grand risque de développer un HSIL (72). La caractérisation des variations propres à la sous-lignée A2 est présentée dans le deuxième chapitre de cette thèse. Brièvement, nous montrons que C7732G par elle-même augmente l'activité transcriptionnelle du LCR du VPH33 de 25% seulement ((705). Toutefois, la combinaison de C7732G avec une délétion de 79-pb et deux variations spécifiques à A2 contribue à augmenter l'activité du LCR des variantes A2 de deux fois celle du prototype de la sous-lignée A1 (705).

### **1.5.2 Pathogénicité des variantes**

Les VPHs sont associés à des potentiels oncogéniques différents; le type 16 étant le plus prévalent dans le CC (706). De plus en plus d'études montrent que quelques variantes du VPH16 sont étroitement associées à un risque plus élevé de progression vers le cancer invasif. Il s'agit des variantes non-européennes classifiées dans les lignées B, C et D, originellement les variantes Af-1, Af-2 et asiatique-américaine (AA), respectivement.

L'association d'une lignée de variantes avec un plus grand risque de développer un CC pour les autres types de VPH est moins documentée. Du fait que plusieurs groupes utilisent des études de cas/contrôles et se basent sur un nombre faible de participantes, il n'existe donc pas de consensus sur l'apport des variantes des VPHs de type 18, 31, 33, 52 et 58 au processus de carcinogénèse. Toutefois, il semblerait qu'en général l'infection aux variantes non-prototypiques des types 18, 31, 33, 35, 52 et 58 est associé avec un plus grand risque de persistance de l'infection et/ou de développer des lésions de haut-grade (CIN2/3 ou HSIL selon l'étude), deux facteurs de risque importants pour la progression du CC (72, 698, 700, 701, 707-710). D'autres études sur les types moins prévalents sont nécessaires pour mieux déterminer l'implication des variations génomiques sur la pathogénicité des VPHs. Les tableaux 1.1 à 1.4 récapitulent les conclusions générales des études sur les variantes des types dont l'apport dans le développement de CC est significatif.

### **1.5.3 Implications fonctionnelles des variations**

Tel que discuté dans une section antérieure, les variations dans le génome viral peuvent augmenter le potentiel oncogénique des variantes en altérant les niveaux d'expression des gènes viraux et de la réplication virale (696, 711-715), en potentialisant les activités d'immortalisation et de transformation des oncoprotéines E6 et E7 (716-718) et possiblement en altérant l'immunogénicité des épitopes de L1 (719). Il semblerait également que certains sites pour des microARNs soient créés dans la région non-codante entre les cadres de lecture de E5 et L2 et qui pourraient altérer l'expression de L2 (720).

Une association existe entre les variantes oncogéniques et une augmentation de ces activités, en particulier de l'activité transcriptionnelle des variantes non-Européennes du VPH16 (58, 711, 714). Ainsi, les variations nucléotidiques peuvent affecter l'histoire naturelle de l'infection au VPH, rendant certains isolats plus oncogéniques que d'autres.

Les articles présentés dans cette thèse indiquent que ceci est possiblement le cas pour les variantes du VPH33 de la sous-lignée A2 et pour la variante rare LCR10 de la lignée B. Nos résultats indiquent que la contribution de quelques variations est suffisante pour moduler significativement l'expression génique virale (705). De plus, l'altération de la séquence d'un site pour C/EBP $\beta$  module positivement la transcription virale du VPH33,

indiquant que C/EBP $\beta$  est un régulateur important de l'expression génique virale. Ainsi, les variations naturelles dans le LCR du VPH33 contribuent à augmenter le potentiel oncogénique des variantes.

**Tableau 1.1 Études sur les variantes du VPH16.**

Région séquencée	Association des variantes avec la dysplasie cervicale	Effets fonctionnelles	Références
<b>Génome complet</b>	Risque élevé de CIN2/3 et CC des variantes non-Européennes (OR, 4.6 [IC 95%, 1.1-20.2]; $p=.05$ )		Freitas et al. (2014) (58)
<b>LCR et E6</b>	Pas d'association significative avec le développement de CC		Marongiu et al. (2014) (695)
<b>E6</b>	Pas d'association significative avec le développement de CC		Qmichou et al. (2013) (721)
<b>Génome complet</b>	Présence de variations dans L2 est associée au CC ( $p=.0002$ )	Perte de sites de liaison pour certains miARNs entre les cadres de lecture de E5 et L2 chez les variantes E	Mandal et al. (2013) (720)
<b>LCR et E6</b>	Risque élevé de HSIL et CC des variantes As chez des Thaïlandaises (OR, 4.4 [IC 95%, 1.0-18.5]; $p=.042$ )	Activité transcriptionnelle accrue chez les variantes E, AS et AA dans des cellules C33A	Pientong et al. (2013) (696)
<b>LCR</b>	Pas d'association significative avec le développement de CC		Niccoli et al. (2012) (722)
<b>LCR et E6</b>	Association des variantes non-prototypiques (L83V) à un risque accru de HSIL ( $p=.04$ )	Activité transcriptionnelle accrue de la variante US56 dans les cellules C33A (2 fois plus que le PT)	Kammer et al. (2002) (697)
<b>E6</b>	ND	Activité de transformation accrue chez les variantes AA	Cornet et al. (2012) (716)
<b>E6</b>	Risque élevé de CIN2/3 des variantes AA chez des Costaricaines (OR, 1.7 [IC 95%, 1.0-3.2]; $p=.07$ )		Smith et al. (2011) (723)
<b>LCR</b>	ND	Activité transcriptionnelle réduite des variantes non-prototypique dans les cellules HeLa et HaCat (jusqu'à 3 fois moins que celle du PT)	Lei et al. (2011) (724)



**Tableau 1.1 (suite) Études sur les variantes du VPH16.**

Région séquencée	Association des variantes avec la dysplasie cervicale	Effets fonctionnelles des variations	Références
E6	Pas d'association significative avec le développement de CC		Sanchez et al. (2011) (725)
E6 et E7	ND	Activité de transformation et d'immortalisation de kératinocytes primaires accrue chez la variante AA de E7	Richard et al. (2010) (717)
E6	Association des variantes AA avec le développement d'adénocarcinome ( $p < 0.05$ )		Quint et al. (2010) (726)
E6	Risque élevé de HSIL et CC des variantes non-Européennes chez des Américaines (OR, 3.8 [IC 95%, 1.3-10.7])		Zuna et al. (2009) (727)
E6	Risque élevé de CC des variantes As (OR, 7.3 [IC 95%, 1.3-40.5]; $p = .022$ ) Risque moins élevé de CC des variantes E (OR, 0.137 [IC 95%, 0.025-0.76]; $p = .022$ )		Chopjitt et al. (2009) (728)
E6	Association des variantes L83V avec la persistance (OR, 3.0 [IC 95%, 1.4-6.7]) et la progression (OR, 6.2 [IC 95%, 2.7-14.3]) de l'infection		Grodzi et al. (2006) (729)
LCR		Activité répliquative accrue des variantes NA et AA (jusqu'à 20 fois par rapport au PT)	Hubert (2005) (714)
E6	Risque élevé de CC des variantes non-Européennes chez des Mexicaines (OR, 5.1 [IC 95%, 1.1-26.6])		Del Refugio Gonzalez-Losa et al. (2004) (730)
E6	Risque élevé de CC des variantes AA à chez des Mexicaines (OR, 27.0 [IC 95%, 6.4-113.7])		Berumen et al. (2001) (731)
E6	Association des variantes L83V à un risque accru de CC ( $p < 0.001$ )		Andersson et al. (2000) (732)

**Tableau 1.1 (suite) Études sur les variantes du VPH16.**

Région séquencée	Association des variantes avec la dysplasie cervicale	Effets fonctionnelles	Références
<b>E6</b>	Risque élevé de CIN2/3 des variantes non-Européennes (OR, 6.5 [IC 95%, 1.6-27.2])		Xi et al. (1997) (733)
<b>LCR et E6</b>	Risque élevé de CIN2/3 des variantes As chez des Chinoises (OR, 5.0 [IC 95%, 2.9-8.3]) Risque moins élevé de CIN2/3 des variantes E chez des Chinoises (OR, 0.2 [IC 95%, 0.1-0.3])		Sun et al. (2013) (734)
<b>LCR</b>	ND	Activité du LCR des variantes AA et NA est 2-3 fois plus élevée que la variante E dans les C33A	Kammer et al. (2000) (711)
<b>LCR et E2</b>	Risque élevé de HSIL et CC des variantes non-Européennes	Activité du LCR de la variante AA est 2 fois plus élevée que la variante E dans les C33A	Veress et al. (1999) (712)
<b>LCR</b>	Pas d'association significative avec le développement de CC		Junes-Gill et al. (2008) (735)
<b>LCR</b>	Risque élevé de CC des variantes non-Européennes chez des Taiwanaises (OR, 10.7 [IC 95%, 1.6-451.1]; $p=.0049$ )		Chang et al. (2013) (736)
<b>LCR et E6</b>	Risque élevé de CC des variantes Af chez des Portugaises ( $p=.006$ )		Pista et al. (2007) (737)
<b>E6</b>		Activité de dégradation de p53 est accrue chez les variantes de E6 du VPH16 dans un essai de dégradation de p53	Mesplède et al. (2012) (718)
<b>LCR et E6</b>	Risque élevé de CIN3 de la variante D (OR, 3.1 [IC 95%, 1.1-12.9]; $p=.04$ ) chez des Portugaises		Perez et al. (2014) (738)
<b>LCR, E6 et E7</b>	Risque élevé de CIN3 des variantes Af-2 (OR ajusté, 2.7 [IC 95%, 1.0-7.0]) et AA (OR, 3.1 [IC 95%, 1.6-6.0])		Xi et al. (2007) (698)

**Tableau 1.2 Études sur les variantes du VPH18.**

Région séquencée	Association des variantes avec la dysplasie cervicale	Effets fonctionnelles	Références
<b>LCR et E6</b>	Pas d'association significative avec le développement de CC		Chen et al. (2015) (699)
<b>LCR</b>	ND	Activité transcriptionnelle réduite de la variante Af Variation de l'activité transcriptionnelle de la variante AA varie selon le type cellulaire (plus active dans les C33A et SiHa que les HeLa)	Lopez-Saavedra et al. (2009) (739)
<b>E1</b>	ND	Activité répliquative accrue de la variante Af (2-3 fois plus élevée que la variante E)	Amador-Molina et al. (2013) (715)
<b>LCR</b>	ND	Activité transcriptionnelle accrue des variantes non-européennes (jusqu'à 8 fois plus pour variante AA comparée à la variante E)	Sichero, Franco et Villa (2005) (713)
<b>LCR, E6 et E7</b>	Risque élevé de CIN3 des variantes Af (OR ajusté, 3.8 [IC 95%, 0.9-17.2]) et E (OR ajusté, 4.8 [IC 95%, 1.0-23.6])		Xi et al. (2007) (698)
<b>LCR et E6</b>	Association de la variante E avec le développement de CC ( $p=0.039$ ) chez des Mexicaines		Lizano et al. (2006) (700)

**Tableau 1.3 Études sur les variantes des VPH31 et VPH33.**

Génome	Région séquencée	Association des variantes avec la dysplasie cervicale	Effets fonctionnelles	Références
HPV31	LCR	ND	Activité transcriptionnelle accrue des variantes B et C dans les cellules C33A (jusqu'à 12 fois plus que les variantes A)	Ferenczi et al. (2013) (740)
HPV31	LCR, E6 et E7	Pas d'association significative avec le développement de CC		Gagnon et al. (2005) (741)
HPV31	E6 et E7	Pas d'association significative avec le développement de CC		Chagas et al. (2013) (742)
HPV31	LCR, E6 et E7	Association des variantes B avec la persistance de l'infection chez des Africaines-Américaines (OR, 2.1 [IC 95%, 1.2-3.5])		Xi et al. (2013) (707)
HPV31	LCR, E6 et E7	Risque élevé de CIN2/3 des variantes A (OR, 1.7 [IC 95%, 1.0-2.9]) et B (OR, 2.2 [IC 95%, 1.2-3.9])		Xi et al. (2012) (708)
HPV31	E6 et E7	Pas d'association significative avec le développement de CC		Chagas et al. (2011) (743)
HPV31	LCR, E6 et E7	Risque élevé de CIN2/3 des variantes A (OR, 1.7 [IC 95%, 1.0-2.8]; $p=.02$ ) et B (OR, 2.2 [IC 95%, 1.3-4.1]; $p=.02$ )		Xi et al. (2014) (744)
HPV33	E6 et E7	Risque élevé de CC des variantes A en Afrique (OR, 9.4 [IC 95%, 1.9-51.4]) et en Europe (OR, 43.2 [IC 95%, 4.0-1945])		Chen et al. (2014) (745)
HPV33	LCR, E6 et E7	Risque élevé de HSIL des variantes non-prototypiques (C7732G) (OR, 9.2 [IC 95%, 1.8-45.9])		Khouadri et al. (2006) (72)
HPV33	LCR, E6 et E7	Association des variantes non-prototypiques (délétion de 79pb) avec la persistance de l'infection (OR, 1.8 [IC 95%, 1.2-2.7])		Gagnon et al. (2004) (701)

**Tableau 1.4 Études sur les variantes des VPH35, VPH52 et VPH58.**

<b>Génome</b>	<b>Région séquencée</b>	<b>Association des variantes avec la dysplasie cervicale</b>	<b>Effets fonctionnelles</b>	<b>Références</b>
<b>HPV35</b>	LCR, E6 et E7	Association des variantes non-prototypiques avec la persistance de l'infection (OR, 2.6 [IC 95%, 1.4-4.6])		Gagnon et al. (2004) (701)
<b>HPV52</b>	LCR, E6, E7 et E1	Pas d'association significative avec le développement de CC		Zhang et al. (2014) (746)
<b>HPV52</b>	LCR et E6	Risque élevé de CIN2/3 des variantes de la délétion 7695 (LCR) (OR, 4.9 [IC 95%, 1.2-20.8] et des variations T7744C (LCR) (OR, 5.7 [IC 95%, 1.1-32.0] et K93R (E6) (OR, 6.9 [IC 95%, 1.3-36.8])		Formentin et al. (2013) (709)
<b>HPV52</b>	LCR, E6, E7 et L1	Pas d'association significative avec le développement de CC		Sun et al. (2012) (747)
<b>HPV52</b>	LCR	Association des variantes non-prototypiques à la persistance de l'infection (OR, 14.1 [IC 95%, 1.1-200.0])		Aho et al. (2004) (748)
<b>HPV52</b>	LCR, E6 et E7	Risque élevé de CIN2/3 des variantes C tel que confirmé par cytologie (OR, 5.2 [IC 95%, 1.0-27.6]) et par histologie (OR, 7.6 [IC 95%, 1.3-43.8]) chez des Taisanaises		Chang et al. (2011) (749)
<b>HPV56</b>	LCR	Pas d'association significative avec le développement de CC		Rodrigues-Coutlée et al. (2013) (750)
<b>HPV58</b>	E6 et E7	Association de la variante E6/E7-HK-2 au développement de CIN2/3		Chan et al. (2002) (710)
<b>HPV58</b>	LCR, E6, E7, E5, E2 et L1	Pas d'association significative avec le développement de CC		Chan et al. (2011) (751)

## 1.6 Buts du projet de recherche

Nous avons précédemment investigué l'association entre des variations naturelles du LCR retrouvées chez des isolats du VPH33 et le grade de lésions intraépithéliales *in vivo* (72, 701). La présence d'une délétion de 79-pb et de C7732G ont été associées à la persistance de l'infection et au développement de HSIL, respectivement, deux facteurs prédictifs de la progression de l'infection au VPH vers le cancer. Tel que discuté précédemment, les variantes des VPH16 et 18 associées avec un potentiel oncogénique plus élevé présentent également une augmentation de leur activité de transcription (696, 711, 713, 714, 716, 717, 739), ce qui pourrait être le cas pour les variantes de LCR du VPH33. Le premier but de ce projet de recherche est d'analyser les effets fonctionnels des variations du LCR, en particulier pour la délétion de 79-pb et C7732G, sur l'activité transcriptionnelle du LCR afin de déterminer si leur association avec un risque plus élevé de progression de l'infection corrèle avec une activité transcriptionnelle accrue. Afin d'analyser l'activité de ces variantes, nous avons conçu un essai de transcription transitoire basé sur l'expression d'un gène rapporteur qui permettra de mesurer le niveau d'activité des LCRs des variantes du VPH33 suite à la transfection dans les cellules C33A et HeLa. Brièvement, nous montrons que trois variations propres à la sous-lignée A2 (C7537A, C7732G et A7879G) ainsi qu'une délétion de 79-pb contribuent à augmenter l'activité transcriptionnelle des variantes du groupe A2 par rapport au prototype. Les conclusions de notre étude sont présentées dans le deuxième chapitre de cette thèse.

La plupart des études sur la variabilité des VPHs se basent uniquement sur l'analyse de la séquence nucléotidique afin de déterminer quels sont les sites de facteurs de transcription altérés par les polymorphismes, mais peu entreprennent des études de validation de concept pour confirmer l'implication réelle d'un facteur cellulaire dans la régulation de l'expression des gènes viraux (679, 719, 734, 739, 752). Ainsi, le deuxième but de ce projet est de caractériser la variante rare du VPH33, LCR10, qui présente une activité transcriptionnelle élevée par rapport au prototype et d'identifier le ou les éléments régulateurs responsables de ce phénotype. Nous montrons que l'augmentation de l'activité du LCR10 est due à une variation unique T7791C qui altère la séquence d'un site répresseur C/EBP $\beta$ . Nous identifions un second site C/EBP $\beta$ , cette fois-ci, impliqué dans

l'activation de l'activité du LCR. Les conclusions de cette étude sont présentées dans le chapitre troisième de cette thèse.

Finalement, le troisième but de nos recherches est de déterminer la prévalence des sites C/EBP $\beta$  chez 175 types de VPH. Nos études de prédictions qui se basent sur l'utilisation d'une matrices des énergies de l'interaction entre C/EBP $\beta$  et l'ADN indiquent que les sites C/EBP $\beta$  à haute-affinité sont conservés chez les VPHs muqueux, en particulier deux groupes d'éléments qui seraient homologues aux deux sites répresseurs et activateurs retrouvés chez le VPH33. Les conclusions de cette dernière étude sont présentées dans le chapitre quatrième de cette thèse.

# **Chapitre deuxième : La combinaison de variations nucléotidiques spécifiques module l'activité transcriptionnelle des variantes de la longue région de contrôle du virus du papillome humain de type 33**

**Title:** The Combination of Specific Nucleotide Variations Modulates the Transcriptional Activity of the Human Papillomavirus Type 33 Long Control Region

**Short title:** Activity of HPV33 Variant LCRs

Jennifer Alvarez<sup>a</sup>, Francois Coutlée<sup>b</sup>, Jacques Archambault<sup>a\*</sup>.

<sup>a</sup> Institut de Recherches Cliniques de Montréal (IRCM) and department of Biochemistry and Molecular Medicine, Université de Montréal, Montréal, Canada <sup>b</sup> Centre de Recherche et Département de Microbiologie Médicale et Infectiologie, Centre Hospitalier de l'Université de Montréal (CHUM), Université de Montréal, Montréal, Canada

Abstract word count: 243

Article word count: 3418



## **2.1 Mise en contexte de l'article**

Il est connu que certaines variantes du VPH16 sont plus pathogènes du fait de leur activité de transcription élevée. Un de nos laboratoires a précédemment montré que certaines variantes du VPH33 contenant la variation C7732G et une délétion de 79 pb dans le LCR étaient associées à un risque plus élevé de développer un HSIL et de persistance de l'infection au VPH33, respectivement (72, 700). Ces variations dans le LCR peuvent affecter l'activité transcriptionnelle si elles sont situées dans des sites pour des FT. Le but de cette étude est d'investiguer l'effet de C7732G, de la délétion et d'autres polymorphismes dans le LCR sur l'expression génique virale. À noter que cet article sera soumis pour publication prochainement.

## **2.2 Contribution des auteurs**

J'ai réalisé toutes les expériences présentées dans ce manuscrit. J'ai également rédigé en grande partie cet article en collaboration avec Dr Jacques Archambault.

Dr François Coullée a participé dans la conception et la supervision de ce projet. Il a également contribué à la correction du manuscrit.

Dr Jacques Archambault a participé à la conception et à la supervision du projet. Il a également grandement contribué à la rédaction et à la correction de cet article.

## **2.3 Texte de l'article**

### **2.3.1 Abstract**

The human papillomavirus (HPV) Long Control Region (LCR) contains several binding-sites for cellular transcription factors that regulate expression of the viral genome. Few studies have been conducted on the activity of the LCR from HPV type 33. Here, we sequenced the LCRs of the prototype and of variants representative of the worldwide diversity of HPV33 and assessed their transcriptional activities in C33A and HeLa cervical

carcinoma cells using a luciferase reporter-gene assay. The strength of the LCR was found to vary according to the phylogenetic classification of HPV33, with the LCRs from A2-sublineage variants being the most active in C33A cells. Site-directed mutagenesis was used to establish that this enhanced activity arises from the combined effect of several weakly-acting variations, including a 79-bp deletion and a C7732G transversion previously associated with the persistence of the infection and the development of high-grade lesions, respectively, working in conjunction with two other polymorphisms, C7537A and A7879G. The stimulatory role of these variations was not as apparent in HeLa cells where the A7879G variation showed a strong repressive, rather than enhancing effect on transcription from the LCR. A fifth variation, A7874C, present in a single A1-sublineage variant was also identified that decreases the activity of the LCR in both C33A and HeLa cells. Collectively, these results indicate that the transcriptional activity of the HPV33 LCR is modulated by at least five distinct positively- and/or negatively-acting variations, whose combinatorial effect likely alter the oncogenic potential of A-lineage variants.

**Keywords.** HPV, LCR, variants, phylogeny, gene expression.

### **2.3.2 Introduction**

The persistence of high-risk human papillomavirus (HR-HPV) is an established risk factor for the development of cervical cancer (CC) [1, 2]. HR-HPV types, of which HPV16 is the most prevalent, favor cancer development by stimulating the proliferation of infected cells and preventing their apoptotic death, thus causing them to become immortal and genetically unstable. These effects result primarily from the ability of the HPV E6 and E7 oncoproteins to antagonize the p53 and pRb tumor suppressor pathways, respectively (reviewed in [3, 4]). Expression of these viral oncogenes is driven by the long control region (LCR), a regulatory sequence which contains binding sites for the viral E2 transcriptional repressor and for several cellular transcription factors (TF) such as AP-1, Sp1, YY1 and NF1 [5-7]. The LCR also contains binding sites for the viral E1 helicase which, together with E2, is necessary for replication of the viral episome. Overexpression of E6 and E7, a hallmark of HPV-associated cancer cells, often

occurs following integration of the viral genome into the host DNA, especially when the E2 open reading frame (ORF) is disrupted [5, 6].

The nucleotide sequence of the L1 ORF has been used to classify HPV isolates into types and variants. Viruses whose L1 sequences diverge by more than 2% are categorized as independent types, while those with  $\leq 2\%$  divergence are defined as variants. The LCR nucleotide sequence can differ by as much as 5% between variants as it is less evolutionarily constrained than coding genes [8]. Polymorphisms in the LCR and elsewhere in the viral genome can alter the behavior of individual HPV isolates. Studies on HPV16 and HPV18 identified variations in the LCR that increase E6 and E7 expression, as well as potentiate the ability of E6 to target p53 for proteasomal degradation [9-11].

Polymorphism in the LCR is commonly identified in cervical samples from women worldwide [12-15]. These variations could affect the association of TF with the LCR by creating, modifying or abolishing their binding sites, resulting in LCR variants with altered transcriptional activity. YY1 binding sites are often mutated or deleted in CC, leading to increased LCR activity, greater expression of the early viral genes and higher replication of the viral episome. [16, 17]. Similarly, the Asian-American and North-American HPV16 LCR variants exhibit higher transcriptional activity than the European prototype [18], which translates into enhanced viral DNA replication in cell culture models [19]. These findings provide a plausible rationale for the association of non-European HPV16 and HPV18 variants with a higher risk of developing high-grade squamous intraepithelial lesions (HSILs) and CC [20-22].

The less studied variants of HPV31 and HPV33 have recently become the focus of more extensive analyses indicating that they are not associated with geographical locations or ethnicity, unlike HPV16 and HPV18 [12-15, 23]. We have previously reported that HPV33 variants which lack one of the perfect 79-bp tandem repeats or which contain a C7732G transversion in their LCR, are associated with persistence of infection and progression to HSILs, respectively [13, 23], two well-established risk factors for CC development [24]. These variations were also found in HPV33 isolates described by other groups [12, 14, 15]. Chen et al. proposed to classify HPV33 variants into two main lineages, A and B, with the former being

divided into two sublineages, A1, which contains the prototype, and A2 [12, 25]. This classification highlights the C7732G variation as a hallmark of the A2-sublineage.

In this study, the impact of the C7732G substitution and/or of the 79-bp deletion, on the transcriptional activity of the HPV33 LCR was analyzed using a luciferase-reporter gene assay to provide evidence that these polymorphisms can alter viral gene expression. Our results indicate that the C7732G variation and the 79-bp deletion function together with other polymorphisms to alter the activity of the LCR. This was particularly manifest for the A2-sublineage variants which exhibited the highest transcriptional activity in C33A cells compared to the prototypical LCR (LCR-PT). Moreover, the introduction of these variations into the LCR-PT was sufficient to reproduce the activity level of the A2-sublineage LCR6 variant in the C33A and HeLa cells. Together, these results emphasize the need to account for the combinatorial effect of weakly-acting variations when studying the activity of the LCR. They also provide evidence that the phylogenetic classification of HPV33 variant reflects differences in LCR activity and raises the possibility that the A2-sublineage variants have a higher oncogenic potential due to increased oncogene expression.

### **2.3.3 Materials and Methods**

#### **DNA Sequence Analysis of HPV33 LCR Variants**

HPV33 LCRs were amplified from cervical samples and sequenced, each in two independent experiments (see [13, 23] for sample processing). All variations are reported according to the numbering scheme of the HPV33 reference genome (GenBank accession M12732.1).

#### **Phylogenetic Analysis**

LCR sequences were aligned using Clustal Omega [12, 14, 15] and used for the construction of a phylogenetic tree with EvolView [26].

## Plasmid Construction and Mutagenesis

The LCR-PT-Fluc and LCR-PT-Rluc plasmids encoding the prototypical HPV33 LCR (termed LCR-PT; nt 7094-7909 and 1-108 of the HPV33 genome) in front of the Firefly luciferase (Fluc) and Renilla luciferase (Rluc) coding region, respectively, were constructed by PCR amplification of LCR-PT from the reference genome using primers 33LCR-MluI-FWD (5'-GGGACGCGTACGCAAAAAGGTTAAAAAATAACACTTTGTG-3', MluI site underlined) and 33LCR-NcoI-RV (5'-GGCGTCTTCCATGGAGTCGTGCAGTACCTTACTGC-3', NcoI site underlined) and cloning of the resulting amplicon between the MluI and NcoI restriction sites of the promoterless pGL3-Basic (Promega) and pGL3-Rluc plasmids. pGL3-Rluc was generated by amplifying the Rluc gene from the pRL vector [27] using primers Rluc-NcoI-FWD (5'-GGGCCATGGATGACCAGCAAGGTG-3', McoI site underlined) and Rluc-XbaI-RV (5'-CCGCAGTCTAGATTACTGCTCGTTC-3', XbaI site underlined) and subcloning of this amplicon between the NcoI and XbaI restriction sites of pGL3-Basic such as to replace the Fluc ORF by that of Rluc. HPV33 LCR variants were amplified and cloned into pGL3-Rluc as described above for LCR-PT. When indicated, mutations in the LCR were introduced using the QuikChange Site-Directed Mutagenesis kit (Stratagene). All constructs were verified by DNA sequencing.

## Cell Culture and Transfection

The human cervical carcinoma C33A and human cervical adenocarcinoma HPV18-positive HeLa cell lines were grown in Dulbecco's Modification of Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 50 I.U. of penicillin/ml, 50 µg of streptomycin/ml and 2 mM of L-glutamine (Wisent Bioproducts). Transfections were performed using the Lipofectamine 2000 reagent according to the manufacturer's recommendations (cat. no. 11668-500, Life Technologies).

## Luciferase-reporter Gene Transcription Assay

C33A and HeLa cells were plated 24 h prior to transfection in white flat-bottom 96-well plates (cat. no. 3917, Corning) at a density of 25 000 and 15 000 cells/well, respectively. Cells were transfected with increasing amounts of each LCR-Rluc plasmid (25, 50, 100 and 200 ng) and a fixed amount of LCR-PT-Fluc (25 ng) as an internal control. The pGL3-Rluc plasmid (200 ng) was used as a negative control. Both Fluc and Rluc activities were measured with the Dual-Glo luciferase assay system and a GloMax™ 96-well luminometer (Promega) 24h post-transfection. For each condition, the activity of the LCR was determined from the Rluc/Fluc ratio and reported relative to the ratio obtained with 100 ng of LCR-PT, which was assigned a reference value of 1.0. Each LCR-Rluc plasmid was tested in duplicates in at least 5 independent experiments.

### **Statistical Analysis**

Statistical significance was tested by ANOVA with Dunnett's post-hoc analysis. Statistics and AUC calculations were performed using GraphPad Prism version 6.00.

### **Computational Analysis of Transcription Factor Binding Sites**

Putative TF binding sites in the HPV33 LCR were predicted with the PROMO software and version 8.3 of the TRANSFAC database [28, 29].

## **2.3.4 Results**

### **Nucleotide Sequence Diversity of HPV33 Long Control Region Variants**

We had previously sequenced a portion of the HPV33 LCR (nt 7337-7878) from cervical samples [13, 23]. The 79-bp deletion and C7732G variation (Fig. 1A) were associated with HPV33 persistence and HSILs, respectively. The complete LCR (nt 7094-108) from eight variants containing the 79-bp deletion and/or C7732G variation (variants 4, 5, 6, 7, 8, 9, 12 and 14 in [13, 23]) was sequenced. New and previously identified variations

were detected for a total of 26 polymorphisms when considering substitutions, insertions and deletions as individual variations (Table 2.1). LCR8 and LCR9 were the most polymorphic each with a total of 18 variations resulting in a 1.9% change in the LCR sequence. All LCRs contained the A81C transversion found in other HPV33 isolates; a C at this position is considered the correct nucleotide although it is replaced by an A in the reference sequence (GenBank: M12732.1).

We found that several of our variants were also isolated in Asian and African cohorts consistent with the absence of a preferred geographical distribution for HPV33 [12, 14, 15]. Namely, LCR5, LCR6, LCR7 and LCR8 correspond to INJP06456 (Thailand), INCC0137 (Thailand), IN241890 (Thailand) and RW136 (Rwanda), respectively. The cumulative number of variations found amongst HPV33 isolates indicate that up to 5.7% of the LCR sequence is subject to change. The phylogenetic tree presented in Fig. 2.1B shows that our LCR variants belong to the A- and B-lineages [12] and thus are representative of the worldwide HPV33 diversity. LCR12, which contains the intact 79-bp duplication and two variations including the previously unnoticed C7732T transition, was the closest to the A1-sublineage prototype [13, 23]. The LCR6 and LCR14 variants, which contain the 79-bp deletion and the C7732G variation, were grouped within the A2-sublineage, together with LCR5 which lacks the latter polymorphism. The remaining variants (LCR4, LCR7, LCR8 and LCR9) belong to the B-lineage and contain the 79-bp deletion with additional B-lineage variations.

### **Transcriptional Activity of HPV33 LCR Variants**

To determine if the transcriptional activity of the LCR varies within and/or between (sub)lineages, the transcriptional activities of LCR-PT and variants were measured in a luciferase-reporter transcription assay. Briefly, plasmids encoding Rluc under the control of LCR-PT and variant LCRs were constructed and tested for Rluc expression in C33A and HeLa cells. A plasmid expressing Fluc from LCR-PT was used as an internal control to avoid the potential titration of cellular TF when Fluc is expressed from a stronger promoter such as pCMV. Transcriptional activities were determined from the Rluc/Fluc ratios measured at increasing amounts of LCR-Rluc plasmid and are presented relative to the ratio

obtained with 100 ng of LCR-PT, which was assigned a value of 1.0. The dose-response curves for each LCR were used to calculate the area under the curve (AUC) as an overall measure of transcriptional activity.

Several observations were made from the results presented in Fig. 2.2. First, LCR variants that belong to the same (sub)lineage tend to have a similar transcriptional activity. For example, the A2-sublineage LCRs were the most active in C33A cells, with transcriptional levels 1.5 to 2-fold higher than those from the A1-sublineage and B-lineage. Second, the activity of some LCRs is noticeably different in C33A cells than in HeLa cells, as best exemplified by the A2-sublineage variants which were the strongest in C33A cells but similar to LCR-PT in HeLa cells. This likely reflects differences in the TF proteome of these cells and highlights the value of assessing LCR activity in different cell lines.

The effect of specific variations could be inferred by comparing the activities of LCRs which diverge only at a few nucleotide positions. For example, LCR12 differs from LCR-PT by the presence of the C7732T and A7874C variations, indicating that its weaker activity in C33A and HeLa cells must be caused by either one or a combination of these differences. Similarly, comparison of the three A2-sublineage variants suggests that the A7686C variation is silent and that the C7732G transversion increases the transcriptional activity of LCR6 and LCR14. The comparison of LCR7 with LCR4 in HeLa cells indicated that deletion of T7404 accounts for the lower activity of LCR4. Comparison of LCR7 with LCR9 and LCR8 showed that the deletion and insertion at nucleotides 7412 and 7529 have little to no effect on its activity. Collectively, these results indicate that changes within the HPV33 LCR can alter its transcriptional activity, often in a cell type-dependent manner. While these studies suggest that the C7732G variation previously associated with HSILs can indeed increase the transcriptional activity of the LCR [13, 23], they provided little information on the effect of the 79-bp deletion that is linked to the persistence of the infection, as none of the variants contained this polymorphism alone. This prompted us to analyze the effect of the 79-bp deletion by introducing it into LCR-PT.



## **The 79-bp Deletion Potentiates the Transcriptional Stimulatory Effect of the C7732G Variation**

To determine its impact on transcriptional activity, the 79-bp deletion was introduced either individually or in combination with C7732G in LCR-PT. The deletion did not alter the activity of LCR-PT significantly in either cell lines (Fig. 2.3A-B). However, it was able to potentiate the stimulatory effect caused by C7732G in C33A cells as the combination of both variations increased LCR activity by approximately 50% (Fig. 2.3C). The deletion of both 79-bp copies reduced the LCR activity by 20-30% in both cell lines (Fig. 2.3C). Alone, the C7732G variation increased LCR activity by approximately 25% in C33A cells with little to no effect in HeLa cells.

The stimulatory effect of C7732G prompted us to investigate if the related C7732T variation in LCR12 could also enhance transcription. Reversion of C7732T to the prototypical sequence (T7732C) was found to have little to no effect on the activity of LCR12 in C33A and in HeLa cells (Fig. 2.4A-B). In parallel, the introduction of T7732G in LCR12 increased its transcriptional activity by 25% in C33A cells (Fig. 2.4C). Thus, only a G at position 7732 increases the activity of the HPV33 LCR. Finally, LCR12 also contains the A7874C variation, which upon reversion, was found to reduce the activity of this LCR by 30% in C33A and by 80% in HeLa cells (Fig. 2.4A-B).

## **Additional Variations Contribute to the Transcriptional Activity of A2-sublineage LCRs**

The results presented above indicate that the A2-lineage specific variation C7732G is sufficient to increase the activity of LCR-PT by 25%, and up to 50% when combined with the 79-bp deletion. LCR6 and LCR14 from the A2-sublineage contain this combination but their activity is almost twice that of LCR-PT in C33A cells (Fig. 2.2A and 2.2C). The third member of the A2-sublineage, LCR5, lacks C7732G but is otherwise identical to LCR6. Accordingly, LCR5 is about 25% less active than LCR6 but is still significantly more potent than LCR-PT (Fig. 2.2 and confirmed in Fig. 2.5A and 2.5C). These observations suggested that additional variations contribute to the higher

transcriptional activity of A2-sublineage LCRs. To test this possibility, the four other A2-sublineage specific variations, C7443T, C7537A, A7879G and C6G were reverted individually to the prototypical sequence in the context of LCR6. The results presented in Fig. 2.5 show that reversion of C7443T and C6G had little to no significant effect whereas reversion of C7537A or A7879G resulted in a 15-25% drop in LCR6 activity, in C33A cells. Interestingly, the A7879G variation (Fig. 2.5B and 2.5C) was found to have a strong repressive effect in HeLa cells as its reversion almost doubled (+ 95%) the activity of LCR6, thus explaining why the three A2-sublineage LCRs are not stronger than LCR-PT in this cell line.

Finally, the introduction of the 79-bp deletion and of the A2-sublineage C7537A, C7732G and A7879G variations in the prototypical LCR resulted in a two-fold increased transcriptional activity in C33A cells but a 35% reduction in activity in HeLa cells (Fig. 2.6). The combination of these variations accounts for the higher transcriptional activity of the A2-sublineage variants in C33A cells but this stimulatory effect is largely masked by the repressive nature of the A7879G variation in HeLa cells. Together, these results led us to conclude that the transcriptional activity of LCR6 in C33A and in HeLa cells arises from the combination of four weakly-acting variations, namely, the 79-bp deletion, C7732G, C7537A and A7879G.

### **2.3.5 Discussion**

In this study, the complete LCRs from different HPV33 isolates found in cervical samples were sequenced and their transcriptional activity measured in C33A and HeLa cervical carcinoma cell lines using a reporter-gene assay. To our knowledge, this is the first study examining the structure and function of the HPV33 LCR. Phylogenetic analysis showed that our LCR variants span the three previously-described clades, the A1- and A2-sublineages and the B-lineage, indicating that they are representative of HPV33 diversity [12]. This classification was originally based on the comparison of complete viral genome sequences [25, 30]. However, our classification of HPV33 variants solely based on the LCR sequence was sufficient to assign each

isolate to its correct clade. Thus, the LCR may be representative of the global evolutionary divergence of each variant.

Several of the A1-sublineage and B-lineage LCR variants exhibited similar transcriptional activities as LCR-PT in C33A and in HeLa cells. The A2-sublineage LCRs exhibited transcriptional levels that were as much as 2-fold higher or lower than those of LCR-PT (e.g. LCR6 in C33A cells and LCR5 in HeLa cells; see Fig. 2.2)). Although a 2-fold increase or decrease in transcriptional activity may seem modest, we note that Kämmer et al. measured a similar difference in LCR activity, of 3-fold, between HPV16 Asian-American and North-American variants and the European reference in comparable reporter-gene assays [18]. Remarkably, this 3-fold increase in LCR activity resulted in greater replication of these variant genomes (20-fold higher copy number than the prototype) in immortalized keratinocytes [31]. Small differences in LCR activity can therefore profoundly influence replication of the viral episome, presumably by modulating transcription of the early viral genes, including those encoding the replication proteins E1 and E2. By analogy, we would expect that the enhanced LCR activity of A2-sublineage variants would stimulate their replication in cell culture models, although additional studies will be needed to ascertain this prediction.

A large part of this study was dedicated to dissecting the contribution of individual variations to the transcriptional activity of the HPV33 LCR. Our results indicate that the 2-fold higher activity of A2-sublineage LCR variants in C33A cells is caused by the combinatorial effect of the specific variations, C7537A, C7732G and G7879A and the loss of a 79-bp region (Fig. 2.5 and 2.6). The C7732G variation and 79-bp deletion are sufficient to stimulate the activity of the LCR by 50% (Fig. 2.3) while the C7537A and A7879G variations each confer an additional 25% increase in transcription in C33A cells (Fig. 2.4). The C7732G variation alone was sufficient to increase the activity of the HPV33 LCR by 25% in C33A cells (Fig. 2.3). Similar experiments indicated that the 79-bp deletion had no effect on its own but that it could potentiate the stimulatory effect of C7732G. It is possible that deletion of a single copy of the duplicated 79-bp region is not sufficiently deleterious to alter the activity of the LCR under our assay conditions, due to the redundant function of the remaining copy. Accordingly, we found that removal of both 79-bp copies

reduced LCR activity by 20-30%, suggesting an important contribution of this region in viral gene expression (Fig. 2.3). Collectively, these results provide a molecular explanation for the association of the C7732G variation and the 79-bp deletion with an increased risk of developing HSILs and with the persistence of the infection, respectively [13, 23].

Our results also raise the possibility that these variations act in conjunction with other co-segregating variations, such as C7537A and G7879A (see Fig. 2.4), which are specific to the A2-sublineage variants and modulate the strength of their LCR. The contribution of A7879G to the LCR activity is particularly interesting as it can be either stimulatory (+25% in C33A cells) or strongly repressive (- 95% in HeLa cells), depending on the cellular context (see Fig. 2.4). The A7874C variation, present only in the A1-sublineage LCR12, is also strongly repressive in HeLa cells only. It would be of interest to determine if this variation accounts for the increased risk of progression to HSILs or worse reported for the A1-sublineage variants [32, 33].

Variations that affect the HPV33 LCR activity may create, alter or abolish binding sites for cellular TF. The 79-bp region contains putative binding sites for the cellular TFs USF (upstream stimulatory factor), NF-Y and c-Myb while the A2-lineage specific variation C7732G could abolish a binding site for USF and/or c-Myc, members of the basic helix-loop-helix-leucine zipper family which regulates genes involved in proliferation, stress response and cell cycle progression [34-36]. The A2-sublineage variations, C7537A, may increase LCR activity by altering a binding site for USF. Interestingly, the other A-lineage variations, A7879G and A7874C, could abolish a putative non-canonical site for AP1 (TAGGTCA, nt 7873-7879) which is not conserved in other HPV types. In addition, we noted that many of the variations that affect the LCR activity differentially in C33A versus HeLa cells lie in putative binding sites for transcription factors that can function either as activators or repressors of transcription (e.g. USF and NF-Y), depending on the cellular context [34, 35, 37].

The differential effect of these variations may thus reflect the distinct TF proteomes and signalling pathways found in cervical carcinoma (C33A) and adenocarcinoma (HeLa) cells. In conclusion, this study indicates that HPV33 gene expression can be modulated by

several naturally-occurring variations that often act in combination to significantly alter the activity of the LCR. Such combination of polymorphisms may enhance the oncogenic potential of HPV33 variants, in particular those from the A2-sublineage.

### **Acknowledgements**

The authors would like to thank Eduardo Franco for his constant support and members of the Archambault laboratory for their helpful comments on this manuscript. This work was supported by grants from the Canadian Cancer Society Research Institute (CCSRI grant #020298), Canadian Institutes of Health Research (CIHR team grant CRN-83320) and FRQS Réseau SIDA MI.

### **Potential conflicts of interest**

The authors have no conflict of interest to declare relevant to the topic of this article.

### **2.3.6 Bibliography**

- [1] H. U. Bernard, R. D. Burk, Z. Chen, K. van Doorslaer, H. zur Hausen, and E. M. de Villiers, "Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments," *Virology*, vol. 401, pp. 70-9, May 25 2010.
- [2] F. X. Bosch, A. N. Burchell, M. Schiffman, A. R. Giuliano, S. de Sanjose, L. Bruni, G. Tortolero-Luna, S. K. Kjaer, and N. Munoz, "Epidemiology and natural history of human papillomavirus infections and type-specific implications in cervical neoplasia," *Vaccine*, vol. 26 Suppl 10, pp. K1-16, Aug 19 2008.
- [3] S. B. Vande Pol and A. J. Klingelutz, "Papillomavirus E6 oncoproteins," *Virology*, vol. 445, pp. 115-37, Oct 2013.
- [4] A. Roman and K. Munger, "The papillomavirus E7 proteins," *Virology*, vol. 445, pp. 138-68, Oct 2013.

- [5] H. U. Bernard, "Regulatory elements in the viral genome," *Virology*, vol. 445, pp. 197-204, Oct 2013.
- [6] P. Howley and D. Lowy, "Papillomaviruses," in *Fields Virology*. vol. Vol. 5, D. Knipe and P. Howley, Eds., ed Philadelphia, PA: Lippincott Williams & Wilkins, 2007, pp. 2299–2354.
- [7] M. J. O'Connor, S. Y. Chan, and H. U. Bernard, "Transcription factor binding sites in the long control regions of genital HPVs, part III-A," in *Human papillomaviruses 1995 compendium*, B. H. U. Myers G, Delius H, Baker C, Icenogle J, Halpern A, Wheeler C, Ed., ed Los Alamos, N.M. USA: Los Alamos National Laboratory, 1995, pp. pp. 21-40.
- [8] L. Ho, S. Y. Chan, R. D. Burk, B. C. Das, K. Fujinaga, J. P. Icenogle, T. Kahn, N. Kiviat, W. Lancaster, P. Mavromara-Nazos, and et al., "The genetic drift of human papillomavirus type 16 is a means of reconstructing prehistoric viral spread and the movement of ancient human populations," *J Virol*, vol. 67, pp. 6413-23, Nov 1993.
- [9] T. Mesplede, D. Gagnon, F. Bergeron-Labrecque, I. Azar, H. Senechal, F. Coutlee, and J. Archambault, "p53 degradation activity, expression, and subcellular localization of E6 proteins from 29 human papillomavirus genotypes," *J Virol*, vol. 86, pp. 94-107, Jan 2012.
- [10] J. W. Yi, M. Jang, S. J. Kim, S. S. Kim, and J. E. Rhee, "Degradation of p53 by natural variants of the E6 protein of human papillomavirus type 16," *Oncol Rep*, vol. 29, pp. 1617-22, Apr 2013.
- [11] M. C. Stoppler, K. Ching, H. Stoppler, K. Clancy, R. Schlegel, and J. Icenogle, "Natural variants of the human papillomavirus type 16 E6 protein differ in their abilities to alter keratinocyte differentiation and to induce p53 degradation," *J Virol*, vol. 70, pp. 6987-93, Oct 1996.
- [12] Z. Chen, M. Schiffman, R. Herrero, R. Desalle, K. Anastos, M. Segondy, V. V. Sahasrabudde, P. E. Gravitt, A. W. Hsing, and R. D. Burk, "Evolution and taxonomic classification of human papillomavirus 16 (HPV16)-related variant genomes: HPV31, HPV33, HPV35, HPV52, HPV58 and HPV67," *PLoS One*, vol. 6, p. e20183, 2011.
- [13] S. Gagnon, C. Hankins, C. Tremblay, P. Forest, K. Pourreaux, and F. Coutlee, "Viral polymorphism in human papillomavirus types 33 and 35 and persistent and

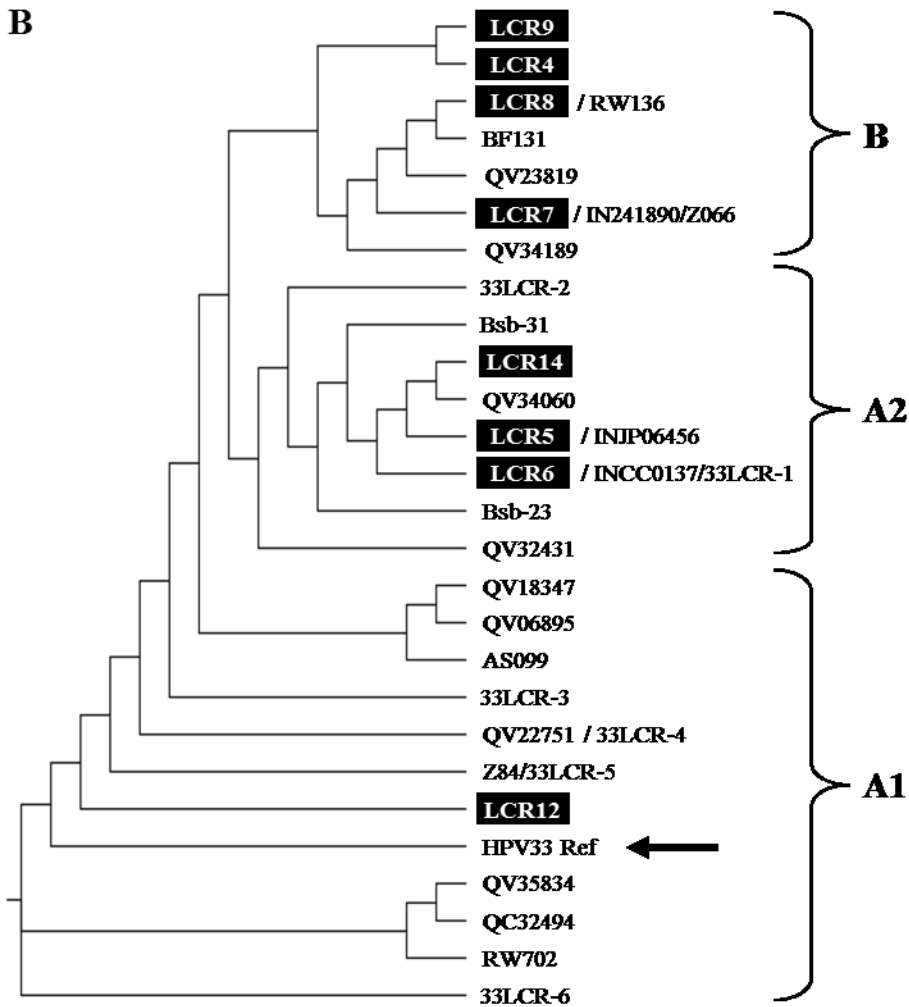
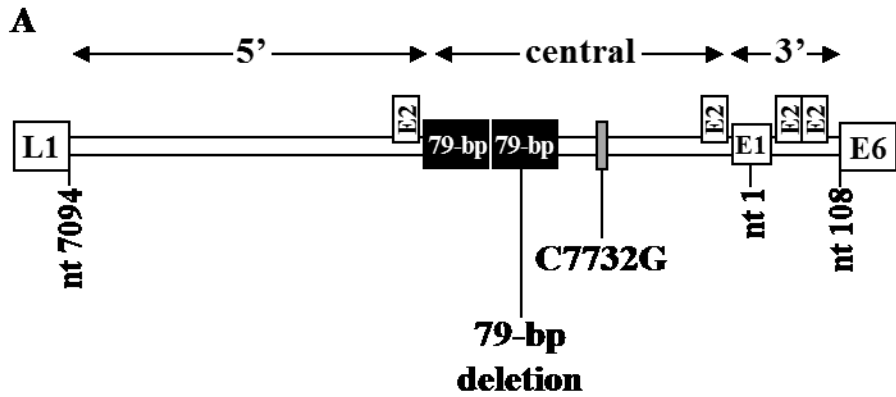
- transient infection in the genital tract of women," *J Infect Dis*, vol. 190, pp. 1575-85, Nov 1 2004.
- [14] E. Vrtacnik Bokal, B. J. Kocjan, M. Poljak, Z. Bogovac, and N. Jancar, "Genomic variants of human papillomavirus genotypes 16, 18, and 33 in women with cervical cancer in Slovenia," *J Obstet Gynaecol Res*, vol. 36, pp. 1204-13, Dec 2010.
- [15] T. Raiol, P. S. Wyant, R. M. de Amorim, D. M. Cerqueira, N. Milanezi, M. Brigido Mde, L. Sichero, and C. R. Martins, "Genetic variability and phylogeny of the high-risk HPV-31, -33, -35, -52, and -58 in central Brazil," *J Med Virol*, vol. 81, pp. 685-92, Apr 2009.
- [16] M. J. Lace, Y. Yamakawa, M. Ushikai, J. R. Anson, T. H. Haugen, and L. P. Turek, "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*, vol. 90, pp. 2402-12, Oct 2009.
- [17] M. May, X. P. Dong, E. Beyer-Finkler, F. Stubenrauch, P. G. Fuchs, and H. Pfister, "The E6/E7 promoter of extrachromosomal HPV16 DNA in cervical cancers escapes from cellular repression by mutation of target sequences for YY1," *Embo j*, vol. 13, pp. 1460-6, Mar 15 1994.
- [18] C. Kammer, U. Warthorst, N. Torrez-Martinez, C. M. Wheeler, and H. Pfister, "Sequence analysis of the long control region of human papillomavirus type 16 variants and functional consequences for P97 promoter activity," *J Gen Virol*, vol. 81, pp. 1975-81, Aug 2000.
- [19] W. G. Hubert, "Variant upstream regulatory region sequences differentially regulate human papillomavirus type 16 DNA replication throughout the viral life cycle," *J Virol*, vol. 79, pp. 5914-22, May 2005.
- [20] L. F. Xi, L. A. Koutsky, A. Hildesheim, D. A. Galloway, C. M. Wheeler, R. L. Winer, J. Ho, and N. B. Kiviat, "Risk for high-grade cervical intraepithelial neoplasia associated with variants of human papillomavirus types 16 and 18," *Cancer Epidemiol Biomarkers Prev*, vol. 16, pp. 4-10, Jan 2007.
- [21] L. Sichero, S. Ferreira, H. Trottier, E. Duarte-Franco, A. Ferenczy, E. L. Franco, and L. L. Villa, "High grade cervical lesions are caused preferentially by non-European variants of HPVs 16 and 18," *Int J Cancer*, vol. 120, pp. 1763-8, Apr 15 2007.

- [22] L. B. Freitas, Z. Chen, E. F. Muqui, N. A. Boldrini, A. E. Miranda, L. C. Spano, and R. D. Burk, "Human papillomavirus 16 non-European variants are preferentially associated with high-grade cervical lesions," *PLoS One*, vol. 9, p. e100746, 2014.
- [23] S. Khouadri, L. L. Villa, S. Gagnon, A. Koushik, H. Richardson, S. Ferreira, P. Tellier, J. Simao, G. Matlashewski, M. Roger, E. L. Franco, and F. Coutlee, "Human papillomavirus type 33 polymorphisms and high-grade squamous intraepithelial lesions of the uterine cervix," *J Infect Dis*, vol. 194, pp. 886-94, Oct 1 2006.
- [24] N. F. Schlecht, S. Kulaga, J. Robitaille, S. Ferreira, M. Santos, R. A. Miyamura, E. Duarte-Franco, T. E. Rohan, A. Ferenczy, L. L. Villa, and E. L. Franco, "Persistent human papillomavirus infection as a predictor of cervical intraepithelial neoplasia," *Jama*, vol. 286, pp. 3106-14, Dec 26 2001.
- [25] Z. Chen, M. Terai, L. Fu, R. Herrero, R. DeSalle, and R. D. Burk, "Diversifying selection in human papillomavirus type 16 lineages based on complete genome analyses," *J Virol*, vol. 79, pp. 7014-23, Jun 2005.
- [26] H. Zhang, S. Gao, M. J. Lercher, S. Hu, and W. H. Chen, "EvolView, an online tool for visualizing, annotating and managing phylogenetic trees," *Nucleic Acids Res*, vol. 40, pp. W569-72, Jul 2012.
- [27] A. Fradet-Turcotte, G. Morin, M. Lehoux, P. A. Bullock, and J. Archambault, "Development of quantitative and high-throughput assays of polyomavirus and papillomavirus DNA replication," *Virology*, vol. 399, pp. 65-76, Mar 30 2010.
- [28] X. Messeguer, R. Escudero, D. Farre, O. Nunez, J. Martinez, and M. M. Alba, "PROMO: detection of known transcription regulatory elements using species-tailored searches," *Bioinformatics*, vol. 18, pp. 333-4, Feb 2002.
- [29] D. Farre, R. Roset, M. Huerta, J. E. Adsuara, L. Rosello, M. M. Alba, and X. Messeguer, "Identification of patterns in biological sequences at the ALGEN server: PROMO and MALGEN," *Nucleic Acids Res*, vol. 31, pp. 3651-3, Jul 1 2003.
- [30] Z. Chen, R. DeSalle, M. Schiffman, R. Herrero, and R. D. Burk, "Evolutionary dynamics of variant genomes of human papillomavirus types 18, 45, and 97," *J Virol*, vol. 83, pp. 1443-55, Feb 2009.

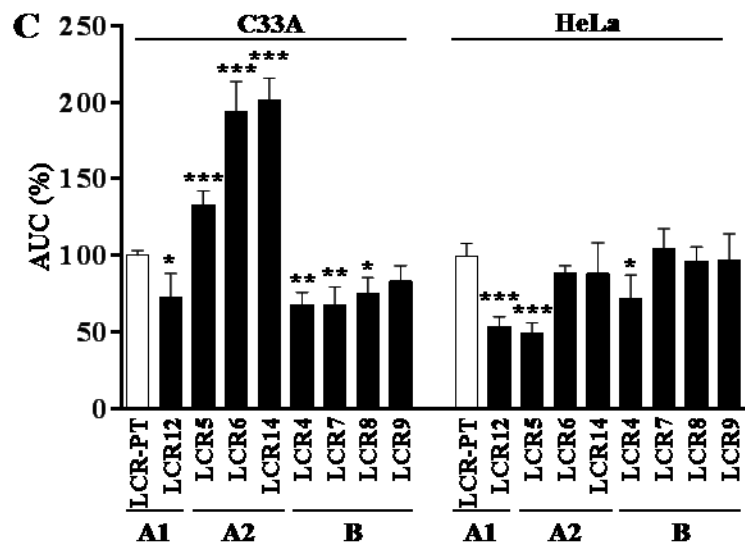
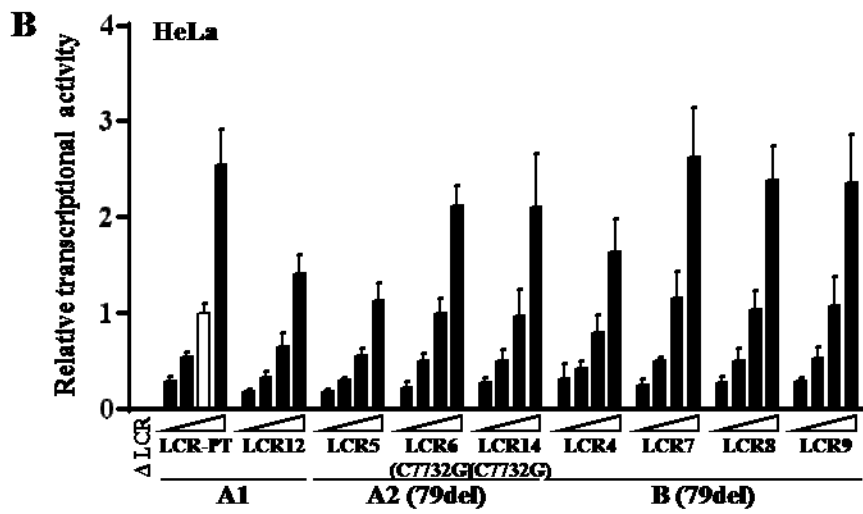
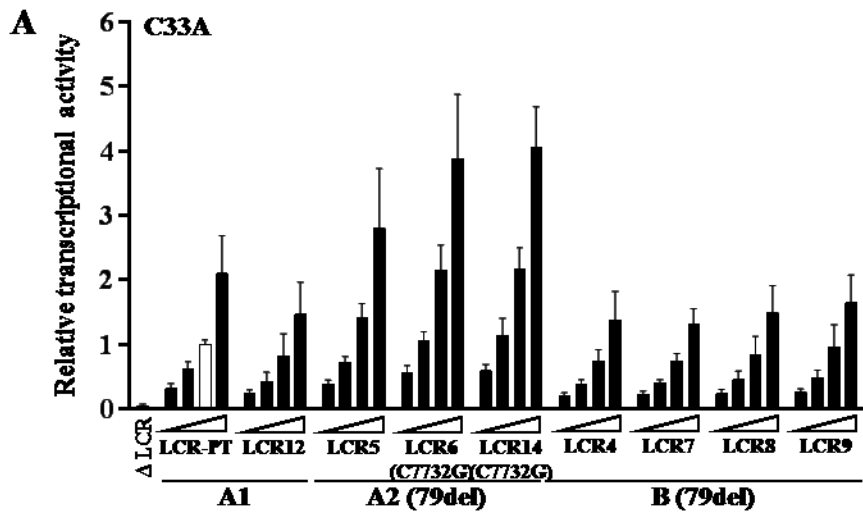


- [31] M. J. Lace, C. Isacson, J. R. Anson, A. T. Lorincz, S. P. Wilczynski, T. H. Haugen, and L. P. Turek, "Upstream regulatory region alterations found in human papillomavirus type 16 (HPV-16) isolates from cervical carcinomas increase transcription, ori function, and HPV immortalization capacity in culture," *J Virol*, vol. 83, pp. 7457-66, Aug 2009.
- [32] L. F. Xi, M. Schiffman, L. A. Koutsky, J. P. Hughes, R. L. Winer, C. Mao, A. Hulbert, S. K. Lee, Z. Shen, and N. B. Kiviat, "Lineages of oncogenic human papillomavirus types other than type 16 and 18 and risk for cervical intraepithelial neoplasia," *J Natl Cancer Inst*, vol. 106, Oct 2014.
- [33] A. A. Chen, D. A. Heideman, D. Boon, Z. Chen, R. D. Burk, H. De Vuyst, T. Gheit, P. J. Snijders, M. Tommasino, S. Franceschi, and G. M. Clifford, "Human papillomavirus 33 worldwide genetic variation and associated risk of cervical cancer," *Virology*, vol. 448, pp. 356-62, Jan 5 2014.
- [34] S. Corre and M. D. Galibert, "Upstream stimulating factors: highly versatile stress-responsive transcription factors," *Pigment Cell Res*, vol. 18, pp. 337-48, Oct 2005.
- [35] Y. Qyang, X. Luo, T. Lu, P. M. Ismail, D. Krylov, C. Vinson, and M. Sawadogo, "Cell-type-dependent activity of the ubiquitous transcription factor USF in cellular proliferation and transcriptional activation," *Mol Cell Biol*, vol. 19, pp. 1508-17, Feb 1999.
- [36] C. V. Dang, "MYC on the path to cancer," *Cell*, vol. 149, pp. 22-35, Mar 30 2012.
- [37] M. Ceribelli, D. Dolfini, D. Merico, R. Gatta, A. M. Vigano, G. Pavesi, and R. Mantovani, "The histone-like NF-Y is a bifunctional transcription factor," *Mol Cell Biol*, vol. 28, pp. 2047-58, Mar 2008.

### 2.3.7 Figures and Tables

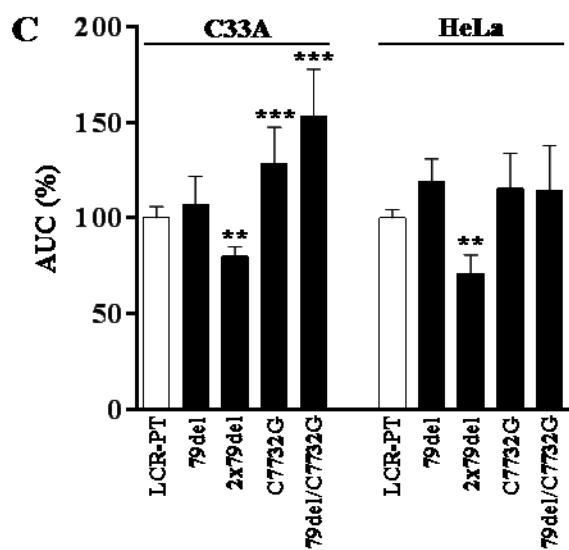
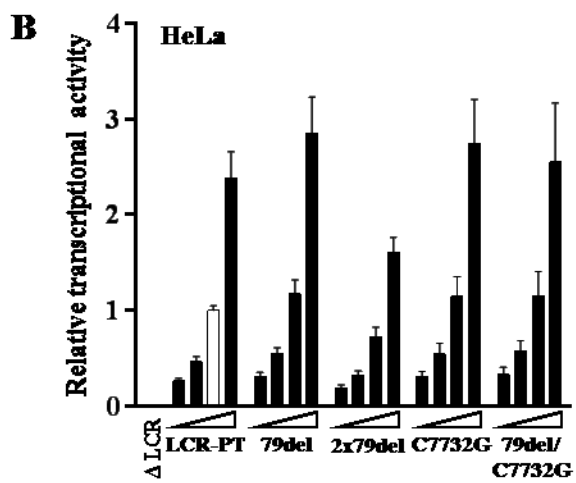
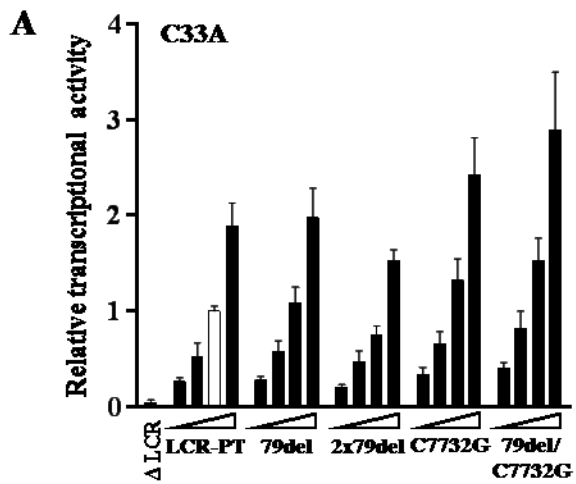


**Figure 2.1. Phylogeny of HPV33 Long Control Region variants.** (A) Schematic representation of the HPV33 LCR. The 924-bp LCR is represented by a white line flanked on its 5' and 3' side by the L1 and E6 open reading frames, respectively. The locations of 4 binding sites for the E2 transcriptional regulator and of a cluster of binding sites for the E1 helicase, which are common features of the LCR from anogenital HPV types, are indicated. Black boxes in the central portion of the LCR represent the duplicated 79-bp region present in the HPV33 prototype. The location of the 79-bp deletion and of the C7732G variation (grey box) are indicated. Nucleotide (nt) are numbered according to the HPV33 reference genome (GenBank accession M12732.1). (B) Phylogenetic tree of HPV33 LCR variants. The complete sequences of the LCR from the indicated HPV33 isolates were aligned using the Clustal Omega program and their phylogenetic relationships represented by a tree constructed with the EvolView software. LCRs are grouped into the A1 and A2 sublineages and B lineage, as indicated. The arrow points to the HPV33 reference LCR. The eight LCR variants characterized in this study are boxed.

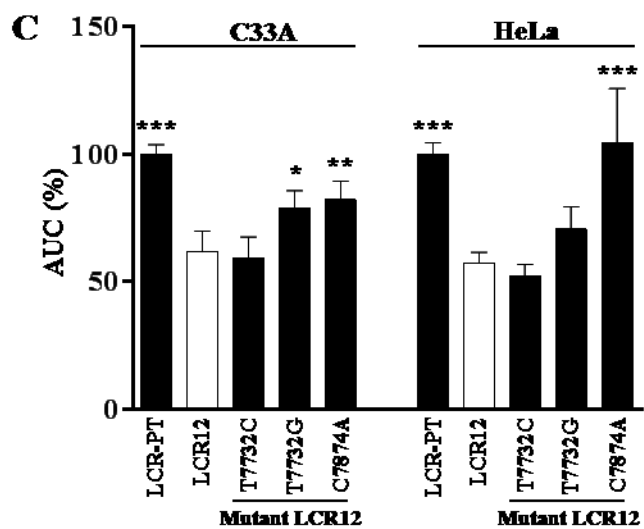
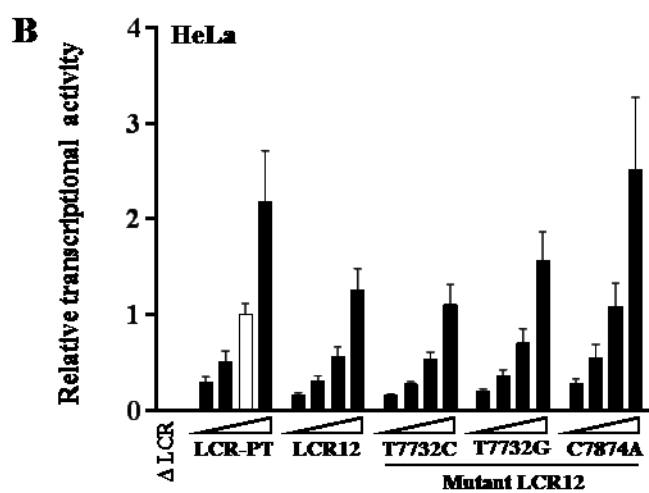
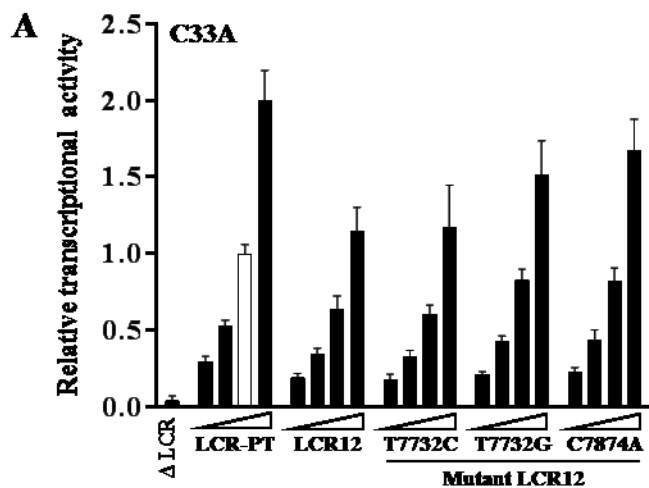


**Figure 2.2 Transcriptional activity of HPV33 LCR variants in C33A and HeLa cells.**

A) Bar graph representing the transcriptional activity of the indicated variant LCRs, relative to that of the prototype (LCR-PT), in either C33A (A) or HeLa (B) cells. The activity of each LCR was measured in luciferase reporter-gene assays using increasing quantities of LCR-Rluc plasmid (25, 50, 100 and 200 ng). Each bar represents the mean value obtained from a minimum of 4 independent experiments performed in duplicates. All values are presented relative to the level of transcriptional activity obtained with 100 ng of LCR-PT, which was used as the reference (white bar) and assigned a value of 1.0. Standard deviations are indicated by error bars. The level of transcriptional activity measured with 200 ng of Rluc plasmid lacking a LCR was used as a negative control ( $\Delta$ LCR). The (sub)lineage of each LCR is indicated at the bottom of the graph. The presence of the C7732G variation and/or the 79del in the variants is indicated. (C) Integrated measure of LCR transcriptional activity using the area under the curve (AUC). AUC values were calculated from the dose-response curves presented in panels A and B and used as an integrated measure of LCR activity in C33A and HeLa cells, as indicated. The AUC value of each LCR variant is reported as a percentage of the AUC value calculated for LCR-PT, which was set as 100. Statistical analysis of AUC values was performed using a one-way ANOVA followed by Dunnett's post-hoc analysis to compare the transcriptional activity of LCR variants with that of the prototype (white bar). \* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$ . Variant (sub)lineages are indicated at the bottom of the graph.

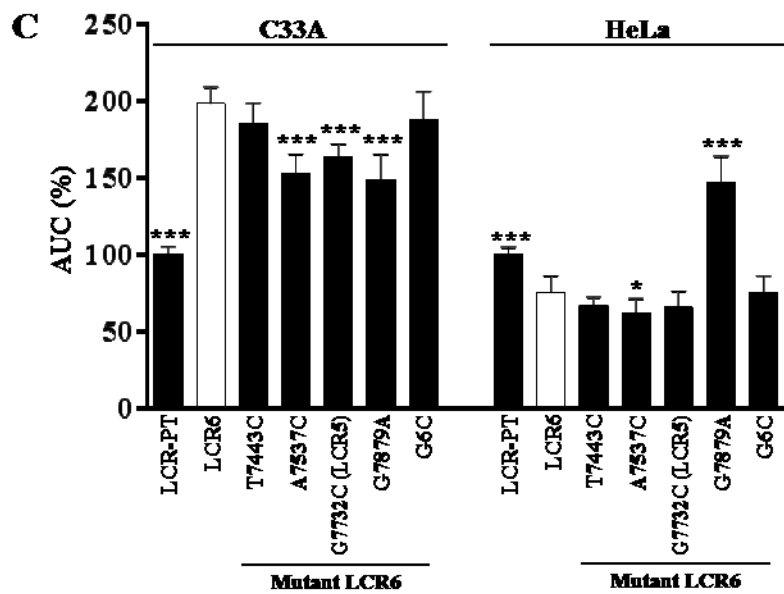
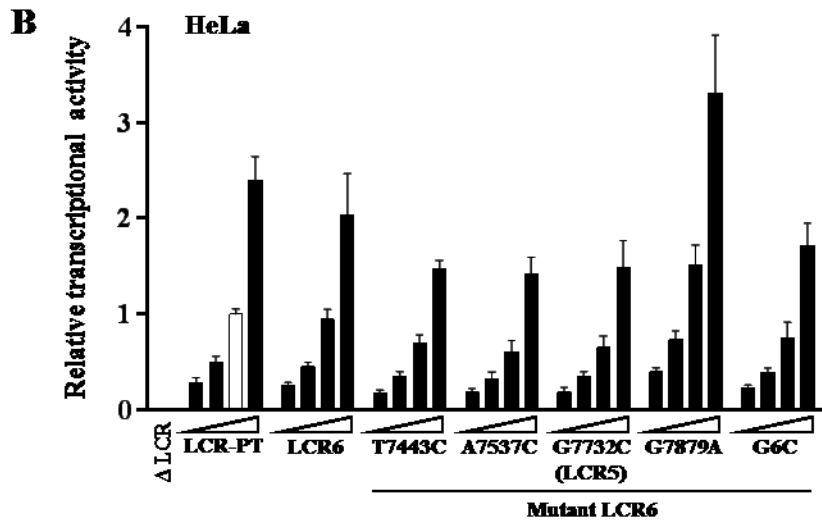
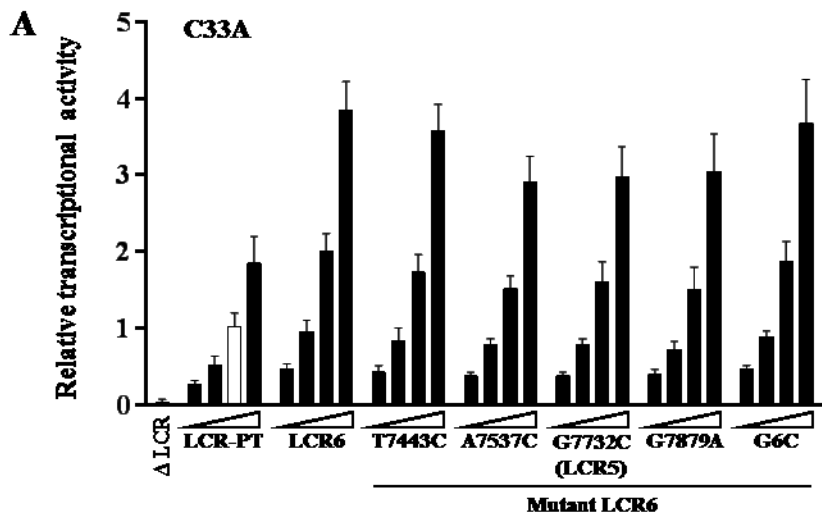


**Figure 2.3. Effect of the 79-bp deletion and of the C7732G variation on the transcriptional activity of the prototype HPV33 LCR.** The 79-bp deletion (79del) and the A2-sublineage specific variation C7732G were introduced, either separately or together, into the prototype HPV33 LCR (LCR-PT) by site-directed mutagenesis. A double 79-bp deletion was introduced into the prototype HPV33 LCR by site-directed mutagenesis. The transcriptional activities of the resulting LCRs were measured using increasing quantities of LCR-Rluc plasmid in either C33A (A) or HeLa (B) cells, exactly as described in the legend of Fig. 2.2. An integrated measure of the transcriptional activity of each LCR (C) was obtained by calculating the area under each dose-response curve (AUC), as described in the legend of Fig. 2.2.C.

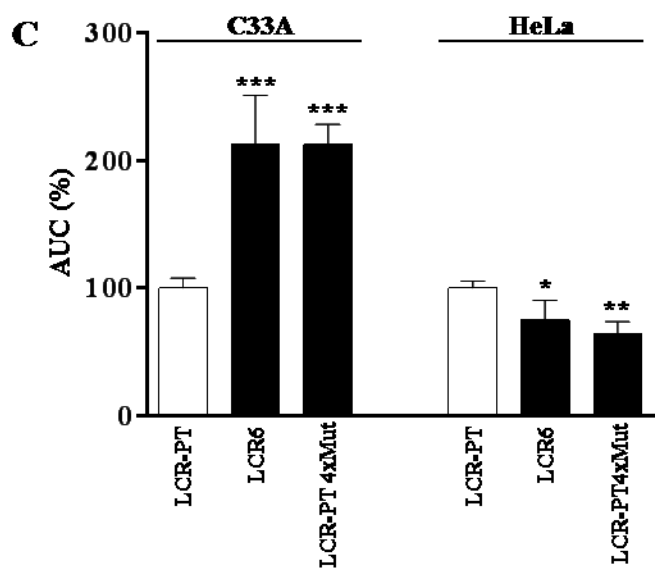
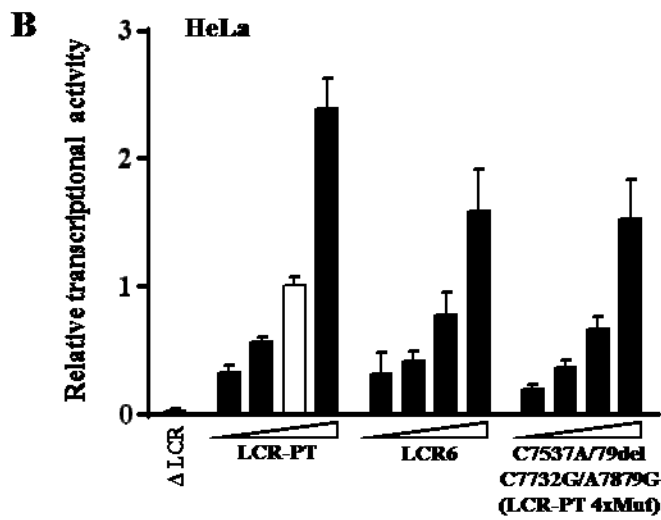
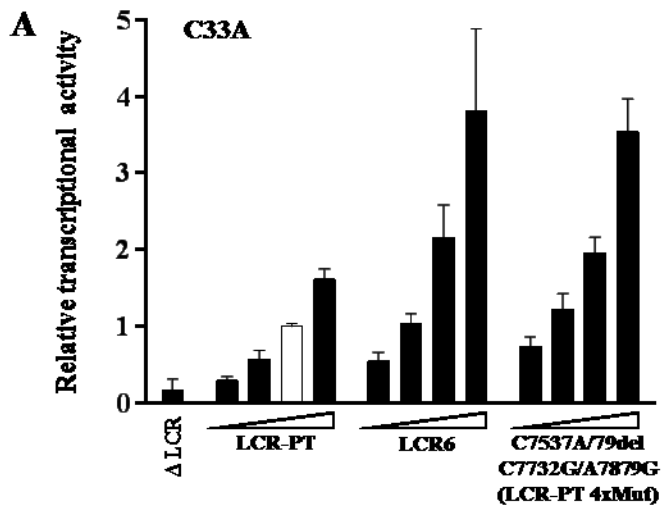




**Figure 2.4. Effect of the C7732T and A7874C variations on the transcriptional activity of the HPV33 LCR12 variant.** The contribution of the C7732T and A7874C variations to the transcriptional activity of LCR12 was tested by reverting each of these two variations, individually, back to the sequence of the prototype (LCR-PT), by site-directed mutagenesis. As a control, the C7732G variation was also introduced in LCR12 to confirm its stimulatory effect in C33A cells. The transcriptional activities of the resulting mutant LCR12 derivatives (carrying, respectively, the T7732C, T7732G and C7874A mutation, as indicated) were then measured using increasing quantities of LCR-Rluc plasmid in either C33A (A) or HeLa (B) cells, as described in the legend of Fig. 2.2. An integrated measure of the transcriptional activity of each LCR (C) was obtained by calculating the area under each dose-response curve (AUC), as described in the legend of Fig. 2.2.C, with the values obtained for the intact LCR12 variant (white bar) used as the reference for statistical analysis.



**Figure 2.5. The specific C7537A, C7732G and A7879G variations in combination with the 79 bp deletion are responsible for the enhanced transcriptional activity of HPV33 A2-sublineage variant LCRs.** The contribution of the five A2-specific variations, C7443T, C7537A, C7732G, A7879G and C6G to the transcriptional activity of LCR6 was tested by reverting each of these variations, individually, back to the sequence of the prototype (LCR-PT), by site-directed mutagenesis. Note that reversion of C7732G would yield a LCR6-derivative that is identical in sequence to LCR5; thus the latter was used in these experiments. The transcriptional activities of the resulting LCR6 derivatives (carrying, respectively, the T7443C, A7537C, G7732C, G7879A and G6C mutation, as indicated) were measured using increasing quantities of LCR-Rluc plasmid in either C33A (A) or HeLa (B) cells, exactly as described in the legend of Fig. 2.2. An integrated measure of the transcriptional activity of each LCR (C) was obtained by calculating the area under each dose-response curve (AUC), as described in the legend of Fig. 2.2.C, with the values obtained for the intact LCR6 variant (white bar) used as the reference for statistical analysis.



**Figure 2.6. The 79bp deletion increases the activity of the prototypical LCR harbouring the C7537A, C7732G and A7879G variations in C33A cells, but induces a decrease in gene expression in HeLa cells.** The 79-bp deletion and the three A2-specific variations (C7537A, C7732G and A7879G) were introduced into the prototype HPV33 LCR by site-directed mutagenesis to obtain the quadruple mutant LCR-PT 4xMut. The transcriptional activities of the resulting mutant LCR was measured using increasing quantities of LCR-Rluc plasmid in either C33A (A) or HeLa (B) cells, exactly as described in the legend of Fig. 2.2, and compared to that of LCR-PT and the LCR6 variant. An integrated measure of the transcriptional activity of each LCR (C) was obtained by calculating the area under each dose-response curve (AUC), as described in the legend of Fig. 2.2.C.

**Table 2.1. Naturally-Occurring Polymorphisms in the Long Control Region of Human Papillomavirus Type 33**

		Position	7116	7128	7174	7182	7198	7227	7362	7404	7412	7422	7424	7425	7442	7443	7454	7481	7529	7535	7537	7595	7686	7732	7874	7879	6	18
		Prototype	T	T	T	G	T	G	T	T	TA	G	A	C	G	C	G	T	-	G	C	79-bp	A	C	A	A	C	G
Lineage	Accession number	Variants	G	C	C	C	C	A	C	A/del	del	T	G	T	A	T	A	G	CCCTAATA	A	A	del	C	T/G	C	G	G	A
A1	KT827351	LCR12	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	C	-	-	-
A2	KT827346	LCR6	-	-	-	-	-	A	-	A	-	T	-	-	-	T	A	-	-	A	A	del	-	G	-	G	G	A
A2	KT827345	LCR5	-	-	-	-	-	A	-	A	-	T	-	-	-	T	A	-	-	A	A	del	-	-	-	G	G	A
A2	KT827352	LCR14	-	-	-	-	-	A	-	A	-	T	-	-	-	T	A	-	-	A	A	del	C	G	-	G	G	A
B	KT827347	LCR7	G	C	C	C	C	A	C	A	-	T	G	T	T	-	A	G	-	A	-	del	-	-	-	-	-	A
B	KT827348	LCR8	G	C	C	C	C	A	C	A	-	T	G	T	T	-	A	G	CCCTAATA	A	-	del	-	-	-	-	-	A
B	KT827344	LCR4	G	C	C	C	C	A	C	del	-	T	G	T	T	-	A	G	-	A	-	del	-	-	-	-	-	A
B	KT827349	LCR9	G	C	C	C	C	A	C	A	del	T	G	T	T	-	A	G	-	A	-	del	-	-	-	-	-	A

**NOTE.** Variant positions of the LCR are numbered according to the sequence of the HPV33 prototype genome. Nucleotides found at these positions of the prototype and variant LCRs are indicated by a letter (A, C, G or T) or by the sequence “CCCTAATA” in the case of the insertion at position 7529 of LCR8. Deletions of nucleotide 7404, 7412 or of one copy of the duplicated 79-bp region are indicated by the abbreviation “del”. Hyphens indicate nucleotide positions that are identical to the prototype.

**Table 2.2. Nucleotide Variations affecting Transcriptional Activity of the Long Control Region of Human Papillomavirus Type 33**

Variation(s)	LCR transcriptional activity		Putative transcription factor
	C33A	HeLa	
<b>C7443T<sup>b</sup></b>	silent	silent	unknown
<b>C7537A<sup>b</sup></b>	+ 25 %	+ 20 %	USF
<b>79del<sup>a</sup></b>	silent	silent	USF, NF-Y, c-Myb
<b>2x79del<sup>a</sup></b>	- 20 %	- 30 %	USF, NF-Y, c-Myb
<b>C7732G<sup>a</sup></b>	+ 25 %	silent	USF, c-Myc
<b>79del/C7732G<sup>a</sup></b>	+ 50 %	silent	USF
<b>C7732T<sup>c</sup></b>	silent	silent	USF, c-Myc
<b>A7874C<sup>c</sup></b>	- 35 %	- 80 %	AP1
<b>A7879G<sup>b</sup></b>	+ 25 %	- 95 %	AP1
<b>C6G<sup>b</sup></b>	silent	silent	unknown
<b>4xMut (79del/ C7537A/C7732G/A7879G)<sup>a</sup></b>	+ 110%	- 35%	Various

**NOTE.** Percentages refer to the degree of activation (+) or repression (-) of LCR transcriptional activity and were calculated from the AUC values.

<sup>a</sup> When introduced in the HPV33 prototypical LCR.

<sup>b</sup> When reverted to the prototypical sequence in the LCR6 variant.

<sup>c</sup> When reverted to the prototypical sequence in the LCR12 variant.

# **Chapitre troisième : Un polymorphisme rare augmente l'activité transcriptionnelle d'une variante de la longue région de contrôle du VPH33 en abolissant un des deux sites de liaison pour le facteur cellulaire C/EBP $\beta$**

**Title:** A Rare Polymorphism Enhances the Transcriptional Activity of the HPV33 Long Control Region by Disrupting One of Two Regulatory Binding Sites for the Cellular Transcription Factor C/EBP $\beta$

**Running title:** A C/EBP $\beta$  binding site variant of the HPV33 LCR (limit of 40 characters)

Jennifer Alvarez<sup>1,2</sup>, Francois Coullée<sup>3</sup>, Jacques Archambault<sup>1,2\*</sup>.

<sup>1</sup> Institut de Recherches Cliniques de Montréal (IRCM), <sup>2</sup> Department of Biochemistry and Molecular Medicine, Université de Montréal, <sup>3</sup> Centre de Recherche et Département de Microbiologie Médicale et Infectiologie, Centre Hospitalier de l'Université de Montréal (CHUM), Université de Montréal, Montréal Québec, Canada.

Abstract word count: 269

Text word count: 5026

## **FOOTNOTE**

Presented in part: DNA Tumour Virus Meeting, Montreal, Canada, 15-21 July 2012 and ICGEB DNA Tumour Virus Meeting, Trieste, Italy, 20-26 July 2015.

Correspondence: Jacques Archambault, Laboratory of Molecular Virology, Institut de Recherches Cliniques de Montréal, 110 Pine Avenue West, Montréal, QC, Canada, H2W 1R7 ([jacques.archambault@ircm.qc.ca](mailto:jacques.archambault@ircm.qc.ca))



## NOTES

**Acknowledgements.** The authors would like to thank Eduardo Franco for his constant support, Janice Clements and Lou Laimins for providing the C/EBP $\beta$  LAP and LIP expression plasmids and members of the Archambault laboratory for their helpful comments on this manuscript.

**Financial support.** This work was supported by grants from the Canadian Cancer Society Research Institute (CCSRI grant #020298), Canadian Institutes of Health Research (CIHR team grant CRN-83320) and FRQS Réseau SIDA MI.

**Potential conflicts of interest.** The authors have no conflict of interest to declare relevant to the topic of this article.

### 3.1 Mise en contexte de l'article

Dans le chapitre précédent, nous avons montré que plusieurs variations contribuaient à augmenter l'activité transcriptionnelle des variantes de la sous-lignée A2 du VPH33. Dans ce chapitre, nous avons identifié une nouvelle variante de la lignée B qui était plus polymorphique que les isolats que nous avons préalablement décrits. Ainsi, nous avons voulu caractériser cette nouvelle variante. Du fait que cette variante montre une activité accrue par rapport au prototype, nous l'avons utilisée comme modèle d'étude de validation de concept que les variations modulent l'expression des gènes viraux en altérant la liaison de certains FT. À noter que cet article sera soumis pour publication prochainement.

### 3.2 Contribution des auteurs

J'ai réalisé toutes les expériences présentées dans ce manuscrit. J'ai également rédigé en grande partie cet article en collaboration avec Dr Jacques Archambault.

Dr François Coutlée a participé dans la conception et la supervision de ce projet. Il a également contribué à la correction du manuscrit.

Dr Jacques Archambault a participé à la conception et à la supervision du projet. Il a également grandement contribué à la rédaction et à la correction de cet article.

### 3.3 Texte de l'article

#### 3.3.1 Abstract

**Background.** Expression of the HPV genome is regulated by the viral long control region (LCR) which contains binding sites for several cellular transcription factors and for the viral E2 protein. Variants of the HPV33 LCR exhibiting altered transcriptional levels have been identified. The characterization of a rare HPV33 LCR variant, LCR10, which displays significantly higher activity than the prototypical LCR (LCR-PT) is presented here.

**Methods.** LCR activity was measured in luciferase reporter gene assays. The contribution of individual variations to the activity of LCR10 was assessed by reverting them to the prototypical sequence. The consequences of a specific LCR10 variation and of other mutations in predicted C/EBP $\beta$  binding sites was determined by introducing them into LCR-PT, as well as by assessing their effect on the ability of the LIP and LAP isoforms of C/EBP $\beta$  to inhibit and activate the LCR, respectively. The effect of nucleotide changes on the affinity of C/EBP $\beta$  for its target binding sequence was determined in vitro, by fluorescence polarization.

**Results.** The enhanced activity of LCR10 could be attributed to the specific T7791C variation that disrupts a negatively-acting C/EBP $\beta$  binding site (BS1). Loss of BS1 alleviates C/EBP $\beta$ -mediated repression of the HPV33 LCR. A second functional C/EBP $\beta$  binding site (BS2) was also identified which, in contrast to BS1, increases the activity of the LCR.

**Conclusions.** Two C/EBP $\beta$  binding sites have been identified that negatively and positively regulate the transcriptional activity of the HPV33 LCR, respectively. The rare T7791C variation increases the activity of the LCR by disrupting the binding of C/EBP $\beta$  to the negatively-acting site. These results highlight C/EBP $\beta$  as an important regulator of HPV33 gene expression.

**Keywords.** HPV, LCR, variants, C/EBP $\beta$ , transcription factor, gene expression

### 3.3.2 Introduction

Persistent infections by high-risk human papillomaviruses (HR-HPVs) have been associated with an increased risk of developing cervical cancer (CC) and other malignancies [1, 2]. The E6 and E7 viral oncogenes not only promote cancer progression but remain essential for CC cells to maintain their transformed phenotype, as first demonstrated in HeLa cells [3]. E6 and E7 are expressed from the long control region (LCR), a regulatory part of the viral genome that contains promoter and enhancer sequences comprised of binding sites (BS) for activating or inhibitory transcription factors (TF) such as Sp1, AP-1, Oct1, YY1 and C/EBP $\beta$  and for the viral E2 transcriptional repressor [4, 5]. E6 and E7 expression is frequently up-regulated in CC and other HPV-associated cancers, often following integration of the viral genome into cellular DNA and concomitant disruption of the E2 open reading frame (ORF), which results in derepression of the LCR.

The basic leucine zipper TF C/EBP $\beta$  (CCAAT/Enhancer-binding Protein  $\beta$ ) regulates many cellular processes including adipogenesis, hematopoiesis and epithelial differentiation and is also usurped by certain HPV types to regulate viral gene transcription (reviewed in [5-9]). Translation initiation at different sites of the C/EBP $\beta$  mRNA gives rise to three isoforms of 38, 35 and 20 kDa known as the liver-enriched activator proteins LAP\* and LAP (herein termed LAP) and the liver-enriched inhibitory protein LIP, respectively [10, 11]. The ability of C/EBP $\beta$  to activate or repress transcription is determined by the relative expression (i.e. ratio) of the LAP and LIP isoforms, by interacting TFs and by post-translational modifications such as phosphorylation which regulates its DNA-binding activity [5-10, 12, 13].

Discrepancies have been noted in how C/EBP $\beta$  regulates distinct HPV types in different cell lines. Analysis of the HPV18 regulatory region revealed that its activity in HeLa cells depends on the assembly of a specific C/EBP $\beta$ -YY1 complex on the so-called “switch region” of the LCR [14, 15]. In contrast, overproduction of C/EBP $\beta$  has been shown to repress the HPV18 and HPV16 LCRs in HeLa cells and the HPV11 LCR in primary human keratinocytes [16-19]. For reasons unknown, overexpression of C/EBP $\beta$  has the opposite effect in C33A cells, resulting

in transactivation of the HPV11 and HPV16 LCRs [12, 18]. An elegant study in organotypic raft cultures revealed that mutation of either one of two C/EBP sites in the HPV11 LCR increases its activity in basal cells, suggesting that C/EBP $\beta$  functions primarily as a repressor of viral gene expression in undifferentiated keratinocytes [20]. A similar study with HPV31 indicated that LIP is the predominant C/EBP $\beta$  isoform in undifferentiated cells but that its expression decreases upon differentiation such as to favor LAP-induced transactivation of the viral late promoter [13]. These studies emphasize the dual role that C/EBP $\beta$  plays as an activator or repressor of HPV gene expression, depending on the LCR structure, the LAP/LIP ratio and the cellular complement of TFs.

Our previous study of the HPV33 LCR has indicated that variants from the B-lineage have transcriptional activities that are comparable to that of the prototypical LCR (LCR-PT). In the present study, we characterized an unusual B-lineage variant, termed LCR10, which shows significantly higher transcriptional activity than LCR-PT. Our results revealed that the enhanced activity of LCR10 is caused primarily by a single variation, T7791C, which abolishes an inhibitory binding site for C/EBP $\beta$ , thus partly relieving repression of the LCR. Interestingly, a second C/EBP $\beta$  BS was also identified which, in contrast to the first one, increases the activity of the HPV33 LCR. These results suggest that C/EBP $\beta$  is both an activator and a repressor of the HPV33 LCR, depending on the context of its binding site.

### **3.3.3 Materials and Methods**

#### **DNA sequence of the HPV33 LCR10 variant**

The HPV33 LCR10 variant (accession number KT827350) previously isolated from a cervical sample from a woman co-infected with the human immunodeficiency virus (HIV) [21, 22] was amplified by PCR and completely sequenced as described in [23].

#### **Computational analysis of transcription factor binding sites**

Putative TF BS that are altered in the HPV33 LCR10 variant were predicted with the PROMO software (TRANSFAC version 8.3 database) [24, 25].

## **Plasmid construction and mutagenesis**

The plasmid expressing Renilla luciferase (Rluc) from the prototypical HPV33 LCR (LCR-PT) as well as the internal control plasmid encoding firefly luciferase (Fluc) downstream of LCR-PT (LCR-PT-Fluc) have been described in [23]. The plasmid expressing Rluc from the LCR10 variant was obtained similarly. Variations were introduced in LCR-PT or reverted in LCR10 using the QuikChange Site-Directed Mutagenesis kit (Stratagene). Mutagenesis was also used to create a mutant version of LCR-PT, mBS1, in which the C/EBP $\beta$  BS spanning T7791 was inactivated by four mutations (GCTTACACAATT changed to GCCCACACCCTT, mutated nucleotides underlined). Two other putative C/EBP $\beta$  BS (BS2 and BS3) were similarly mutated, either individually or in combination with mBS1. A similar mutagenesis strategy was used to create plasmid LCR- $\Delta$ C/EBP $\beta$ -Fluc, a derivative of LCR-PT-Fluc in which the three putative C/EBP $\beta$  BS in LCR-PT were each inactivated as described above. Plasmids used for bacterial expression of rat C/EBP $\beta$  fragments fused to glutathione-S-transferase (GST) were constructed by inserting an EcoRI-SalI-digested PCR product encoding either the basic region alone (B, amino acids 262-298) or the basic region and leucine zipper domain (BZ, amino acids 262-345) into the pGEX-4T1 vector. All constructs were verified by DNA sequencing. Additional information on the construction of these plasmids will be made available upon request.

## **Cell culture and transfection**

The two cervical carcinoma cell lines C33A (HPV-negative) and HeLa (HPV18-positive), as well as the osteosarcoma-derived cell line U2OS, were grown and transfected using Lipofectamine 2000 (Life Technologies, cat: 11668-500) as previously described [23].

## **Luciferase-reporter transcription assay**

This protocol is described in [23]. Briefly, C33A, HeLa or U2OS cells were co-transfected with increasing quantities of LCR-Rluc plasmid (25, 50, 100 and 200 ng) and 25 ng of the LCR-PT-Fluc control plasmid. Fluc and Rluc activities were measured with the Dual-Glo luciferase assay system and a GloMax<sup>TM</sup> 96-well luminometer (Promega) 24h post-transfection. LCR transcriptional activities are presented either as Rluc/Fluc ratios or

relative to the transcriptional activity measured with 100 ng of LCR-PT, which was assigned a value of 1.0. In C/EBP $\beta$  overexpression studies, 100 ng of LCR-Rluc plasmid (LCR-PT, T7791C and C/EBP $\beta$  binding mutants) was co-transfected into C33A, HeLa and U2OS cells with a rat C/EBP $\beta$ -expression plasmid encoding either the LAP (5 ng) or LIP (10 ng) isoform (kindly provided by Janice Clements via Lou Laimins). The internal control LCR- $\Delta$ C/EBP $\beta$ -Fluc plasmid was used in these experiments. All other steps were performed as described above.

### **Statistical analysis**

Statistical significance was tested using a one-way ANOVA with Dunnett's post-hoc analysis. Statistics and AUC calculations were performed using GraphPad Prism version 6.00. The Bonferonni correction was applied where multiple comparisons were made.

### **Western blot and antibodies**

Plasmids encoding the rat C/EBP $\beta$  LAP and LIP isoforms (0.5  $\mu$ g) were transfected into C33A, HeLa and U2OS cells. Transfected-cell extracts were prepared, separated by SDS-PAGE, transferred onto polyvinylidene fluoride membranes and the endogenous and transiently-expressed C/EBP $\beta$  proteins detected using a rabbit polyclonal antibody (Santa Cruz Biotechnology, cat: sc-150) and a horseradish peroxidase-conjugated sheep anti-rabbit antibody (GE healthcare, cat: NA934).  $\beta$ -tubulin was detected using a mouse monoclonal antibody (Sigma-Aldrich, cat: T4026) and a horseradish peroxidase-conjugated sheep anti-mouse antibody (GE healthcare, cat: NA931). All proteins were detected using an enhanced chemiluminescence detection kit (GE Healthcare). Relative protein levels were determined with ImageJ [26].

### **Protein expression and purification**

GST-C/EBP $\beta$  fusion proteins were expressed and purified as previously described [27] with the following modifications. Briefly, bacterial cultures of *E. coli* BL21 cells (Novagen) transformed with the GST, GST-C/EBP $\beta$ -B and GST-C/EBP $\beta$ -BZ expression plasmids were grown in Luria-Bertani medium to an optical density of 0.5 at 595 nm and induced for 4 hours by the addition of 0.5 mM of isopropyl- $\beta$ -D-thiogalactopyranoside.

Cells harvested from a 2-liter culture were resuspended in 30 mL of lysis buffer (50mM Tris-HCl [pH 7.6], 1M NaCl, 5mM EDTA, 5mM DTT, 10% glycerol, 0.1% NP-40, 10ug/mL antipain, 2ug/mL leupeptin, 1ug/mL pepstatin A, 2ug/mL aprotinin and 1mM phenylmethanesulfonylfluoride). Cells were disrupted by sonication and lysates were cleared by centrifugation at 9000g for 20 minutes and incubated with 2 mL of glutathione-Sepharose 4B beads (GE Healthcare, cat: 17-0756-01) for 3 hours at 4°C. Beads were washed with 15 bed volumes of high-salt buffer (50 mM Tris-HCl [pH 8.0], 0.5M NaCl, 5mM EDTA, 5mM DTT, 10% glycerol, 0.1% NP-40) followed by 15 bed volumes of low-salt buffer (0.2 M NaCl). Proteins were eluted in elution buffer (25mM Tris-HCl [pH 8.0], 0.2M NaCl, 1mM EDTA, 5mM DTT, 10% glycerol, 0.1% NP-40, 0.01M reduced glutathione) and recovered after centrifugation at 5000g for 5 minutes. Purified protein concentrations were determined by Bradford analysis.

### **Fluorescence polarization**

Binding reactions were performed in OptiPlate-96F high-binding-affinity black 96-well plates (Perkin Elmer) in a final volume of 150 $\mu$ L of assay buffer (25mM Tris-HCl [pH 7.6], 0.1mM DTT, 0.01% NP-40, 75mM NaCl, 150mM KCl and 5mM MgCl<sub>2</sub>) containing 15nM of fluorescein-labeled BS1, T7791C or mBS1 probe and the indicated concentration of purified GST, GST-C/EBP $\beta$ -B or GST-C/EBP $\beta$ -BZ protein, as described previously [28]. Fluorescence readings were taken after 20 minutes of incubation at room temperature with a Victor<sup>3</sup>V multilabel plate-reader (Perkin Elmer) equipped with a 485-nm/535-nm filter set. The background fluorescence of the probe was subtracted from all values. Competition binding assays were performed as described in [28] using a final concentration of 15 nM of fluorescein-labeled BS1 probe, 125 nM of GST-C/EBP $\beta$ -BZ protein and increasing amount of unlabeled competitors (0, 5.9, 11.7, 23.4, 46.9, 93.8, 187.5, 375, 750 and 1500 nM) in a final volume of 150 $\mu$ L of assay buffer (25mM Tris-HCl [pH 7.6], 0.1mM DTT, 0.01% NP-40, 75mM NaCl, 150mM KCl and 5mM MgCl<sub>2</sub>). Fluorescence readings were taken as described above.

### **Fluorescent DNA probe and competitor oligonucleotides**

The fluorescein-labeled oligonucleotide with the fluorophore attached at the 5' end by a six-carbon linker and the unlabeled oligonucleotides were purchased from Integrated

DNA Technologies (IDT). The nucleotide sequence of the fluorescein-labeled probe is as follows (C/EBP $\beta$  site underlined and mutated nucleotide in bold): BS1, TTTGGCTTACACAATTGCTTT. Unlabeled duplex DNAs and the fluorescein-labeled probes were prepared by annealing complementary oligonucleotides at a ratio of 1:1 and of 1:1.5, respectively. The mixture was heated at 95°C for 5 minutes in annealing buffer (25mM Tris-HCl [pH 7.6], 1mM DTT, 0.1M NaCl, 0.01% NP-40) and allowed to cool slowly to room temperature as previously described [28]. The sequences of the unlabeled oligonucleotide are reported in Table 3.1: Consensus (shorter version of the C/EBP $\beta$ -BS-containing oligonucleotide described in [29]), BS1, BS2, BS3 T7791C, mBS1, mBS2 and mBS3.

### **K<sub>D</sub> and K<sub>i</sub> determination**

K<sub>D</sub> values were obtained by fitting the data from direct binding curves with nonlinear least-squares regression to a standard equation describing a single binding-site equilibrium, using GraphPad Prism version 6.00, as described previously [28]. Values were obtained from three independent experiments with each data point performed in duplicate. No corrections for emission intensity were performed, since the quantum yield of the fluorescent probes did not change significantly upon C/EBP $\beta$  binding. Inhibitory constant (K<sub>i</sub>) values were calculated with the GraphPad Prism version 6.00 software by fitting the data from inhibitory dose-response curves to a single binding-site equilibrium model using nonlinear least-squares regression analysis, as previously described [28]. K<sub>i</sub> values were obtained from three independent experiments with each data point performed in duplicate.

### **3.3.4 Results**

#### **The enhanced transcriptional activity of the HPV33 LCR10 variant is attributable to the rare T7791C variation**

Our previous characterization of several HPV33 LCR variants indicated that those from the B-lineage have transcriptional activities similar to that of LCR-PT, in C33A and HeLa cervical carcinoma cells [21-23]. Here, we report on an unusual LCR from this same lineage, LCR10, which exhibits a 2-fold increase in transcriptional activity compared to



LCR-PT and other related B-lineage variants (Fig. 3.1 and [23]). LCR10 contains several polymorphisms specific to B-lineage variants together with three unique variations, T7365C, C7531T and T7791C. To determine if these three variations account for the higher activity of LCR10, they were reverted, individually or in combination, to the prototypical sequence. The activity of the resulting mutant LCRs was then measured in C33A and in HeLa cells using our previously described luciferase reporter gene assay [23]. These mutant LCRs were also analyzed in U2OS cells, which have been shown recently to support episomal replication and amplification of the HPV genomes from several different HPV types, suggesting that the transcriptional regulatory elements of the LCR are fully functional in these cells [30]. As shown in Fig. 3.1, reversion of T7791C (C7791T) was sufficient to decrease the activity of LCR10 to levels that were similar to those of LCR-PT in C33A, HeLa and U2OS cells. Reversion of the other two polymorphisms did not accentuate this effect, showing that T7791C is primarily responsible for the enhanced activity of LCR10 in all three cell lines.

### **The T7791C variation is sufficient to enhance the transcriptional activity of the HPV33 LCR**

The T7791C variation was then introduced into LCR-PT to determine if it is sufficient to increase its activity. Figure 3.2B and 3.2C show that T7791C does indeed stimulate the transcriptional activity of LCR-PT, by about 75% in C33A, HeLa and U2OS cells. A bioinformatic analysis of the LCR region was then performed, which suggested that the T7791C variation could potentially alter a binding site for the cellular transcription factor C/EBP $\beta$  (consensus sequence: 5'-(A/G)TT(A/G)CGCAA(C/T)-3' [31]). Figure 3.2A depicts the position of this site in LCR-PT (C/EBP $\beta$ -BS; 5'-CTTACACAAT-3', T7791 underlined). If this sequence constitutes a genuine C/EBP $\beta$  BS, it should be completely inactivated by mutation of four conserved nucleotides (5'-CCCACACCCT-3', mutated nucleotides underlined) that are critical for interaction with C/EBP $\beta$  ([31] and data presented below). Introduction of this quadruple mutation into LCR-PT (mBS1) increased its activity to the same extent as T7791C in C33A and U2OS cells and to an even greater degree in HeLa cells (Fig. 3.2B and 3.2C), thus providing the first evidence that T7791C may abolish a negatively-acting C/EBP $\beta$  BS.

### **The T7791C variation weakens the affinity of C/EBP $\beta$ for its binding site in LCR-PT**

To confirm that the T7791C variation causes the loss of a C/EBP $\beta$  BS, a biochemical DNA binding assay based on fluorescence polarization was developed. C/EBP $\beta$  contains a leucine zipper domain (Z) and a basic region (B) that allows the protein to dimerize and contact DNA, respectively [11]. C/EBP $\beta$  fragments encompassing the B or BZ region were expressed and purified as fusions to glutathione-S-transferase (GST-C/EBP $\beta$ -B or GST-C/EBP $\beta$ -BZ, Fig. 3.3A). The capacity of these GST-C/EBP $\beta$  fusion proteins, and of GST alone, to bind a fluorescein-labeled DNA probe containing the C/EBP $\beta$  BS found in LCR-PT (BS1) was then determined. Titration of purified GST-C/EBP $\beta$ -BZ resulted in a dose-dependent increase in polarization indicative of its binding to the DNA probe (Fig. 3.3B). From this curve, a dissociation constant ( $K_D$ ) of  $62 \pm 3$  nM was measured for the interaction of GST-C/EBP $\beta$ -BZ with LCR-PT. This interaction was specific as neither GST-C/EBP $\beta$ -B (which lacks the Z domain) nor GST was able to bind to the probe (Fig. 3.3B).

The assay described above was then used to measure the affinity of GST-C/EBP $\beta$ -BZ for different binding sites, in competition experiments using unlabelled oligonucleotides containing either the C/EBP $\beta$  BS found in LCR-PT (BS1), the T7791C variant, the inactive mBS1 or a previously described high-affinity consensus C/EBP $\beta$  binding sequence [29]. The inhibition curves presented in Fig. 3.3C revealed that the GST-C/EBP $\beta$ -BZ protein bound to BS1 with an affinity of  $96 \pm 25$  nM that is only 2-fold lower than its affinity for the consensus site ( $40 \pm 2$  nM). Importantly, C/EBP $\beta$ -BZ bound only weakly to the T7791C variant site, with an affinity comparable to that of the negative control mBS1 ( $K_i > 1269$  nM). The results and the nucleotide sequence of each oligonucleotide are summarized in Table 3.1. Collectively, these results indicate that LCR-PT contains a high-affinity BS for C/EBP $\beta$  that is abrogated by the T7791C variation.

### **Identification of two additional binding sites for C/EBP $\beta$ in the LCR of HPV33**

Our computational analysis of the HPV33 LCR predicted two additional sites (BS2 and BS3) related to the C/EBP $\beta$  consensus binding-sequence 5'-ATTGCGCAAT-3' [29]. Figure 3.4A indicates the positions of BS2 and BS3 in the 5' and central regions of the HPV33 LCR, respectively, relative to the locations of BS1 and the four highly-conserved

E2-binding sites. The ability of C/EBP $\beta$  to interact with BS2 and BS3 was tested in vitro, using our competitive fluorescence polarization DNA binding assay.

The results presented in Figure 3.4B and in Table 3.1 show that GST-C/EBP $\beta$ -BZ bound tightly to the oligonucleotide carrying BS2 ( $K_i = 90 \text{ nM} \pm 28$ ) but only weakly to the one carrying BS3 ( $K_i = 181 \text{ nM} \pm 84$ ). These protein-DNA interactions appeared to be specific as they were not observed with mutant oligonucleotides in which BS2 and BS3 were inactivated by a quadruple mutation (mBS2 and mBS3, respectively). Taken together, these results indicated that the affinity of C/EBP $\beta$  for BS2 is similar to that for BS1 and close to its affinity for a consensus binding sequence. They also showed that BS3 is a rather weak BS.

### **Mutation of C/EBP $\beta$ BS1 enhances the transcriptional activity of the HPV33 LCR, whereas the mutation of BS2 reduces it**

The results presented above raised the possibility that C/EBP $\beta$  could modulate the transcriptional activity of the HPV33 LCR through its interaction with BS1 and BS2, but is unlikely to do so through BS3 given its weak affinity for this site. To test these predictions, BS2 and BS3 were inactivated (with the quadruple mutation described above) in LCR-PT, either separately or in combination with each other or with mBS1. The transcriptional activities of the resulting mutant LCRs were then measured using a luciferase reporter-gene assay in C33A, HeLa and U2OS cells.

The results presented in Figure 3.5A and 3.5D reveal that the inactivation of BS1 enhances the LCR activity either by itself or in combination with mBS2 and mBS3 in C33A cells, showing that the inhibitory effect of this site is dominant in this cell line. In contrast, the inactivation of BS2 had little to no effect on the activity of the LCR in C33A cells, but significantly decreased its activity in HeLa cells, by approximately 4-fold. Mutation of BS2 also led to a small decrease in LCR-activity in U2OS cells, although this difference did not reach statistical significance. The positive contribution that BS2 plays in U2OS cells, as well as in HeLa cell, became more obvious when this site was mutated in combination with BS1; the double mutant LCR beings approximately 3-fold less active than the single mBS1 LCR ( $p < 0.002$ ), showing that the function of BS2 is dominant in

these cell lines. Thus, in contrast to BS1 which negatively regulates the LCR in all three cell lines, BS2 appears to play a stimulatory role in HeLa and U2OS cells. As for BS3, its inactivation had no significant effect in all three cell lines, consistent with its low affinity for C/EBP $\beta$  in vitro. Collectively, these results indicate that BS2 can be an important activator of the HPV33 LCR depending on the cellular context, whereas BS1 acts as a negative regulator of viral gene expression.

### **Modulation of the HPV33 LCR by exogenous expression of the C/EBP $\beta$ LAP and LIP isoforms.**

The results presented above suggest that the HPV33 LCR is regulated by C/EBP $\beta$ . To further substantiate this notion, we investigated the effect of overproducing C/EBP $\beta$ , either the activating LAP or inhibitory LIP isoform, on the activity of LCR-PT. Expression plasmids for LAP or LIP were co-transfected with the LCR-PT-Rluc plasmid in C33A, HeLa and U2OS cells. LCR- $\Delta$ C/EBP $\beta$ -Fluc, in which the three potential C/EBP $\beta$  BS were inactivated by mutations, was used as the internal control plasmid in these experiments to mitigate any potentially confounding effect that overproduced C/EBP $\beta$  may exert on Fluc expression from LCR-PT. As anticipated, LAP increased the activity of LCR-PT by 5.9-, 12.4- and 12.4-fold in C33A, HeLa and U2OS cells, respectively (Fig. 3.6A). As for LIP, it repressed LCR-PT by 3.1-fold in HeLa cells but not in C33A and U2OS cells (Fig. 3.4A), likely because the latter two cell lines express low levels of endogenous C/EBP $\beta$  at a low LAP/LIP ratio (Fig. 3.6B). Indeed, immunoblot analysis revealed that while the levels of transiently overexpressed LAP and LIP were comparable in all cell lines, those of the endogenous proteins were lower in C33A and U2OS cells compared to HeLa cells (Fig. 3.6B, LAP/LIP ratio of approximately 0.5 in C33A and U2OS versus 2 in HeLa). Collectively, these results provided further evidence for the regulation of the HPV33 LCR by C/EBP $\beta$  isoforms in cervical carcinoma cell lines as well as in U2OS osteosarcoma cells which support episomal replication of the HPV genome.

### **Contribution of BS1 and BS2 to the activity of the LCR under repressive and activating conditions**

Having determined that the HPV33 LCR can be activated and repressed by LAP and LIP, respectively, we wished to examine the effect of inactivating mutations in BS1 and

BS2 under these conditions. Results presented in Fig. 3.7 show the effect of the mBS1 and mBS2 mutations, either alone or in combination (mBS1/2), on the activity of the LCR in C33A, HeLa and U2OS when LAP and LIP are overexpressed. In C33A cells, the inactivation of BS1 or BS2 resulted in reduced transactivation activity of overexpressed LAP (Figure 3.7A) with 3.9- and 4.1-fold stimulation compared to 6.1fold seen with LCR-PT, respectively. This effect was further accentuated when mBS1 was introduced in combination with mBS2 (fold-activation of 3.1). The overexpression of LIP did not markedly alter the effect of the individual mBS1 and mBS2 mutations relative to LCR-PT in this cell line (Fig. 3.7B).

In HeLa cells, the mutation of BS1 also resulted in decreased LAP activity from a 14.8-fold to 5-fold (Fig. 3.7C). Interestingly, the capacity of exogenous LAP to transactivate the LCR was increased to 43.0-fold in a BS2 mutant, indicating that BS1 plays a stimulatory role when LAP is abundant and BS2 is mutated. However, when both sites are mutated, the transactivation activity of overexpressed LAP was reduced with 12.7-fold stimulation. Finally, we noticed that the repressive effect of LIP was also reduced in BS1 and BS2 mutants from 2.6-fold to 1.5- and 1.2-fold, respectively (Fig. 3.7D). The ability of LIP to repress the LCR was completely abrogated in the double mutant (mBS1/2).

Finally, the function of overexpressed LAP on the activity of the LCRs in U2OS cells was similar as seen in HeLa cells (Fig. 3.7E). Mainly, the LAP-induced transactivation activity was reduced in mBS1 (9.7-fold) but was increased in the BS2 and BS1/2 mutants (15.7 and 15.4-fold stimulation, respectively) compared to LCR-PT (11.8-fold). The overexpression of LIP did not significantly alter the effect of the individual mBS1 and mBS2 mutations relative to LCR-PT in this cell line (Fig. 3.7F). Collectively, the results presented above confirmed that the two C/EBP $\beta$  binding sites have opposite activities, with BS1 being a repressor and BS2 an activator. In addition, they showed that the influence (i.e. potency) of each site can be altered by modulating the intracellular levels of the LAP and LIP isoforms.

### 3.3.5 Discussion

In this study, we characterized the rare HPV33 LCR10 variant, a member of the B-lineage that was previously detected in consecutive cervical samples from a HIV-positive woman [21, 22]. LCR10 exhibited a 2-fold higher transcriptional activity than LCR-PT in C33A, HeLa and U2OS cells (Fig. 3.1). Our results indicate that the enhanced activity of LCR10 is caused primarily by the single T7791C variation which abolishes a negatively-acting C/EBP $\beta$  BS (Fig. 3.2). The discovery that C/EBP $\beta$  regulates the HPV33 LCR may help explain, at least in part, why this LCR is less active in C33A and U2OS cells than in HeLa cells. Indeed, the observation that C/EBP $\beta$  is expressed only in small amounts and at a lower LAP/LIP ratio in C33A and U2OS cells provides a partial explanation for the weaker activity of LCR-PT in these cells and its inability to be further repressed by overexpression of LIP. The previous report that only substoichiometric amounts of LIP are sufficient to antagonize the levels of transcriptional activation brought about by LAP lends further support to this possibility [10]. Finally, it is possible, if not likely, that other factors impinge on the activity of C/EBP $\beta$  in C33A, HeLa and U2OS cells, such as phosphorylation which modulates its DNA-binding activity [31, 32]. Further studies will be needed to examine this possibility.

Our previous analysis of several A2-sublineage variants indicated that the combination of four-weakly-acting variations is needed to enhance the activity of the HPV33 LCR by almost 2-fold [23]. In this study, we found that the single T7791C variation was sufficient to increase transcription from the LCR by same extent. Naturally-occurring variations that disrupt important TF BS in various LCRs from different HPV types have been reported in the literature. For HPV16, these include mutations and deletions in several YY1 BS that stimulate transcription from variant LCRs [33-35]. Importantly, it has been shown in cell culture models that a gain in LCR activity can increase the transforming and replication activities of HPV16 variants [34, 36-41]. Although it is currently unknown if the higher activity of LCR10 has any effect on the oncogenic potential and/or replication capacity of the HPV33 episome, we note that this variant LCR was detected in five specimens (2 cervicovaginal lavages and 3 tampons) collected over a period of 288 days, indicative of a persistent infection. However, no squamous intraepithelial lesion was ever detected on cytology smears and the infection by

this variant was eventually cleared. Others have also recently described C/EBP $\beta$  BS variations in the LCR of HPV16, HPV31 and HPV58 [42]. However, these studies neither investigated the functional impact of these variations nor their association with the development of CC. Consequently, it is still unclear if C/EBP $\beta$  BS variations can modulate HPV carcinogenesis.

A major finding of this study is the critical role of C/EBP $\beta$  for regulation of the HPV33 LCR through two distinct BS with antagonistic activity (Fig. 3.5). This ability of C/EBP $\beta$  to function either as an activator or repressor of HPV gene expression has been noted previously for HPV types other than 33. For example, C/EBP $\beta$  has been implicated in the repression of the HPV11 and HPV16 LCRs and in the activation of the HPV18 LCR, through its interaction with YY1 [12, 16-19]. C/EBP $\beta$  has also been suggested to repress transcription from the HPV16 LCR by competition with the transcriptional activator proteins AP-1 and NF-1 whose BS overlaps that of C/EBP $\beta$  in these viruses [16]. In this respect, we note that a NF-1 consensus sequence (5'-TTGGC(N)<sub>6</sub>CC-3', [43]) surrounds the C/EBP $\beta$  BS1 in LCR-PT (5'-TTGGCTTACACAA-3', NF-1 site underlined, C/EBP $\beta$  site italicized) suggesting that C/EBP $\beta$  may exert its repressive effect by competing with NF-1. Furthermore, it is particularly intriguing that a putative AP-1 BS (consensus 5'-TGA(C/G)TCA-3', [44]) is created by T7791C in LCR10 (5'-TGGCTCA-3', T7791C underlined). This observation raises the possibility that the T7791C variation not only prevents C/EBP $\beta$ -binding but that it may also allow AP-1 to further transactivate the HPV33 LCR.

In addition to BS1, we also identified a second C/EBP $\beta$  site (BS2) which, in contrast to BS1, activates the HPV33 LCR in HeLa cells and, to a lesser extent, in U2OS cells (Fig. 3.5). Interestingly, C/EBP $\beta$  BS2 (nucleotides 7700-7709) lies within a region of the HPV33 LCR (nucleotides 7587-7729) that shows sequence similarity with the minimal keratinocyte enhancer region first identified in HPV16 and shown to be essential for activation of the early promoter [45, 46]. Footprinting analysis of this enhancer element in the LCRs of HPV16 and HPV18 revealed a protected region of approximately 27 (7665-7691) and 35 (7641-7675) base pairs which has been proposed to contain binding sites for the yet unidentified PEF-1 (papillomavirus enhancer binding factor-1) and KRF-1

(keratinocyte-specific factor 1) factors in the LCR of HPV16 and HPV18, respectively [46, 47]. A similar site is also found in the LCR of HPV33 (7676-7710). The conservation of the protected region in these LCRs indicates that the same TF may actually bind to this site (Fig. 3.8). Further examination of this region in HPV18 shows the presence of a putative C/EBP $\beta$  that falls within the proposed KRF-1 binding region, suggesting that C/EBP $\beta$  may either be the transcription factor that binds to this region or may act together with KRF-1 to induce the activation of the LCR. Interestingly, the HPV33 BS2 is also found within the corresponding KRF-1 site whereas no apparent C/EBP $\beta$  site is found in the PEF-1 protected region in HPV16. It would thus be interesting to investigate the role of this conserved region in other HPV types in relation to C/EBP $\beta$ .

In conclusion, we have described what we believe to be the first instance of a single naturally-occurring variation, T7791C, which significantly enhances the transcriptional activity of the HPV33 LCR by disrupting an inhibitory C/EBP $\beta$  BS. A second C/EBP $\beta$  element was also identified and shown to be important for the activation of the LCR. Taken together, our findings suggest that C/EBP $\beta$  can both activate and repress transcription of the HPV33 early genes depending on the context of its binding site and the relative expression of the LAP and LIP isoforms.

### **3.3.6 Bibliography**

1. Bosch FX, Burchell AN, Schiffman M, et al. Epidemiology and natural history of human papillomavirus infections and type-specific implications in cervical neoplasia. *Vaccine* **2008**; 26 Suppl 10:K1-16.
2. Walboomers JM, Jacobs MV, Manos MM, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *The Journal of pathology* **1999**; 189:12-9.
3. Goodwin EC, DiMaio D. Repression of human papillomavirus oncogenes in HeLa cervical carcinoma cells causes the orderly reactivation of dormant tumor suppressor pathways. *Proceedings of the National Academy of Sciences of the United States of America* **2000**; 97:12513-8.
4. O'Connor MJ, Chan SY, Bernard HU. Transcription factor binding sites in the long control regions of genital HPVs, part III-A. In: Myers G BHU, Delius H, Baker C, Icenogle



- J, Halpern A, Wheeler C, ed. Human papillomaviruses 1995 compendium. Los Alamos, N.M. USA: Los Alamos National Laboratory, **1995**;pp. 21-40.
5. Bernard HU. Regulatory elements in the viral genome. *Virology* **2013**; 445:197-204.
  6. Nerlov C. The C/EBP family of transcription factors: a paradigm for interaction between gene expression and proliferation control. *Trends in cell biology* **2007**; 17:318-24.
  7. Darlington GJ, Ross SE, MacDougald OA. The role of C/EBP genes in adipocyte differentiation. *The Journal of biological chemistry* **1998**; 273:30057-60.
  8. Johnson PF. Molecular stop signs: regulation of cell-cycle arrest by C/EBP transcription factors. *Journal of cell science* **2005**; 118:2545-55.
  9. Liu Y, Nonnemacher MR, Wigdahl B. CCAAT/enhancer-binding proteins and the pathogenesis of retrovirus infection. *Future microbiology* **2009**; 4:299-321.
  10. Descombes P, Schibler U. A liver-enriched transcriptional activator protein, LAP, and a transcriptional inhibitory protein, LIP, are translated from the same mRNA. *Cell* **1991**; 67:569-79.
  11. Tsukada J, Yoshida Y, Kominato Y, Auron PE. The CCAAT/enhancer (C/EBP) family of basic-leucine zipper (bZIP) transcription factors is a multifaceted highly-regulated system for gene regulation. *Cytokine* **2011**; 54:6-19.
  12. Struyk L, van der Meijden E, Minnaar R, Fontaine V, Meijer I, ter Schegget J. Transcriptional regulation of human papillomavirus type 16 LCR by different C/EBPbeta isoforms. *Molecular carcinogenesis* **2000**; 28:42-50.
  13. Gunasekharan V, Hache G, Laimins L. Differentiation-dependent changes in levels of C/EBPbeta repressors and activators regulate human papillomavirus type 31 late gene expression. *Journal of virology* **2012**; 86:5393-8.
  14. Bauknecht T, Jundt F, Herr I, et al. A switch region determines the cell type-specific positive or negative action of YY1 on the activity of the human papillomavirus type 18 promoter. *Journal of virology* **1995**; 69:1-12.
  15. Bauknecht T, See RH, Shi Y. A novel C/EBP beta-YY1 complex controls the cell-type-specific activity of the human papillomavirus type 18 upstream regulatory region. *Journal of virology* **1996**; 70:7695-705.
  16. Kyo S, Inoue M, Nishio Y, et al. NF-IL6 represses early gene expression of human papillomavirus type 16 through binding to the noncoding region. *Journal of virology* **1993**; 67:1058-66.

17. Wang H, Liu K, Yuan F, Berdichevsky L, Taichman LB, Auborn K. C/EBPbeta is a negative regulator of human papillomavirus type 11 in keratinocytes. *Journal of virology* **1996**; 70:4839-44.
18. Ralph WM, Jr., Liu K, Auborn KJ. CCAAT/enhancer-binding protein beta represses human papillomavirus 11 upstream regulatory region expression through a promoter-proximal YY1-binding site. *The Journal of general virology* **2006**; 87:51-9.
19. Bauknecht T, Shi Y. Overexpression of C/EBPbeta represses human papillomavirus type 18 upstream regulatory region activity in HeLa cells by interfering with the binding of TATA-binding protein. *Journal of virology* **1998**; 72:2113-24.
20. Zhao W, Chow LT, Broker TR. A distal element in the HPV-11 upstream regulatory region contributes to promoter repression in basal keratinocytes in squamous epithelium. *Virology* **1999**; 253:219-29.
21. Gagnon S, Hankins C, Tremblay C, Forest P, Pourreaux K, Coutlee F. Viral polymorphism in human papillomavirus types 33 and 35 and persistent and transient infection in the genital tract of women. *The Journal of infectious diseases* **2004**; 190:1575-85.
22. Khouadri S, Villa LL, Gagnon S, et al. Human papillomavirus type 33 polymorphisms and high-grade squamous intraepithelial lesions of the uterine cervix. *The Journal of infectious diseases* **2006**; 194:886-94.
23. Alvarez J, Coutlee F, Archambault J. The Combination of Specific Nucleotide Variations Modulates the Transcriptional Activity of the Human Papillomavirus Type 33 Long Control Region (unpublished), **2016**.
24. Messeguer X, Escudero R, Farre D, Nunez O, Martinez J, Alba MM. PROMO: detection of known transcription regulatory elements using species-tailored searches. *Bioinformatics (Oxford, England)* **2002**; 18:333-4.
25. Farre D, Roset R, Huerta M, et al. Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN. *Nucleic acids research* **2003**; 31:3651-3.
26. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nature methods* **2012**; 9:671-5.
27. Titolo S, Brault K, Majewski J, White PW, Archambault J. Characterization of the minimal DNA binding domain of the human papillomavirus e1 helicase: fluorescence

anisotropy studies and characterization of a dimerization-defective mutant protein. *Journal of virology* **2003**; 77:5178-91.

28. Fradet-Turcotte A, Vincent C, Joubert S, Bullock PA, Archambault J. Quantitative analysis of the binding of simian virus 40 large T antigen to DNA. *Journal of virology* **2007**; 81:9162-74.

29. Williams SC, Baer M, Dillner AJ, Johnson PF. CRP2 (C/EBP beta) contains a bipartite regulatory domain that controls transcriptional activation, DNA binding and cell specificity. *The EMBO journal* **1995**; 14:3170-83.

30. Geimanen J, Isok-Paas H, Pipitch R, et al. Development of a cellular assay system to study the genome replication of high- and low-risk mucosal and cutaneous human papillomaviruses. *Journal of virology* **2011**; 85:3315-29.

31. Osada S, Yamamoto H, Nishihara T, Imagawa M. DNA binding specificity of the CCAAT/enhancer-binding protein transcription factor family. *The Journal of biological chemistry* **1996**; 271:3891-6.

32. Trautwein C, van der Geer P, Karin M, Hunter T, Chojkier M. Protein kinase A and C site-specific phosphorylations of LAP (NF-IL6) modulate its binding affinity to DNA recognition elements. *The Journal of clinical investigation* **1994**; 93:2554-61.

33. Dong XP, Stubenrauch F, Beyer-Finkler E, Pfister H. Prevalence of deletions of YY1-binding sites in episomal HPV 16 DNA from cervical cancers. *International journal of cancer Journal international du cancer* **1994**; 58:803-8.

34. Schmidt M, Kedzia W, Gozdziacka-Jozefiak A. Intratype HPV16 sequence variation within LCR of isolates from asymptomatic carriers and cervical cancers. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology* **2001**; 23:65-77.

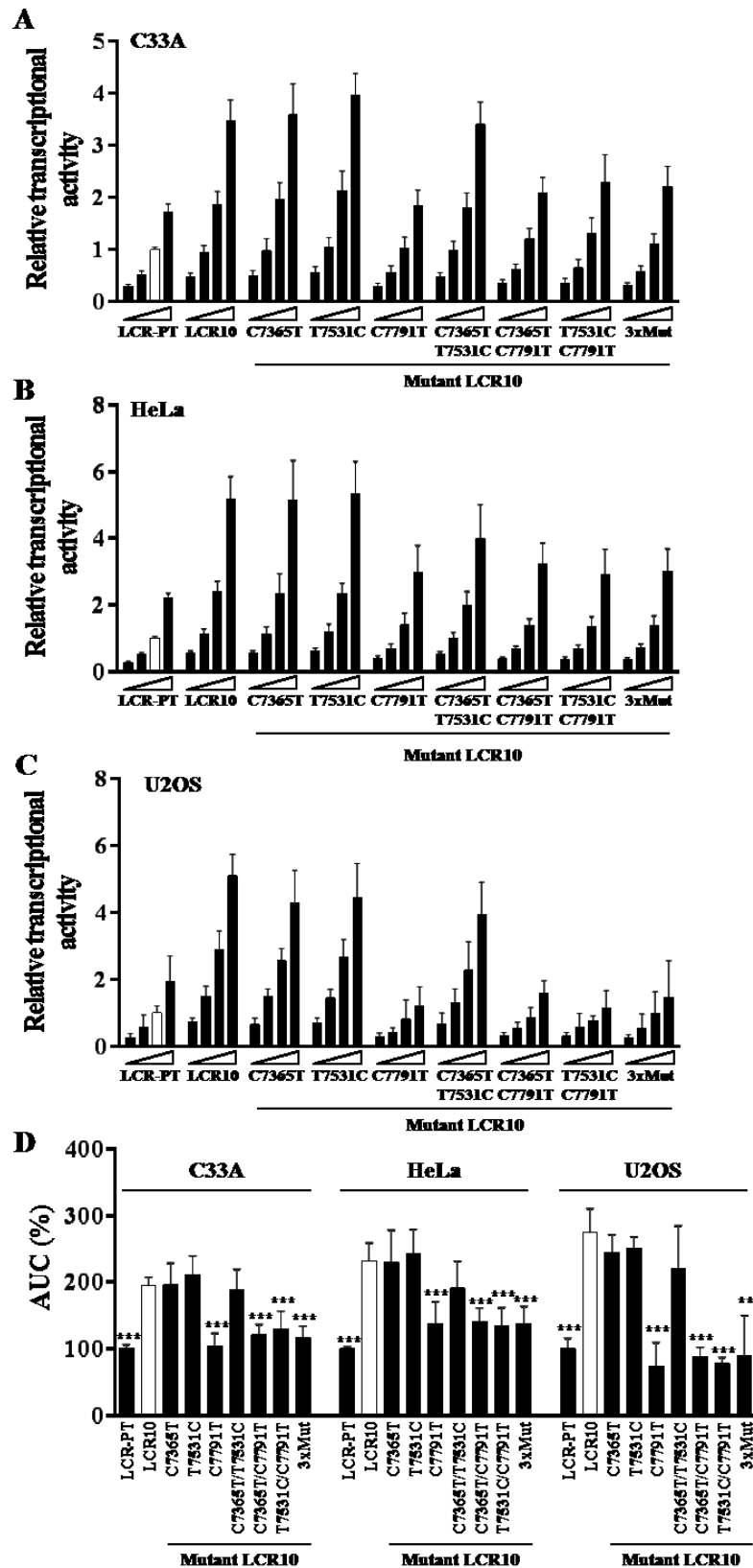
35. Kozuka T, Aoki Y, Nakagawa K, et al. Enhancer-promoter activity of human papillomavirus type 16 long control regions isolated from cell lines SiHa and CaSki and cervical cancer biopsies. *Japanese journal of cancer research : Gann* **2000**; 91:271-9.

36. Xi LF, Koutsky LA, Hildesheim A, et al. Risk for high-grade cervical intraepithelial neoplasia associated with variants of human papillomavirus types 16 and 18. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **2007**; 16:4-10.

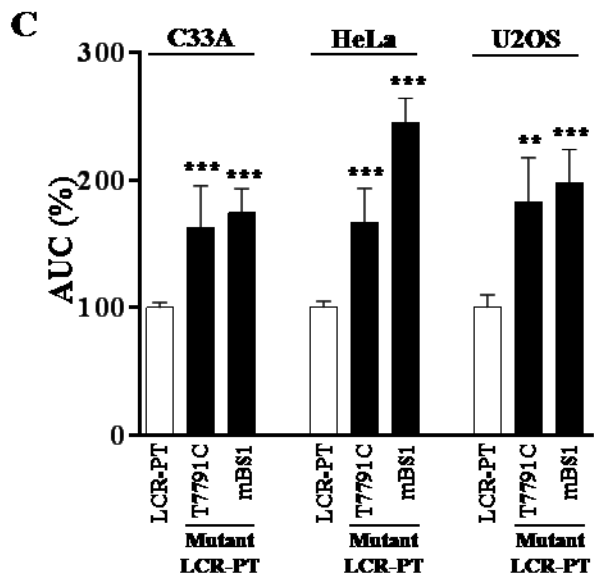
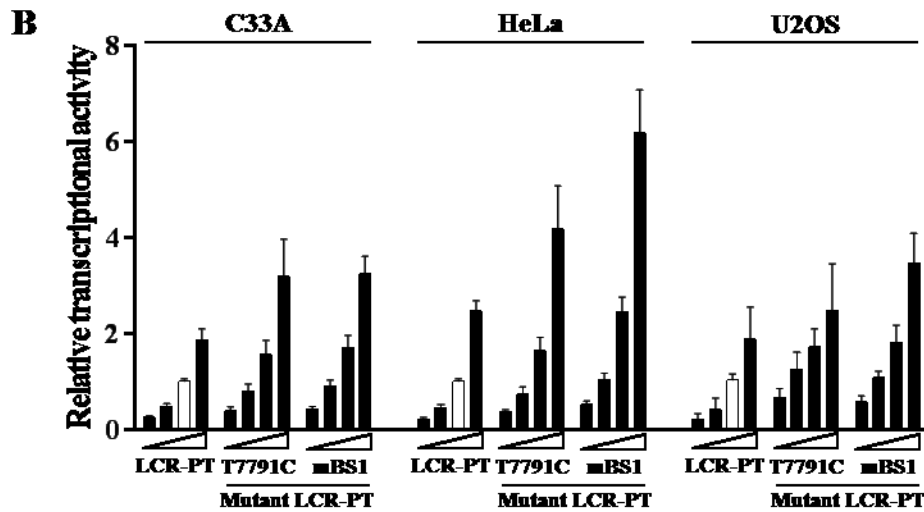
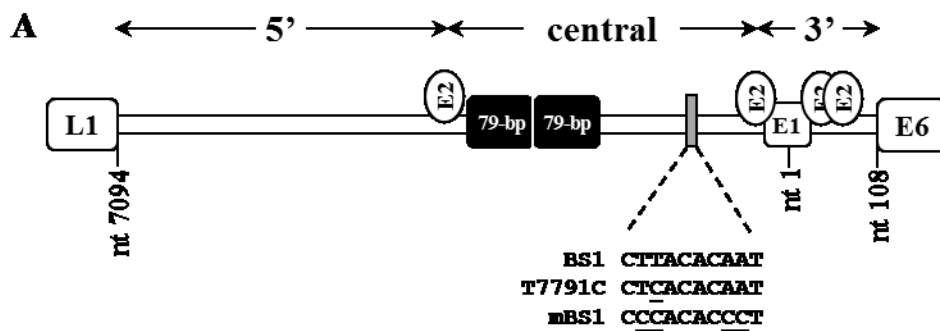
37. Sichero L, Ferreira S, Trottier H, et al. High grade cervical lesions are caused preferentially by non-European variants of HPVs 16 and 18. *International journal of cancer Journal international du cancer* **2007**; 120:1763-8.
38. Sichero L, Sobrinho JS, Villa LL. Oncogenic potential diverge among human papillomavirus type 16 natural variants. *Virology* **2012**; 432:127-32.
39. Lace MJ, Isacson C, Anson JR, et al. Upstream regulatory region alterations found in human papillomavirus type 16 (HPV-16) isolates from cervical carcinomas increase transcription, ori function, and HPV immortalization capacity in culture. *Journal of virology* **2009**; 83:7457-66.
40. Hubert WG. Variant upstream regulatory region sequences differentially regulate human papillomavirus type 16 DNA replication throughout the viral life cycle. *Journal of virology* **2005**; 79:5914-22.
41. May M, Dong XP, Beyer-Finkler E, Stubenrauch F, Fuchs PG, Pfister H. The E6/E7 promoter of extrachromosomal HPV16 DNA in cervical cancers escapes from cellular repression by mutation of target sequences for YY1. *The EMBO journal* **1994**; 13:1460-6.
42. Gurgel AP, Chagas BS, do Amaral CM, et al. Prevalence of Human Papillomavirus Variants and Genetic Diversity in the L1 Gene and Long Control Region of HPV16, HPV31, and HPV58 Found in North-East Brazil. *BioMed research international* **2015**; 2015:130828.
43. Misawa H, Yamaguchi M. Involvement of nuclear factor-1 (NF1) binding motif in the regucalcin gene expression of rat kidney cortex: the expression is suppressed by cisplatin administration. *Molecular and cellular biochemistry* **2001**; 219:29-37.
44. Zhou H, Zarubin T, Ji Z, et al. Frequency and distribution of AP-1 sites in the human genome. *DNA research : an international journal for rapid publication of reports on genes and genomes* **2005**; 12:139-50.
45. Sibbet GJ, Cuthill S, Campo MS. The enhancer in the long control region of human papillomavirus type 16 is up-regulated by PEF-1 and down-regulated by Oct-1. *Journal of virology* **1995**; 69:4006-11.
46. Cuthill S, Sibbet GJ, Campo MS. Characterization of a nuclear factor, papilloma enhancer binding factor-1, that binds the long control region of human papillomavirus type 16 and contributes to enhancer activity. *Molecular carcinogenesis* **1993**; 8:96-104.

47. Mack DH, Laimins LA. A keratinocyte-specific transcription factor, KRF-1, interacts with AP-1 to activate expression of human papillomavirus type 18 in squamous epithelial cells. *Proceedings of the National Academy of Sciences of the United States of America* **1991**; 88:9102-6.

### 3.3.7 Figures and Tables

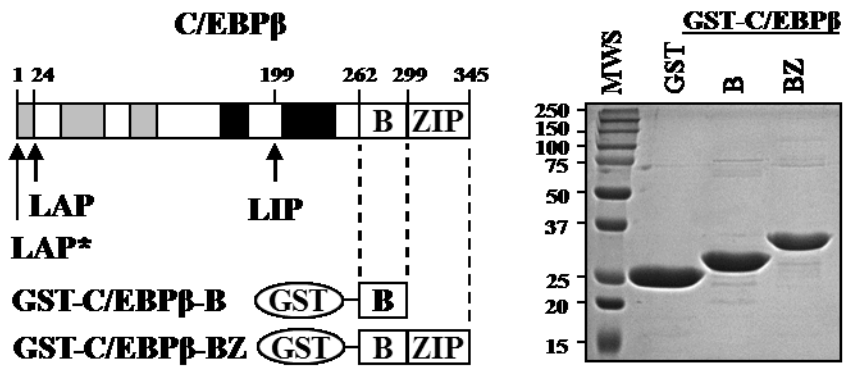
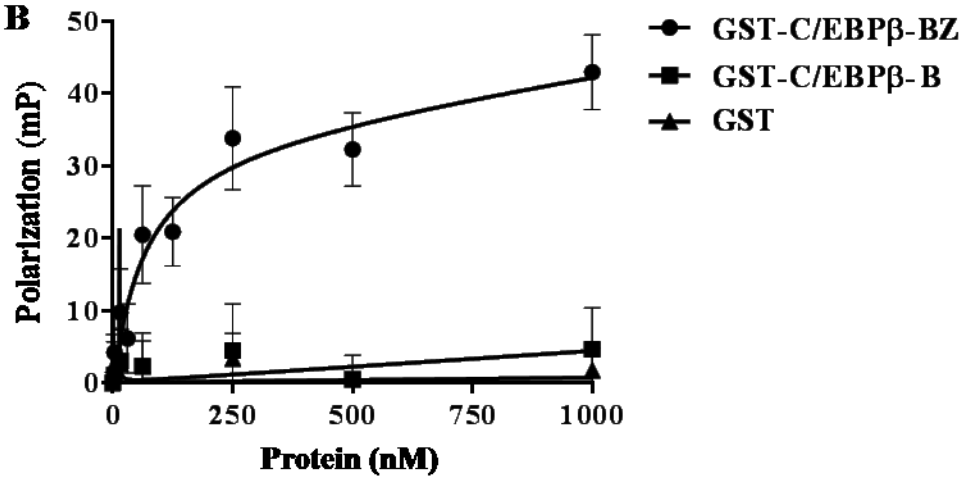
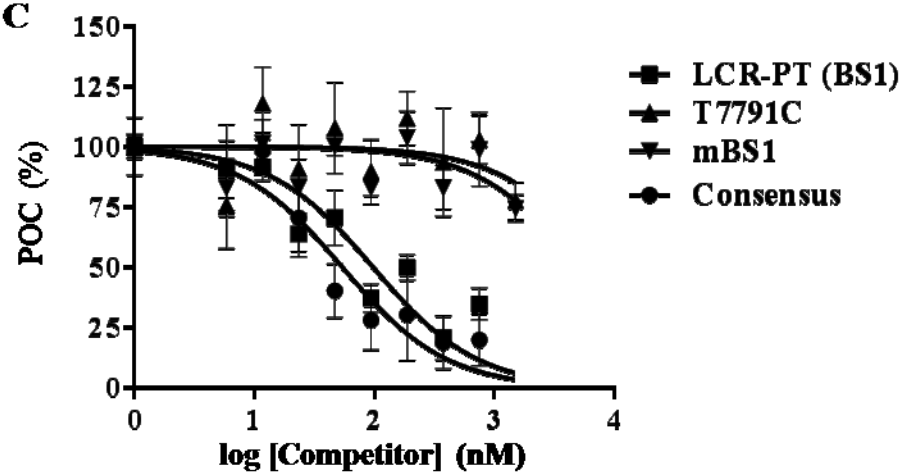


**Figure 3.1. The enhanced transcriptional activity of the HPV33 LCR10 variant is attributable primarily to the T7791C variation.** The contribution of the three LCR10-specific variations, C7365T, T7531C and T7791C, to the transcriptional activity of LCR10 was tested by reverting each of them, individually and in combination, back to the sequence of the prototype (LCR-PT). 3xMut designates the mutant LCR10 in which the C7365T, T7531C and T7791C variations were reverted to the prototypical sequence. Bar graph representing the transcriptional activity of the indicated LCR10 mutant derivatives relative to that of the prototype (LCR-PT), in C33A (panel A), HeLa (panel B) and U2OS (panel C) cells. The activity of each LCR was measured in luciferase reporter-gene assays using increasing quantities of LCR-Rluc plasmid (25, 50, 100 and 200 ng). The LCR-PT-Fluc plasmid (25 ng) was co-transfected as an internal control. Each bar represents the mean value obtained from a minimum of 4 independent experiments performed in duplicates. All values are presented relative to the level of transcriptional activity obtained with 100 ng of LCR-PT, which was used as the reference (white bar) and assigned a value of 1.0. Standard deviations are indicated by error bars. D) Integrated measure of LCR transcriptional activity using the area under the curve (AUC). AUC values were calculated from the dose-response curves presented in panels A-C and used as an integrated measure of LCR activity in C33A, HeLa and U2OS cells, as indicated. The AUC value of each LCR is reported as a percentage of the AUC value calculated for LCR-PT, which was set as 100. Statistical analysis of AUC values was performed using a one-way ANOVA followed by Dunnett's post-hoc analysis to compare the transcriptional activity of LCR10 (white bar) with that of each LCR10 mutant derivatives and of the prototype. \* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$ .





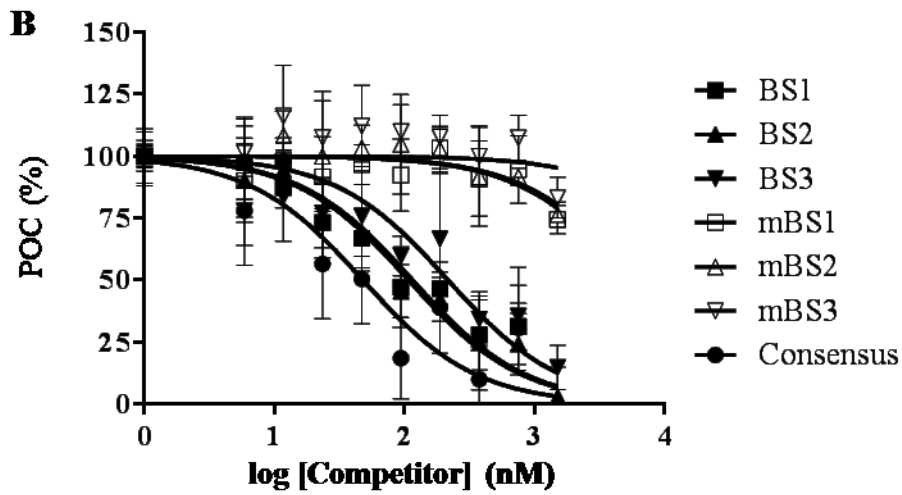
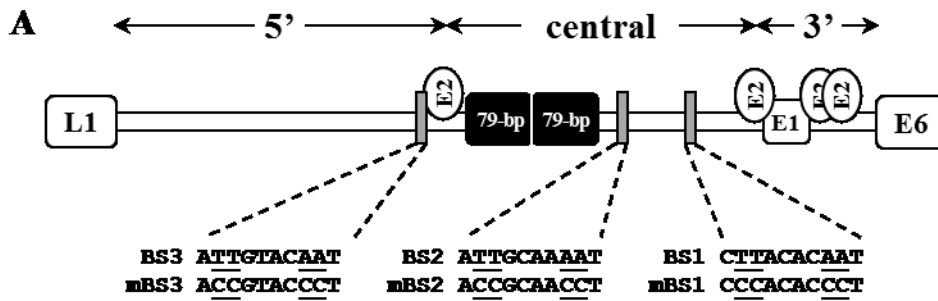
**Figure 3.2. The T7791C variation stimulates the transcriptional activity of the HPV33 LCR to a similar extent as inactivation of the C/EBP $\beta$  binding site.** A) Schematic representation of the HPV33 LCR. The 924-bp LCR is represented by a white line flanked by the L1 and E6 open reading frames at its 5' and 3' extremities, respectively. The locations of 4 binding sites for the E2 transcriptional regulator and binding sites for the viral E1 helicase are indicated. Black boxes represent the perfect 79-bp duplication present in the central portion of the LCR of the prototype. The grey box indicates the location of the putative C/EBP $\beta$  binding site (nucleotide 7789-7798) encompassing nucleotide 7791, which is changed from a T to a C in LCR10. The sequence of this site in LCR-PT (BS1), in the variant LCR10 (T7791C) or when inactivated by a quadruple mutation (mBS1) are indicated, with the nucleotides that differ from the prototype underlined. Nucleotides (nt) are numbered according to the HPV33 reference genome (GenBank accession M12732.1). Nucleotides (nt) are numbered according to the HPV33 reference genome (GenBank accession M12732.1). B) Transcriptional activities of C/EBP $\beta$  binding site mutant LCRs. The T7791C variation and quadruple inactivating mutation (mBS1) were introduced separately into LCR-PT and the transcriptional activities of the resulting mutant LCRs were measured using increasing quantities of LCR-Rluc plasmid, in either C33A (left panel), HeLa (center panel) and U2OS (right panel) cells, exactly as described in the legend of Fig. 3.1. Each bar represents the mean value obtained from 4 independent experiments performed in triplicates. An integrated measure of the transcriptional activity of each LCR (C) was obtained by calculating the area under each dose-response curve (AUC), as described in the legend of Fig. 3.1.C.

**A****B****C**

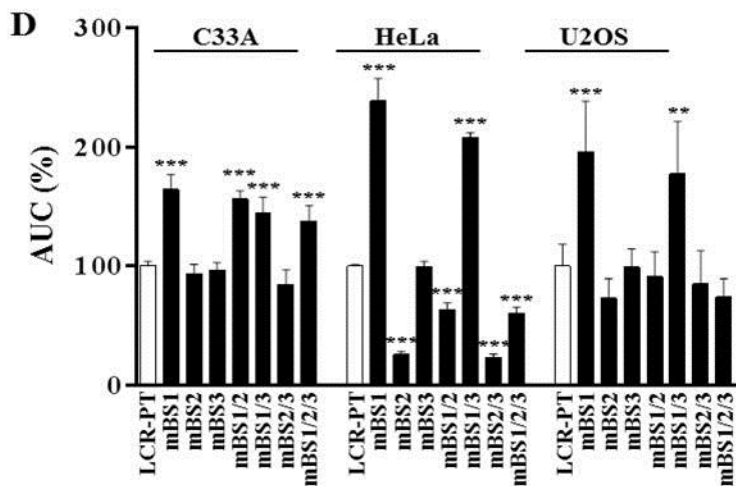
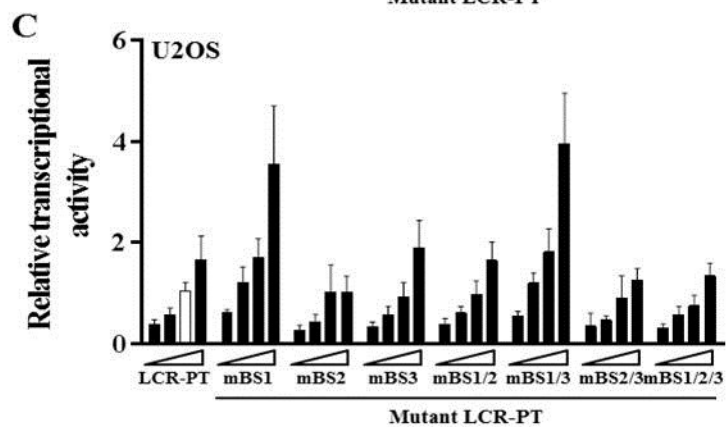
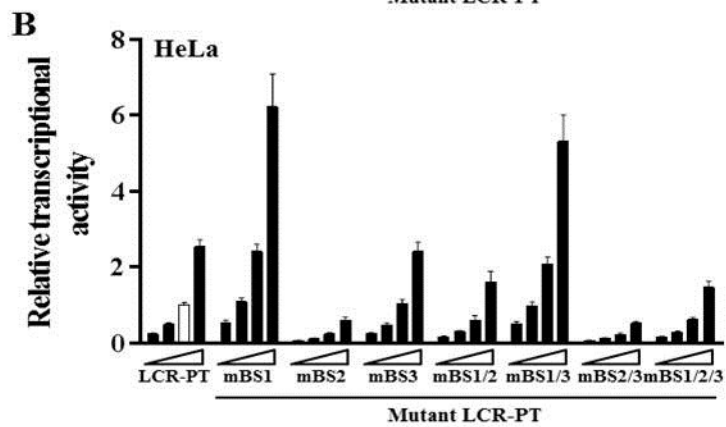
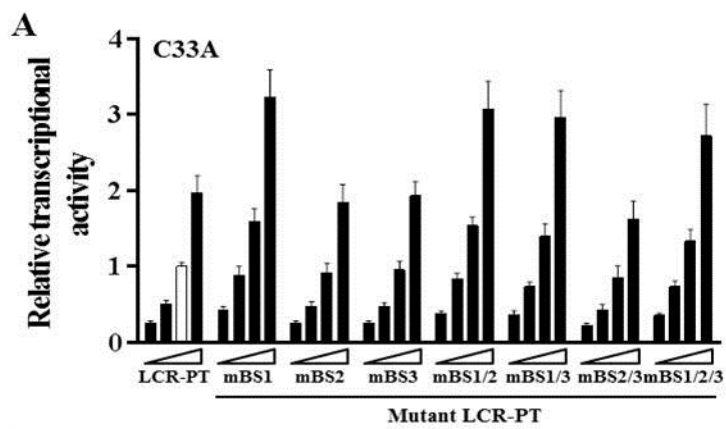
**Figure 3.3. The T7791C variation reduces the affinity of C/EBP $\beta$  for its binding site.**

A) Schematic representation of C/EBP $\beta$  and of the two GST-fusion proteins used in DNA-binding assays. C/EBP $\beta$  is produced as three isoforms: two which activate transcription, LAP\* (Liver-enriched activator protein full length, amino acids 1-345) and LAP (Liver-enriched activator protein, amino acids 24-345), and one which is inhibitory, LIP (Liver-enriched inhibitory protein amino acids 199-345). Grey and black boxes indicate the location of the C/EBP $\beta$  activation and repression domains, respectively. All isoforms contain a basic region (B) and a leucine zipper domain (Z) that are both necessary for DNA-binding. C/EBP $\beta$  fragments encompassing the B and Z regions, or only the B region as a negative control, were bacterially expressed and purified as fusion proteins with GST (as illustrated). The right panel shows a coomassie-stained 12% SDS-PAGE analysis of the purified GST (28 kDa), GST-C/EBP $\beta$ -B (30 kDa) and GST-C/EBP $\beta$ -BZ (36 kDa) proteins (3  $\mu$ g of each protein were loaded on the gel). The position and size (in kDa) of molecular weight standards (MWS) are shown on the left of the gel. B) DNA-binding activity of the purified proteins. The ability of GST, GST-C/EBP $\beta$ -B and GST-C/EBP $\beta$ -BZ to bind to a fluorescein-labeled DNA probe containing the prototype C/EBP $\beta$  binding site BS1 (LCR-PT) was determined using a fluorescence polarization assay performed with increasing concentrations of each protein (3.9, 7.8, 15.6, 31.25, 62.5, 125, 250, 500 and 1000 nM) and 15 nM of probe. Each data point is the average of duplicates values measured in three independent experiments (n=6). Error bars represent standard deviations. C) Fluorescence polarization competition assay. Unlabeled duplex oligonucleotides (BS1, T7791C, mBS1 and Consensus) were used as competitors and tested in increasing amounts (0, 5.9, 11.7, 23.4, 46.9, 93.8, 187.5, 375, 750 and 1500 nM) for their ability to inhibit the binding of GST-C/EBP $\beta$ -BZ (125 nM) to a fluorescein-labeled BS1 probe (15 nM). The T7791C oligonucleotide is identical in sequence to the BS1 oligonucleotide but contains the T7791C variation. The Consensus oligonucleotide contains a high-affinity C/EBP $\beta$  site that is used as a positive control. The mBS1 oligonucleotide corresponds to a defective C/EBP $\beta$  site that was inactivated by a quadruple mutation and is used as a negative control. Competitor concentration is indicated as the log<sub>10</sub> value. Fluorescence polarization levels were normalized to the value measured in absence of competitor DNA, which was set at 100%, and are reported as a percent of control (POC). Each data point is the average of 3

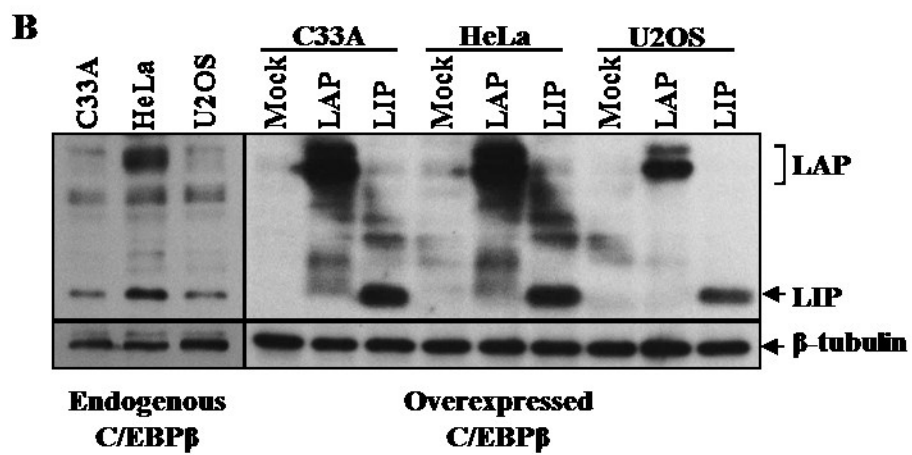
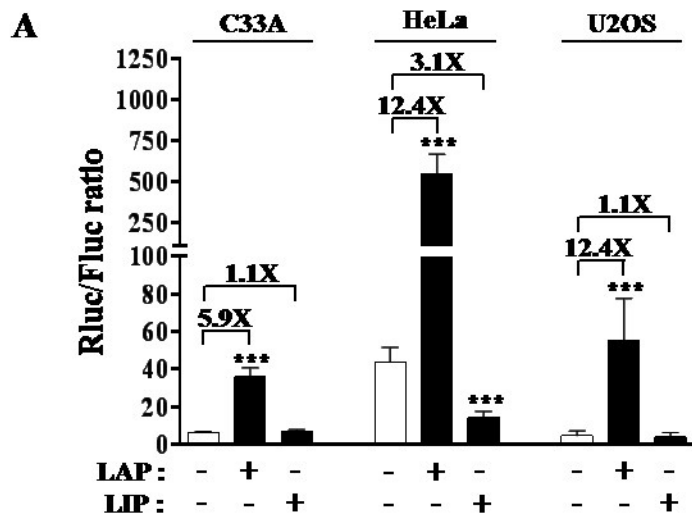
independent experiments, each performed in duplicates (n=6). Error bars represent the standard deviation.



**Figure 3.4. Identification of two additional binding sites in the LCR of HPV33 with different affinities for C/EBP $\beta$ .** The grey boxes show the location of the three predicted C/EBP $\beta$  binding sites termed BS3 (nucleotide 7471-7480), BS2 (nucleotide 7700-7709) and BS1 (nucleotide 7789-7798). The sequences of the three sites are shown below. The bases that are underlined show the position of the nucleotides that were mutated to cytosines to inactivate the C/EBP $\beta$  sites (mBS3, mBS2 and mBS1). Nucleotides (nt) are numbered according to the HPV33 reference genome (GenBank accession M12732.1). B) Competitive fluorescence polarization assay for the binding of GST-C/EBP $\beta$ -BZ. Increasing amount (0, 5.9, 11.7, 23.4, 46.9, 93.8, 187.5, 375, 750 and 1500 nM) of unlabeled oligonucleotides (Consensus, BS1, BS2, BS3 and of the C/EBP $\beta$  binding site quadruple mutants mBS1, mBS2 and mBS3) were used to compete against a fluorescein-labeled BS1 probe (15 nM) for the binding of GST-C/EBP $\beta$ -BZ (125 nM). Competitor concentration is indicated as the log<sub>10</sub> value. Polarization levels were normalized to the value at 0 ng of competitor which was set at 100% and are reported as a percent of control (POC). Each data point is the average of duplicates of 3 independent experiments (n=6). Error bars represent the standard deviation.

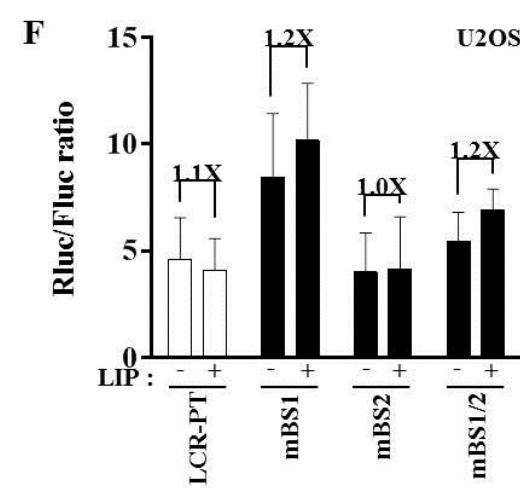
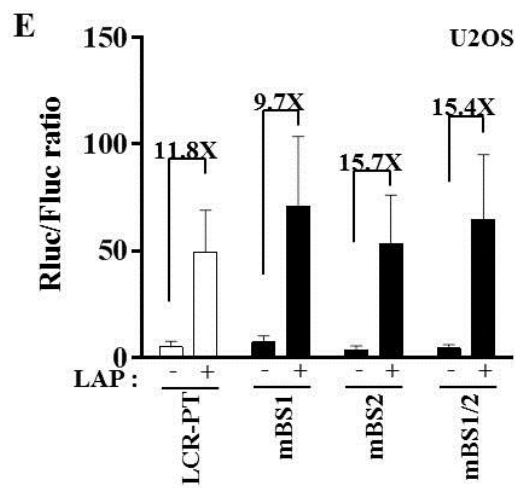
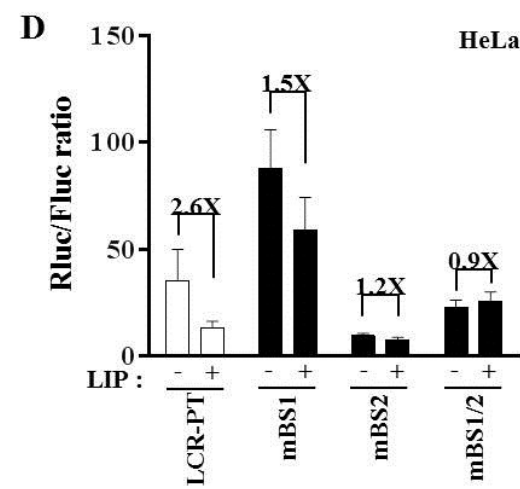
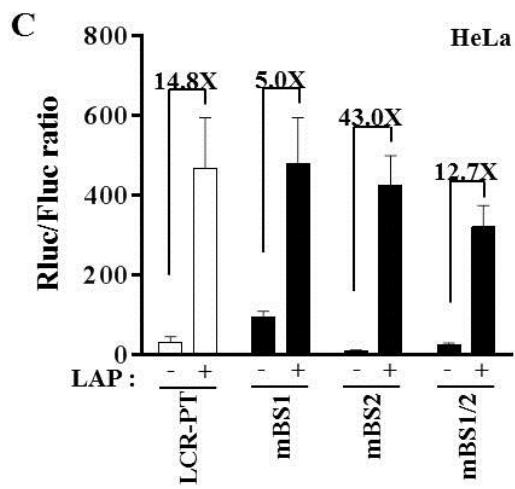
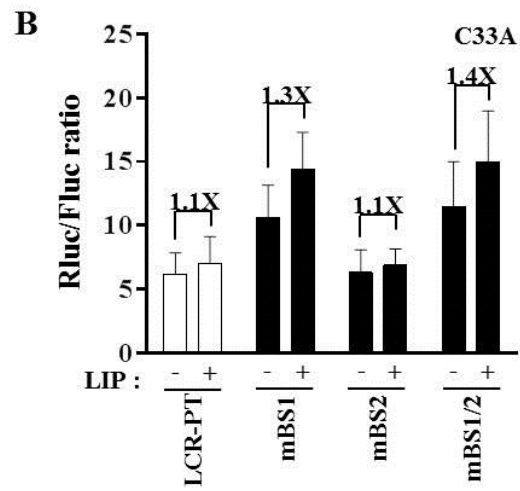
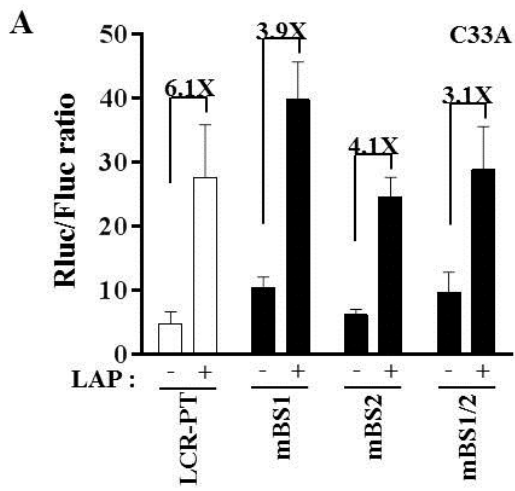


**Figure 3.5. Mutational analysis of the three C/EBP $\beta$  binding sites present in the HPV33 prototypical LCR.** The contribution of the three C/EBP $\beta$  binding sites to the transcriptional activity of LCR-PT was assessed by inactivating each site (with a quadruple mutation) either individually or in combination. Bar graph representing the transcriptional activity of the indicated LCR-PT mutant derivatives, in C33A (A), HeLa (B) and U2OS (C) cells. The activity of each LCR was measured in luciferase reporter-gene assays using increasing quantities of LCR-Rluc plasmid (25, 50, 100 and 200 ng). The LCR- $\Delta$ C/EBP $\beta$ -Fluc (25 ng) was co-transfected as an internal control. Each bar represents the mean values of duplicates from at least 4 independent assays. All values are presented relative to the level of transcriptional activity obtained with 100 ng of LCR-PT, which was used as the reference (white bar) and assigned a value of 1.0. Standard deviations are indicated by error bars. D) Integrated measure of LCR transcriptional activity using the area under the curve (AUC). AUC values were calculated from the dose-response curves presented in panels A-C and used as an integrated measure of LCR activity in C33A, HeLa and U2OS cells, as indicated. The AUC value of each LCR is reported as a percentage of the AUC value calculated for LCR-PT, which was set as 100. Statistical analysis of AUC values was performed using a one-way ANOVA followed by Dunnett's post-hoc analysis to compare the transcriptional activity of LCR-PT (white bar) with that of each LCR-PT mutant. \* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$ .





**Figure 3.6. Effect of the C/EBP $\beta$  LAP and LIP isoforms on the transcriptional activity of the HPV33 prototypical LCR.** A) Effect of overexpression of C/EBP $\beta$  LAP and LIP on LCR-PT activity. Reporter gene assays were performed in C33A (left panel), HeLa (center panel) and U2OS (right panel) cells. These cells were co-transfected with LCR-PT (100 ng), LCR- $\Delta$ C/EBP $\beta$ -Fluc (25 ng) as an internal control, and with an expression plasmid for either LAP (5 ng) or LIP (10 ng), as indicated. Bars represent the mean values of Rluc/Fluc ratios obtained from at least 4 independent experiments each performed in duplicates. Standard deviations are indicated by error bars. LCR-PT in the absence of LAP or LIP was used as the reference (white bars). The numbers indicated above brackets represent the fold-activation and fold-repression values of each LCR brought about by exogenous LAP and LIP, respectively. These values were obtained by comparing the basal Rluc/Fluc ratios measured in absence of exogenous LAP or LIP to those obtained in presence of 5 ng and 10 ng of LAP- and LIP-expression vector, respectively. Statistical analysis of the Rluc/Fluc values was performed using a one-way ANOVA followed by Dunnett's post-hoc analysis. \* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$ . B) Expression levels of C/EBP $\beta$  LAP and LIP. Western blot analysis of endogenous (left panel) and transiently overexpressed (right panel) LAP and LIP in C33A, HeLa and U2OS cells. In overexpression studies, cells were transfected with an expression plasmid for either LAP or LIP, as indicated, or mock-transfected as a control. Both isoforms were detected using a rabbit polyclonal antibody against C/EBP $\beta$ . The bands corresponding to LAP and LIP are indicated on the right.  $\beta$ -tubulin was used as a loading control.



**Figure 3.7. Modulation of the HPV33 LCR activity by LAP and LIP is altered by C/EBP $\beta$  binding site mutations.** The effects of overproducing LAP (5 ng; panels A, C and E) or LIP (10 ng; panels B, D and F) on the transcriptional activities of LCR-PT and mutant derivatives carrying the quadruple inactivating mutation mBS1, mBS2 or a combination of both (mBS1/2), as indicated, were determined in C33A (A and B), HeLa (C and D) and U2OS (E and F) cells exactly as described in the legend of Fig. 3.6.A. The basal activity of each LCR was measured in absence of exogenous LAP or LIP. All values are presented as Rluf/Fluc ratios obtained with 100 ng of LCR for each condition, which was used as the reference (white bar). Standard deviations are indicated by error bars.

**HPV18 (7631)** CGCTACAACAATGCTTGCATAACTATATCCACTCCCTAAGTAATAAAACT  
**HPV16 (7641)** GTACATTGTGTCATATAAAAATAAATCACTATGCGCCAACGCCTTACATACC  
**HPV33 (7686)** ATACATGCAGTGCAATTGCAAAATACTTAATTGTACTAATAGTTTACACAT

**Figure 3.8. Conservation of the minimal enhancer region in the LCR.** Alignment of the partial LCR sequences of HPV18, HPV16 and HPV33 according to the position of the HPV33 BS2 and the putative C/EBP $\beta$  site in the LCR of HPV18. The blue and green boxes show the position of the KRF-1 and PEF-1 sites in the LCR of the HPV18 and HPV16, respectively. The purple and yellow boxes show the position of the HPV33 C/EBP $\beta$  BS2 site and the putative C/EBP $\beta$  site in the LCR of HPV18, respectively.

**Table 3.1. Binding affinities ( $K_D$ ) of C/EBP $\beta$  for different sites as determined in competitive fluorescence polarization assays**

<b>Competitor</b>	<b>Sequence (5' - 3')</b>	<b>IC<sub>50</sub> (nM)</b>	<b>K<sub>i</sub> (nM)</b>
<b>Consensus</b>	CCCTGATTGCGCAATAGGCTC	54 ± 18	45 ± 15
<b>33BS1</b>	CCCTGCTTACACAATAGGCTC	97 ± 0.5	82 ± 0.5
<b>T7791C</b>	CCCTGCTCACACAATAGGCTC	>1500 <sup>a</sup>	>1263
<b>33mBS1</b>	CCCTGCCCCACACCCTAGGCTC	>1500 <sup>a</sup>	>1263

<sup>a</sup> Highest concentration of GST-C/EBP $\beta$ -BZ protein tested.

# **Chapitre quatrième : Construction d'une matrice poids-position basée sur l'énergie de liaison d'un homodimère de C/EBP $\beta$ à l'ADN et son utilisation pour la prédiction de sites à haute affinité dans la longue région de contrôle de 175 types de VPH**

**Title:** Construction of a Binding Energy-Based Position Weight Matrix Describing the affinities of C/EBP $\beta$  Homodimers for different DNA sequences and its Use to Predict High-Affinity Sites in the Long Control Region of 175 HPV Types

**Running title:** C/EBP $\beta$  binding sites in HPV LCRs

Jennifer Alvarez<sup>1,2</sup>, David Gagnon<sup>1</sup>, Anthony Mathelier<sup>3</sup> and Jacques Archambault<sup>1,2\*</sup>.

1- Laboratory of Molecular Virology, Institut de Recherches Cliniques de Montréal (IRCM) Montréal, Québec, Canada

2- Department of Biochemistry and Molecular Medicine, Université de Montréal, Montréal, Québec, Canada

3- Laboratory of Computational Biology and Gene Regulation, Centre for Molecular Medicine Norway, Norway.

Abstract word count: 271

Word count: 5359

\* Corresponding author: Laboratory of Molecular Virology, Institut de Recherches Cliniques de Montréal, 110 Pine Avenue West, Montréal, Québec, Canada, H2W 1R7, Tel: (514) 987-5739; Fax: (514) 987-5741; E-mail: [jacques.archambault@ircm.qc.ca](mailto:jacques.archambault@ircm.qc.ca)

## 4.1 Mise en contexte de l'article

Dans le chapitre précédent, nous avons montré que C/EBP $\beta$  est un régulateur important de l'expression des gènes du VPH33 par son interaction sur deux sites dans le LCR. Ainsi, dans ce chapitre, nous avons voulu déterminer la prévalence des sites C/EBP $\beta$  chez différents types de VPH afin de déterminer si C/EBP $\beta$  est un régulateur commun de l'activité transcriptionnelle virale tel que suggéré dans la littérature et par nos études. À noter que cet article sera soumis pour publication prochainement.

## 4.2 Contribution des auteurs

J'ai réalisé toutes les expériences présentées dans ce manuscrit excepté les expériences de polarisation de la fluorescence en utilisant des doubles mutants de liaison de C/EBP $\beta$  qui ont été produites par Dr David Gagnon. J'ai également rédigé cet article dans son intégralité.

Dr Anthony Mathelier a prédit les sites C/EBP $\beta$  en utilisant la matrice de fréquence.

Dr Jacques Archambault a participé à la conception et à la supervision du projet. Il a également participé à l'analyse des résultats et a rédigé le résumé.

## 4.3 Texte de l'article

### 4.3.1 Abstract

**Background.** The long-control regions (LCRs) of human papillomavirus (HPV) types 11, 16, 18 and 33 are regulated by C/EBP $\beta$ , a transcription factor that binds as homo- or heterodimers to DNA sites of the consensus sequence 5'-[A/G]TTGCG[C/T]AA[C/T]-3'. To assess the importance of C/EBP $\beta$  as a general regulator of HPV gene expression, we sought to determine the prevalence of C/EBP $\beta$  binding sites in the LCRs of 175 HPV types. **Methods.** C/EBP $\beta$  binding sites were predicted with an available frequency-based position weight matrix (PWM) derived from chromatin-immunoprecipitation studies and reflecting the occurrence of the four nucleotides A, C, G and T at each position of the C/EBP $\beta$

binding sequence in mammalian genomic DNA. As a complementary approach, a PWM based on binding energy was also used, which was created by measuring the affinity of a C/EBP $\beta$  homodimer for a series of DNA sites containing all four permutations at each position of the binding sequence in vitro.

**Results.** The two PWMs predicted slightly different but overlapping sets of C/EBP $\beta$  binding sites, with the main difference being at position 5 where a C is almost invariably found in mammalian DNA although a C, T or a G is tolerated for high-affinity binding in vitro. Previously described as well as novel sites were identified, some of which were validated in functional studies. Interestingly, both matrices highlighted the enrichment of C/EBP $\beta$  binding sites in the central region of the LCR of high-risk HPV (HR-HPV) types.

**Conclusions.** These results underscore the value of combining PWMs based on occurrence-frequency and binding-energy for the prediction of TF binding sites and point to C/EBP $\beta$  as being an important regulator of HR-HPV types.

Keywords: HPV, C/EBP $\beta$ , gene expression, transcription, long control region, LCR, binding-affinity, occurrence frequency, fluorescence polarization.

### 4.3.2 Introduction

High-risk human papillomaviruses (HR-HPV), of which HPV16 and HPV18 are the most prevalent, are the root cause of cervical cancer (CC) and other HPV-associated malignancies (1, 2). The viral oncoproteins E6 and E7, which target the p53 and Rb-tumor suppressor pathways, respectively, are necessary for cancer development and to sustain the viability and proliferation of HPV-immortalized cells (3-5). The expression of E6 and E7 is driven by the viral long control region (LCR), a regulatory part of the HPV genome which contains binding sites for cellular transcription factors (TFs) and for the viral E2 transcriptional repressor (reviewed in (6) and (7)). Increased expression of E6 and E7, which often arises following integration of the viral genome into the host chromosomes and disruption of the E2 gene, is a hallmark of HPV-associated cancers (8). Naturally-occurring variations in the LCR can also contribute to enhanced E6 and E7 expression (7, 9). For example, our group has previously described a single variation in the HPV33 LCR, T7791C, which increases its transcriptional activity approximately two-fold by abolishing a



negatively-acting CCAAT-enhancer binding protein  $\beta$  (C/EBP $\beta$ ) binding site normally present in the LCR prototype (LCR-PT,(10)).

C/EBP $\beta$  is a member of the CCAAT family of TFs and is an important regulator of genes involved in immunity, cell proliferation and differentiation (reviewed in (11) and (12)). Three C/EBP $\beta$  isoforms of 38, 35 and 20 kDa have been identified: the liver-enriched activator proteins LAP\* and LAP (herein termed LAP only) and the liver-enriched inhibitory proteins LIP, respectively (12, 13). The latter lacks the transactivation domains but retains the ability to bind DNA via its basic leucine zipper domain. The ability of C/EBP $\beta$  to activate or repress promoters is influenced by several conditions including the relative expression of the LAP and LIP isoforms, the occupancy of adjacent binding sites for other cellular factors and, more generally, the composition of the cellular transcription factor proteome (11, 12). Although C/EBP $\beta$  preferentially binds to the 5'-ATTGCGCAAT-3' consensus sequence as a homodimer (7, 14), complexes containing other factors such as CREB, NF- $\kappa$ B, ATF as well as other members of the C/EBP family (15-17) have been identified. These heterodimers can bind to different DNA targets known as composite elements thus greatly expanding the number of genes that can be regulated by C/EBP $\beta$ .

C/EBP $\beta$  was shown to act primarily as a transcriptional repressor of the HPV16 and HPV11 E6/E7 promoter (18, 19). C/EBP $\beta$  was shown to bind the 5' region of the LCR (nt 7285-7315) in gel retardation and DNA footprinting assays (18) whereas binding of C/EBP $\beta$  onto the LCR of HPV11 (nt 7451-7459) was assessed by EMSA (20). For HPV11, this inhibitory activity of C/EBP $\beta$  was found to require a promoter-proximal YY1 binding site; surprisingly, mutation of this site not only abolished repression but switched the function of C/EBP $\beta$  to that of an activator (20). Additional evidence that YY1 can modulate the activity of C/EBP $\beta$  came from the observation that both factors assemble into a complex at the "switch" region (nt 7709-7718) in the enhancer of the HPV18 LCR to activate viral gene transcription, although this mechanism was found to operate in HeLa cells only (21-23). Additionally, increased binding of C/EBP $\beta$  onto the enhancer region of the HPV31 LCR was observed by chromatin immunoprecipitation (ChIP) analysis of differentiated keratinocytes although the exact sites were not defined (24). Deletion of this region significantly decreased the activity of the late promoter. Binding of C/EBP $\beta$  onto the

HPV31 LCR was also observed in undifferentiated cells. The deletion of the region thought to contain a C/EBP $\beta$  site reduced in the early promoter activity in undifferentiated keratinocytes, suggesting an activating role for C/EBP $\beta$  in early gene expression (25, 26).

We have recently identified three C/EBP $\beta$  binding sites, termed BS1, BS2 and BS3, in the LCR of HPV33. Functional analysis in cultured cell lines and in undifferentiated primary keratinocytes implicated BS1 and BS2 in the repression and activation of viral gene expression, respectively, whereas BS3 did not have any apparent function (10). Using a competitive fluorescence polarization DNA binding assay in vitro, we determined that a recombinant C/EBP $\beta$  homodimer could bind to 33BS1 ( $K_i = 96$  nM) and 33BS2 ( $K_i = 90$  nM) with similar affinities as to the consensus sequence ATTGCGCAAT ( $K_i = 40$  nM) (10). In contrast, the binding of C/EBP $\beta$  to 33BS3 was rather weak ( $K_i = 181$  nM), thus providing a likely explanation for its lack of effect in reporter gene assays. Taken together, these studies in several different HPV types suggested that C/EBP $\beta$  may be a general regulator of HPV gene transcriptional.

In the present study, we have performed a bioinformatic analysis of the occurrence of potential C/EBP $\beta$  binding sites in the LCRs of 175 HPV types. C/EBP $\beta$  sites were predicted using two complementary position weight matrices (PWM). One derived from ChIP studies and depicting the nucleotide composition of C/EBP $\beta$ -bound sequences in mammalian DNA. The other, constructed in this study, reflecting the nucleotide sequences that are bound with high-affinity by a C/EBP $\beta$  homodimer in vitro. Our analysis shows an overrepresentation of C/EBP $\beta$  sites in the LCR of mucosal HPV types, in particular within the high-risk group, suggesting that transcriptional regulation by C/EBP $\beta$  is a conserved feature of oncogenic HPV types.

### **4.3.3 Materials and Methods**

#### **C/EBP $\beta$ expression and purification**

The GST-C/EBP $\beta$ -BZ protein, comprised of the basic region (B) and leucine-zipper (Z) of C/EBP $\beta$  fused to glutathione-S-transferase (GST), was expressed in E. coli BL21

cells and purified exactly as described in (10). Protein concentrations were determined by Bradford analysis.

### **Fluorescence polarization DNA-binding assay**

DNA-binding assays were performed as described in (27) and (10) using a final concentration of 15 nM of fluorescein-labeled BS1 probe and 31.25 nM of GST-C/EBP $\beta$ -BZ protein in a final volume of 150 $\mu$ L of assay buffer (25mM Tris-HCl [pH 7.6], 0.1mM DTT, 0.01% NP-40, 75mM NaCl, 150mM KCl and 5mM MgCl<sub>2</sub>). The affinities of unlabeled duplex DNAs for GST-C/EBP $\beta$ -BZ were measured in competition assays by titrating them in DNA binding reactions prepared as described above. Each unlabeled duplex DNA was tested at increasing concentrations (0, 5.9, 11.7, 23.4, 46.9, 93.8, 187.5, 375, 750 and 1500 nM) such as to generate an inhibitory dose-response curve from which a  $K_i$  value could be calculated (see below). Fluorescence readings were taken in OptiPlate-96F high-binding-affinity black 96-well plates (Perkin Elmer) with a Victor<sup>3</sup>V multilabel plate reader (PerkinElmer) equipped with a 485-nm/535-nm filter set. The background fluorescence of the unbound probe was subtracted from all values. Each DNA competitor was tested in duplicates in two or three independent experiments (n=6) as indicated in the figure legends.

### **Fluorescent DNA probe and competitor oligonucleotides**

The BS1 probe is identical to the previously described LCR-PT probe (10). This probe contains a single C/EBP $\beta$  binding site, corresponding to BS1 (5'-TTTGGCCTTACACAATTGCTTT-3', BS1 underlined), and is linked to a fluorescein molecule at its 5' end by a six-carbon linker. This probe and the following unlabeled competitor oligonucleotides were purchased from Integrated DNA Technologies. Unlabeled duplex DNAs and the fluorescein-labeled BS1 probe were prepared by annealing complementary oligonucleotides at a ratio of 1:1 and of 1:1.5 as previously described, respectively (10, 27). The sequence of the C/EBP $\beta$  consensus site used as a competitor was as follows: 5'-CCCTGATTGCGCAATAGGCTC-3' (C/EBP $\beta$  site is underlined). Derivatives of this consensus sequence were obtained by substituting each base of one ATTGC half-site by the other three possible nucleotides. Double mutants were obtained similarly but this time by substituting each nucleotide in both half-sites simultaneously.

Unlabeled oligonucleotides containing previously described C/EBP $\beta$  binding sites in HPV16 (16BS1, 16BS2), HPV11 (11BS1), HPV18 (18BS1, 18BS2, 18BS3) and in HPV33 (33BS1, 33BS2, 33BS3) were purchased and prepared as described above and their nucleotide sequences indicated in Tables 4.1, 4.3, 4.4 and 4.5.

### **K<sub>i</sub> determination**

Inhibitory constant (K<sub>i</sub>) values were calculated with the GraphPad Prism version 6.00 software by fitting the data from inhibitory dose-response curves to a single binding-site equilibrium model using nonlinear least-squares regression analysis, as previously described (10, 27). K<sub>i</sub> values were obtained from two or three independent experiments with each data point performed in duplicate as described in the figure legends.

### **HPV LCR DNA sequences**

The nucleotide sequences of the LCRs from 175 HPV types were retrieved from the Papillomavirus Episteme website (<http://pave.niaid.nih.gov>) (28).

### **Definition of the binding predictor matrix**

The energetic contribution of each position of the C/EBP $\beta$  consensus sequence (ATTGCGCAAT, [7, 13]) was derived from the K<sub>i</sub> values of mutated sequences in which each nucleotide was changed to the other three possibilities. The changes in Gibbs free energy ( $\Delta\Delta G$ ) were calculated from the following equation 1 as previously published in (29):  $\Delta\Delta G = RT\ln[K_{i(2)}/K_{i(1)}]$  where K<sub>i(1)</sub> is the inhibitory constant measured for the C/EBP $\beta$ -consensus sequence and K<sub>i(2)</sub> is the inhibitory constant of the mutated C/EBP $\beta$ -binding sequence.  $\Delta\Delta G$  are shown in Table 4.2.

### **Computational prediction of C/EBP $\beta$ binding sites**

C/EBP $\beta$  binding sites were predicted with the sequence analysis tools of the MEME suite using the binding energy position weight matrix (PWM) depicted in Fig. 4.2 and, separately, the frequency-based JASPAR transcription factor binding profile MA0466.1 (31). Sites with a p value >0.0005 were discarded. The sequence logo representing the

energy-based PWM was generated with enoLOGO and graphically-rendered using the MEME-Suite (30, 31).

### **Prediction of the affinity of C/EBP $\beta$ -DNA interactions**

The affinity of C/EBP $\beta$  for a given DNA sequence was estimated from the following equation 2:  $\ln K_i(x) = \ln K_i(\text{ref}) + \sum \Delta \ln K_i(i,N)$  where  $K_i(\text{ref})$  and  $K_i(x)$  represent the inhibitory constant values of the consensus C/EBP $\beta$  sequence (ref) and of a mutant derivative (x), respectively, and where  $\Delta \ln K_i(i,N)$  corresponds to the change in Gibbs free energy ( $\Delta\Delta G$ ; calculated as described above) caused by a nucleotide substitution,  $N$ , at the  $i$ th position of the binding sequence (adapted from (32)).

### **HPV LCR sequence alignment**

The LCR sequences from 175 HPV types were aligned using Clustal Omega (EMBL-EB) and visualized with Jalview (33, 34).

### **Plasmid construction and mutagenesis**

The HPV33 LCR-Rluc plasmid expressing Renilla luciferase (Rluc) from the prototypical HPV33 LCR (LCR-PT) as well as the internal control plasmid encoding firefly luciferase (Fluc) downstream of LCR-PT (LCR-PT-Fluc) have been described in (35). The plasmid expressing Rluc from the HPV18 LCR Rluc plasmid encoding the prototypical HPV18 LCR ( nt 7137-7857 and 1-104 of the HPV18 genome) in front of the Rluc coding region was constructed by PCR amplification of the LCR from the reference genome and cloning of the resulting amplicon between the MluI and NcoI restriction sites of the promoterless pGL3-Rluc plasmid described in (35). Sequences of the primers used for amplification will be made available upon request. The HPV33 mBS2 mutant LCR, in which the C/EBP $\beta$  BS2 site was inactivated by mutations (ATTGCAAAT changed to ACCGCACCT), has been described previously (10). The corresponding site in the HPV18 LCR (18BS2) was similarly inactivated (CTTGTACAAC changed to CCCGTACCCC), using the QuikChange Site-Directed Mutagenesis kit (Stratagene). All constructs were verified by DNA sequencing. Additional information on the construction of these plasmids will be made available upon request.

## **Cell culture and transfection**

The C33A (HPV-negative) and HeLa (HPV18-positive) cervical carcinoma cell lines, as well as the osteosarcoma-derived U2OS line, were grown and transfected using Lipofectamine 2000 (Life Technologies, cat: 11668-500) as previously described in (35).

## **Luciferase-reporter transcription assay**

This protocol is described in (35) and (10). Briefly, C33A, HeLa or U2OS cells were co-transfected with increasing quantities of LCR-Rluc plasmid (25, 50, 100 and 200 ng) from HPV33, HPV18 or their respective C/EBP $\beta$  binding site mutants (mBS2) and 25 ng of the LCR-PT-Fluc control plasmid. Fluc and Rluc activities were measured with the Dual-Glo luciferase assay system and a GloMax<sup>TM</sup> 96-well luminometer (Promega) 24h post-transfection. LCR transcriptional activities are presented either as Rluc/Fluc ratios or relative to the transcriptional activity measured with 100 ng of HPV33 LCR-Rluc, which was assigned a value of 1.0.

## **Statistical analysis**

Statistical significance was tested using a one-way ANOVA with Dunnett's post-hoc analysis. Statistics and AUC calculations were performed using GraphPad Prism version 6.00.

## **4.3.4 Results**

### **Choice of C/EBP $\beta$ position weight matrices**

Several consensus motifs are available to predict potential binding sites for C/EBP $\beta$ . As well as having comparable TFs coverage than commercially available resources, the open-access database Jaspar provides non-redundant curated motifs of higher quality than those found in other databases such as Transfac® (36, 37). Additionally, the binding preferences of transcription factors are mainly derived from published collections of datasets obtained by chromatin immunoprecipitation-sequencing (ChIP-seq) experiments. Sites identified using this PWM are defined depending on the frequency of a given nucleotide in the consensus sequence. However, such approach probably does not

distinguish between sites for hetero- or homodimers of C/EBP $\beta$  as well as direct and indirect interactions with DNA. The fact that the Jaspar matrix is not palindromic (Fig. 4.1) i.e the sequence of one half-site is more degenerate than that of the other half-site, lends further supports that this matrix identifies mostly sites for heterodimers. In addition, the Jaspar matrix does not recognize some of the previously reported sites in HPV11 and HPV31 (20, 24-26 and data not shown). To circumvent the limitations described above and as a complementary approach, we develop a PWM that would not be based solely on occurrence (i.e. on ChIP data) but one based on affinity for C/EBP $\beta$  homodimers.

### **Construction of a binding-energy based PWM**

To quantify the importance of each base-pair of the consensus sequence ATTGCGCAAT in binding a homodimer of C/EBP $\beta$ , we systematically changed each of the five nucleotides in one half-site ( $A_1T_2T_3G_4C_5$ ) of this palindrome to the other three possibilities (e.g.  $A_1$  was changed to C, G and T) and measured the affinities of the resulting 15 mutant sites for C/EBP $\beta$  in vitro. Specifically, each mutant site was used as a competitor in our previously described fluorescence polarization assay which measures the interaction of the C/EBP $\beta$  DNA-binding domain (basic region and leucine-zipper fused to GST) with a fluorescent DNA probe carrying a high-affinity C/EBP $\beta$  binding-site (10). This allowed a  $K_i$  value to be measured for each mutant binding site, and for the consensus sequence as a control, which is indicative of its affinity for a homodimer of C/EBP $\beta$ . The results summarized in Table 4.1 show that nucleotide changes at positions 2 and 3, and to a lesser degree at position 4, of the C/EBP $\beta$  consensus sequence resulted in a significant reduction in affinity, thus highlighting the importance of T2, T3 and, albeit more modest, of G4 for specific C/EBP $\beta$  binding. From the data obtained with all mutant sites, we then calculated the changes in Gibbs free energy of binding ( $\Delta\Delta G$ ) caused by each substitution in one half-site to generate the binding predictor matrix described in Table 4.2 and represented by the sequence logo shown in Fig. 4.1. Although this logo shares similarities to the one describing the frequency-based C/EBP $\beta$  binding matrix from JASPAR (MA0466.1) (36), there are also important differences. In particular, the high prevalence of a C at position 5 of the frequency-based JASPAR matrix is not observed in the binding energy matrix, indicating that a C is not essential for high-affinity binding in vitro. Although the reason why a C residue at position 5 is almost invariably observed in ChIP

studies remains to be established, one could speculate that it is required/preferred for the binding of C/EBP $\beta$  heterodimers. Overall, the results from this systematic permutation analysis highlight the importance of T<sub>2</sub>, T<sub>3</sub> and G<sub>4</sub> in both half-sites for strong and specific binding to a C/EBP $\beta$  homodimer and confirm that the consensus sequence ATTGCGCAAT has the greatest affinity for this factor.

### **Validation of the binding energy PWM**

To further assess the sequence requirement for binding a C/EBP $\beta$  homodimer and to validate the accuracy of our binding energy PWM, we performed similar studies as those described above but using mutant sites in which each nucleotide substitution was introduced in both half-sites of the consensus sequence simultaneously. If every nucleotide contributes independently of the others to the binding of C/EBP $\beta$ , the  $\Delta\Delta G$  associated with a given substitution in one half-site (Table 4.1) should be doubled when this same substitution is introduced in both half-sites (Table 4.3). However, we note that the  $\Delta\Delta G$  values were not doubled when two substitutions were introduced in the consensus sequence compared to a single mutant site in particular for the central GC palindromic pair at positions 5 and 6 whose substitution had little to no effect on binding affinity, showing that the identity of these two positions is not important for the binding of homodimeric C/EBP $\beta$ . Our mutational analysis also confirms the importance of T<sub>2</sub>, T<sub>3</sub> and G<sub>4</sub> in conferring high-affinity binding.

Additionally, we found that some permutations in one half-site in combination with their paired mutations on the other half-site potentiate the effect of the single mutant, showing that the identity of a nucleotide in a given position may influence the effect of another nucleotide on the overall binding of C/EBP $\beta$ . This is best exemplified by the T<sub>3</sub> and A<sub>8</sub> pair. As such, when T<sub>3</sub> is changed for G, the binding affinity is not altered. However, when A<sub>8</sub> is simultaneously change for C, binding of C/EBP $\beta$  to the DNA target is abolished. These results show that some changes in the consensus sequence are not necessarily additive to the overall binding as reflected in bindings affinities between single and double mutants. Nonetheless, the correlation between predicted and measured binding affinities was good ( $R^2 = 0.7569$ , Fig. 4.2) as further validated by comparing the predicted and measured  $K_i$  values for the previously described HPV33 BS1, -2 and 3. This shows that



the additivity model, although not perfect, still provides a good approximation of the experimental  $K_i$  values and can therefore be used to predict high-affinity binding sites for C/EBP $\beta$ .

### **Comparison of the energy- and frequency-based C/EBP $\beta$ binding site predictions**

The results above indicate that the JASPAR and binding energy matrices can identify different sets of binding sites especially those that diverge at position 5. Thus both matrices were used to analyse the prevalence of putative C/EBP $\beta$  binding sites in the LCR of 175 HPV types. The energy- and frequency-based matrices identified 55 and 103 sites, respectively, among which 39 sites were found by both methods (Fig. 4.3A). Sequence logos representing the sites identified by both matrices or detected only by the energy-PWM or by the frequency-based PWM are presented Fig. 4.3B, 4.3C and 4.3D, respectively. As expected, the major difference between the sites identified only by the frequency matrix and those detected only by the energy-matrix was at position 5 where a C is almost invariable in the JASPAR matrix but where T, C or G are accommodated by the binding-energy matrix. Small differences were also observed at positions 3, 6 and 7. While T3 was always conserved in the sites identified by the frequency-based matrix, this position could be T or to a lesser extent G or C in sites predicted by the binding energy PWM but only when A8 and A9 are conserved. The identity of the nucleotides at positions 6 and 7 were more variable in sites predicted by the frequency-based PWM than those identified by the binding energy matrix which were preferentially enriched in G and C, respectively. This highlights the need for two high-affinity half-site for the binding of a C/EBP $\beta$  homodimer which is not required in the sites that are identified by the frequency-derived logo and could presumably be heterodimeric sites. The results indicate that although both predictions identify several common sites, a significant number of sites are missed by each matrix, showing that sequences found at high frequency in ChIP data sets are not necessarily high-affinity sites for C/EBP $\beta$  homodimers.

Several observations were noted when further characterizing the sites predicted with the energy- or frequency-based PWM. To facilitate the study, the LCRs of the 175 HPV types were separated into three groups: the high-risk (HR), the low-risk (LR), representing the mucosal HPVs, and the cutaneous (Ct) HPVs. We found that C/EBP $\beta$  binding sites

were overrepresented in the HR-HPVs compared to the LR and Ct types as observed using the energy- (Fig. 4A) and frequency-based PWM (Fig. 4B). When dividing the LCR into three regions determined by the position of the E2 binding sites (the 5', the central and the 3' regions, as described in (6)), the C/EBP $\beta$  binding sites were categorized according to the region they fall in. The results presented in Fig. 4.4C show that the C/EBP $\beta$  binding sites of the HR and LR groups were more often found in the central region of the LCR compared to those of the Ct HPVs which were predominantly found in the 3' region. The majority of the C/EBP $\beta$  binding sites found in the central region of the LCR of the HR groups were identified by both the energy- and frequency-derived PWM, thus lending further support to the notion that they are genuine high-affinity binding sites for C/EBP $\beta$ . Collectively, these results point to C/EBP $\beta$  as being an important regulator of HR-HPV gene expression.

Among the previously described C/EBP $\beta$  sites in the literature, the binding energy PWM accurately identified the one present in HPV11 (here termed 11BS1) and the three found in HPV33 (33BS1, 33BS2 and 33BS3) while the frequency-derived logo only identified 33BS1 and 33BS2 (10, 20, 38). Two distinct sites (18BS2 and 18BS3) were identified in the HPV18 LCR that are different than the C/EBP $\beta$  site previously described (termed here 18BS1) (22). While both matrices predicted 18BS2, 18BS3 was specifically found by the binding energy logo only. Additionally, both methods identified different sites in the LCR of HPV16, 16BS1 (energy) and 16BS2 (frequency) and also failed to identify any binding sites in the LCR of HPV31. The exact sequence of the C/EBP $\beta$  element(s) in the HPV16 LCR was never assessed (18). The sequences as well as the major functions of the previously characterized sites are found in Table 4.4. The sites that were missed by the frequency-based matrix contained different nucleotides than that of the energy logo at positions 1, 4 and particularly 5 which is T in 33BS3 and 11BS1. The 18BS1 site could not be detected by both methods possibly because of C4 which was found to be deleterious to the binding according to the binding energy PWM as well as not being detected *in vivo* using the JASPAR matrix.

### **Affinity of C/EBP $\beta$ for previously reported sites in the LCRs of HPV types 11, 16 and 18**

Based on the results of the prediction, the sites that have been previously described were re-examined for their ability to bind a C/EBP $\beta$  homodimer using the same competition fluorescence polarization binding assay described above. As expected, 11BS1 ( $K_i = 156 \text{ nM} \pm 28$ ) and 16BS1 ( $K_i = 226 \text{ nM} \pm 195$ ) which were identified by the binding energy matrix were bound by C/EBP $\beta$  although with comparable affinity as 33BS1 ( $K_i = 103 \text{ nM} \pm 26$ ). Furthermore, our results also show that homodimeric C/EBP $\beta$  does not bind to 16BS2 and 18BS3 identified by the frequency matrix only. Taken together, these results show that the LCR of HPV16, -11, -18 and -33 contain binding sites for purified homodimeric C/EBP $\beta$ . It is possible that the low-affinity sites identified here are preferably bound by C/EBP $\beta$  as a heterodimer.

### **Conservation of 33BS2-like sites in HR-HPV types**

The presence of the high-affinity 18BS2 was rather intriguing as this site is similarly positioned in the LCR of HPV18 as 33BS2 in the LCR of HPV33, thus suggesting that both sites may regulate their cognate LCR in a similar fashion. A previous study by our laboratory showed that 33BS2 stimulates the activity of the HPV33 LCR in HeLa cells and to a lesser extent in U2OS cells. To investigate if 18BS2 similarly enhance the activity of the HPV18 LCR, this site was inactivated by a quadruple mutation which eliminates binding to C/EBP $\beta$ , as previously reported in (10). The transcriptional activity of the resulting mutant LCR, HPV18 mBS2, was then investigated in reporter gene assays as described in (10). We have previously shown that mutation of BS2 reduces the activity of the HPV33 LCR by approximately 75 % and 30 % in HeLa and U2SO cells, respectively, but has little to no effect LCR activity in C33A cells ((10) and Fig. 4.7 of this study). The inactivation of BS2 in the HPV18 LCR had essentially the same effects. Specifically, mBS2 reduced the activity of the HPV18 LCR by 75 % in both HeLa and U2OS cells while having little to no effect in C33A cells. These results provide further evidence that 33BS2 and 18BS2 play analogous roles in regulating transcription of their cognate LCR.

The presence of this common activator site in HPV33 and HPV18 prompted us to investigate the prevalence of BS2-like sites in other HPV types. To do so, the DNA sequence of the HPV types that contained at least one putative C/EBP $\beta$  binding site in the central region of the LCR were aligned according to 33BS2 (Fig. 4.7A) and 33BS1 (Fig.

4.7B). These sites were found predominantly in the mucosal HPV types 33, 58, 39, 68, 59 and 18 and in the Ct HPV7. Interestingly, all these types also contained a sequence similar to the BS1 site in the HPV33 LCR, which has been shown to repress transcription (10). These observations suggest that the ability of C/EBP $\beta$  to repress and enhance the activity of the LCR through BS1 and BS2, respectively, may be a conserved feature of several HPV types.

### 4.3.5 Discussion

In this study, we have measured the effect of substituting one nucleotide in one half-site of the ATTGCGCAAT consensus sequence on the binding affinity of a C/EBP $\beta$  homodimer. C/EBP $\beta$  binding sites in the LCR of 175 HPV types were then predicted using an energy- and frequency-based PWM and the results obtained from both methods were compared.

C/EBP $\beta$  is a known regulator of HPV gene expression able to repress and to activate the LCR depending on the cellular context (10, 18, 20, 22). The relative expression of the activating LAP and repressor LIP isoforms regulates C/EBP $\beta$  DNA-binding activity (7, 11, 13, 39-43). The DNA sequence of the C/EBP $\beta$  binding site is also a major determinant of C/EBP $\beta$  activity; some sites allowing the binding of C/EBP $\beta$  and interacting partners on heterodimeric sites. Theoretically, every nucleotide in a given binding sequence contributes independently to the overall binding of a transcription factor meaning that the binding affinity can be inferred by the sum of the changes in energy caused by substitutions in the reference sequence. However, we found that some permutations in both half-sites did not alter binding affinities while some changes resulted in a potentiated effect leading to complete disruption of the C/EBP $\beta$ -DNA interaction (Table 4.3). Hence, the contribution of a nucleotide in a given position may be dependent on the identity of a nucleotide in a different position. This phenomenon was also observed for the p53-DNA interaction and for other transcription factor (32, 44, 45). Even when taking this into consideration, we nevertheless found that the binding energy model still provides a good approximation of the binding affinities although the actual  $K_i$  values were underestimated.

Both frequency- and energy-based PWMs identified common sites such as that of HPV33 (BS1 and BS2). However, several sites were identified either by one matrix or the other. As such, the frequency-based matrix failed to identify sites in the LCR of HPV11 whereas the energy-based PWM did so, suggesting that 11BS1 is a site for C/EBP $\beta$  homodimers. In the LCR of HPV18, both matrices identified 18BS2 but only the energy-based matrix predicted 18BS3, also suggesting that the latter is a homodimeric site. As for 18BS1, which was previously shown to bind a YY1-C/EBP $\beta$  complex in HeLa cells (22), it was not predicted by any matrices, suggesting that C/EBP $\beta$  does not physically interact with the DNA. Both the energy- and frequency-based PWMs did not predict any C/EBP $\beta$  site in the LCR of HPV31 whereas others have identified increased binding of C/EBP $\beta$  unto the LCR in chromatin immunoprecipitation assays in differentiated keratinocytes, arguing for the fact that C/EBP $\beta$  heterodimers are preferably recruited only when differentiation is induced (24-26). Finally, different HPV16 sites were identified by both approaches; 16BS1 and 16BS2 being predicted by the energy- and frequency-based PWMs, respectively. The different sets of binding sites specifically identified by the JASPAR matrix such as 16BS1 may represent heterodimeric sites. These sites are enriched with a C at position 5 on one ATTGC half-site, suggesting that the identity of this nucleotide is critical for the binding of a C/EBP $\beta$  heterodimer whereas it can be substituted by any base for binding of a C/EBP $\beta$  homodimer. Several C/EBP $\beta$  interactors have been identified such as YY1, ATF, NFAT, CREB and NF- $\kappa$ B (22, 46-50). However, computational analysis of the second half-site sequence using the PROMO software (TRANSFAC version 8.3 database) did not reveal a particular binding sites for known C/EBP $\beta$  interactors (51, 52). Composite sites that allow the binding of a C/EBP $\beta$  heterodimer are not identified in our binding energy prediction where only the binding of a C/EBP $\beta$  homodimer was assessed. In that sense, it would be interesting to evaluate the change in binding affinities of different C/EBP $\beta$  complexes for a given site in a way that heterodimeric sites can be discriminated from homodimeric ones.

Finally, we note the presence of a conserved region (5'-TGCTTTAGGCACATATTT-3') between the clusters of 33BS2- and 33BS1-like sites (Fig. 4.7). Unfortunately, computational analysis of this region did not identify a candidate

factor that could interact with this region. However, this region partially falls within a segment which is thought to be the binding site for the papilloma enhancing binding factor (PEF-1) described in HPV16, shown to be important for the activation of viral gene expression (53). Our results raise the possibility that both activating and inhibitory C/EBP $\beta$  binding sites are required to regulate the LCR transcriptional in the mucosal HPV types. Regarding this, we found that most of the HR and LR-HPVs contained at least one C/EBP $\beta$  element in the central portion of the LCR which are similar to the 33BS2 and 33BS1 previously described (10). The central region is thought to confer tissue-specificity and may be a determinant of the epithelial tropism of the different HPVs, suggesting that BS2-like sites are involved in the enhancer activity while BS1- like sites serves to counterbalance the activation in epithelial cells.

#### 4.3.6 Bibliography

1. **Bosch FX, Lorincz A, Munoz N, Meijer CJ, Shah KV.** 2002. The causal relation between human papillomavirus and cervical cancer. *Journal of clinical pathology* **55**:244-265.
2. **Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, Snijders PJ, Peto J, Meijer CJ, Munoz N.** 1999. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *The Journal of pathology* **189**:12-19.
3. **Vande Pol SB, Klingelutz AJ.** 2013. Papillomavirus E6 oncoproteins. *Virology* **445**:115-137.
4. **Roman A, Munger K.** 2013. The papillomavirus E7 proteins. *Virology* **445**:138-168.
5. **Munger K, Scheffner M, Huibregtse JM, Howley PM.** 1992. Interactions of HPV E6 and E7 oncoproteins with tumour suppressor gene products. *Cancer surveys* **12**:197-217.
6. **O'Connor MJ, Chan SY, Bernard HU.** 1995. Transcription factor binding sites in the long control regions of genital HPVs, part III-A, p. pp. 21-40. *In* Myers G BHU, Delius H, Baker C, Icenogle J, Halpern A, Wheeler C (ed.), *Human papillomaviruses 1995 compendium*. Los Alamos National Laboratory, Los Alamos, N.M. USA.
7. **Bernard HU.** 2013. Regulatory elements in the viral genome. *Virology* **445**:197-204.
8. **Howley P, Lowy D.** 2007. Papillomaviruses, p. 2299–2354. *In* Knipe D, Howley P (ed.), *Fields Virology*, vol. Vol. 5. Lippincott Williams & Wilkins, Philadelphia, PA.

9. **Lace MJ, Isacson C, Anson JR, Lorincz AT, Wilczynski SP, Haugen TH, Turek LP.** 2009. Upstream regulatory region alterations found in human papillomavirus type 16 (HPV-16) isolates from cervical carcinomas increase transcription, ori function, and HPV immortalization capacity in culture. *Journal of virology* **83**:7457-7466.
10. **Alvarez J, Coutlee F, Archambault J.** 2016. A Rare Polymorphism Enhances the Transcriptional Activity of the HPV33 Long Control Region by Disrupting a Binding Site for the Cellular Transcription Factor C/EBP $\beta$  (unpublished).
11. **Nerlov C.** 2007. The C/EBP family of transcription factors: a paradigm for interaction between gene expression and proliferation control. *Trends in cell biology* **17**:318-324.
12. **Tsukada J, Yoshida Y, Kominato Y, Auron PE.** 2011. The CCAAT/enhancer (C/EBP) family of basic-leucine zipper (bZIP) transcription factors is a multifaceted highly-regulated system for gene regulation. *Cytokine* **54**:6-19.
13. **Descombes P, Schibler U.** 1991. A liver-enriched transcriptional activator protein, LAP, and a transcriptional inhibitory protein, LIP, are translated from the same mRNA. *Cell* **67**:569-579.
14. **Osada S, Yamamoto H, Nishihara T, Imagawa M.** 1996. DNA binding specificity of the CCAAT/enhancer-binding protein transcription factor family. *The Journal of biological chemistry* **271**:3891-3896.
15. **LeClair KP, Blonar MA, Sharp PA.** 1992. The p50 subunit of NF-kappa B associates with the NF-IL6 transcription factor. *Proceedings of the National Academy of Sciences of the United States of America* **89**:8145-8149.
16. **Vallejo M, Ron D, Miller CP, Habener JF.** 1993. C/ATF, a member of the activating transcription factor family of DNA-binding proteins, dimerizes with CAAT/enhancer-binding proteins and directs their binding to cAMP response elements. *Proceedings of the National Academy of Sciences of the United States of America* **90**:4679-4683.
17. **Williams SC, Cantwell CA, Johnson PF.** 1991. A family of C/EBP-related proteins capable of forming covalently linked leucine zipper dimers in vitro. *Genes & development* **5**:1553-1567.
18. **Kyo S, Inoue M, Nishio Y, Nakanishi K, Akira S, Inoue H, Yutsudo M, Tanizawa O, Hakura A.** 1993. NF-IL6 represses early gene expression of human papillomavirus type 16 through binding to the noncoding region. *Journal of virology* **67**:1058-1066.
19. **Wang H, Liu K, Yuan F, Berdichevsky L, Taichman LB, Auborn K.** 1996. C/EBPbeta is a negative regulator of human papillomavirus type 11 in keratinocytes. *Journal of virology* **70**:4839-4844.
20. **Ralph WM, Jr., Liu K, Auborn KJ.** 2006. CCAAT/enhancer-binding protein beta represses human papillomavirus 11 upstream regulatory region expression through a promoter-proximal YY1-binding site. *The Journal of general virology* **87**:51-59.
21. **Bauknecht T, Jundt F, Herr I, Oehler T, Delius H, Shi Y, Angel P, Zur Hausen H.** 1995. A switch region determines the cell type-specific positive or negative action of YY1 on the activity of the human papillomavirus type 18 promoter. *Journal of virology* **69**:1-12.
22. **Bauknecht T, See RH, Shi Y.** 1996. A novel C/EBP beta-YY1 complex controls the cell-type-specific activity of the human papillomavirus type 18 upstream regulatory region. *Journal of virology* **70**:7695-7705.

23. **Bauknecht T, Shi Y.** 1998. Overexpression of C/EBPbeta represses human papillomavirus type 18 upstream regulatory region activity in HeLa cells by interfering with the binding of TATA-binding protein. *Journal of virology* **72**:2113-2124.
24. **Wooldridge TR, Laimins LA.** 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.
25. **Sen E, Bromberg-White JL, Meyers C.** 2002. Genetic analysis of cis regulatory elements within the 5' region of the human papillomavirus type 31 upstream regulatory region during different stages of the viral life cycle. *Journal of virology* **76**:4798-4809.
26. **Sen E, Alam S, Meyers C.** 2004. Genetic and biochemical analysis of cis regulatory elements within the keratinocyte enhancer region of the human papillomavirus type 31 upstream regulatory region during different stages of the viral life cycle. *Journal of virology* **78**:612-629.
27. **Fradet-Turcotte A, Vincent C, Joubert S, Bullock PA, Archambault J.** 2007. Quantitative analysis of the binding of simian virus 40 large T antigen to DNA. *Journal of virology* **81**:9162-9174.
28. **Van Doorslaer K, Tan Q, Xirasagar S, Bandaru S, Gopalan V, Mohamoud Y, Huyen Y, McBride AA.** 2013. The Papillomavirus Episteme: a central resource for papillomavirus sequence data and analysis. *Nucleic acids research* **41**:D571-578.
29. **Alexander KA, Phelps WC.** 1996. A fluorescence anisotropy study of DNA binding by HPV-11 E2C protein: a hierarchy of E2-binding sites. *Biochemistry* **35**:9864-9872.
30. **Workman CT, Yin Y, Corcoran DL, Ideker T, Stormo GD, Benos PV.** 2005. enoLOGOS: a versatile web tool for energy normalized sequence logos. *Nucleic acids research* **33**:W389-392.
31. **Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS.** 2009. MEME SUITE: tools for motif discovery and searching. *Nucleic acids research* **37**:W202-208.
32. **Veprintsev DB, Fersht AR.** 2008. Algorithm for prediction of tumour suppressor p53 affinity for binding sites in DNA. *Nucleic acids research* **36**:1589-1598.
33. **McWilliam H, Li W, Uludag M, Squizzato S, Park YM, Buso N, Cowley AP, Lopez R.** 2013. Analysis Tool Web Services from the EMBL-EBI. *Nucleic acids research* **41**:W597-600.
34. **Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ.** 2009. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics (Oxford, England)* **25**:1189-1191.
35. **Alvarez J, Coutlee F, Archambault J.** 2016. The Combination of Specific Nucleotide Variations Modulates the Transcriptional Activity of the Human Papillomavirus Type 33 Long Control Region (unpublished).
36. **Mathelier A, Zhao X, Zhang AW, Parcy F, Worsley-Hunt R, Arenillas DJ, Buchman S, Chen CY, Chou A, Ienasescu H, Lim J, Shyr C, Tan G, Zhou M, Lenhard B, Sandelin A, Wasserman WW.** 2014. JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. *Nucleic acids research* **42**:D142-147.

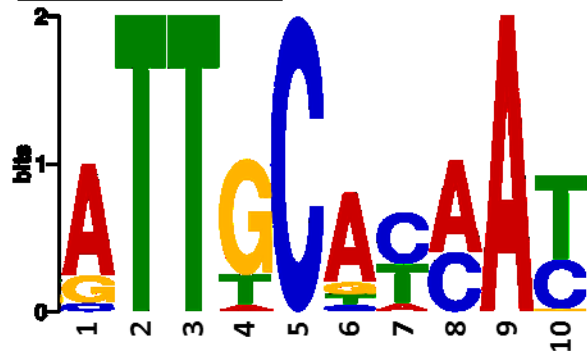


37. **Dabrowski M, Dojer N, Krystkowiak I, Kaminska B, Wilczynski B.** 2015. Optimally choosing PWM motif databases and sequence scanning approaches based on ChIP-seq data. *BMC bioinformatics* **16**:140.
38. **Auborn KJ, Steinberg BM.** 1991. A key DNA-protein interaction determines the function of the 5'URR enhancer in human papillomavirus type 11. *Virology* **181**:132-138.
39. **Struyk L, van der Meijden E, Minnaar R, Fontaine V, Meijer I, ter Schegget J.** 2000. Transcriptional regulation of human papillomavirus type 16 LCR by different C/EBPbeta isoforms. *Molecular carcinogenesis* **28**:42-50.
40. **Gunasekharan V, Hache G, Laimins L.** 2012. Differentiation-dependent changes in levels of C/EBPbeta repressors and activators regulate human papillomavirus type 31 late gene expression. *Journal of virology* **86**:5393-5398.
41. **Darlington GJ, Ross SE, MacDougald OA.** 1998. The role of C/EBP genes in adipocyte differentiation. *The Journal of biological chemistry* **273**:30057-30060.
42. **Johnson PF.** 2005. Molecular stop signs: regulation of cell-cycle arrest by C/EBP transcription factors. *Journal of cell science* **118**:2545-2555.
43. **Liu Y, Nonnemacher MR, Wigdahl B.** 2009. CCAAT/enhancer-binding proteins and the pathogenesis of retrovirus infection. *Future microbiology* **4**:299-321.
44. **Bulyk ML, Johnson PL, Church GM.** 2002. Nucleotides of transcription factor binding sites exert interdependent effects on the binding affinities of transcription factors. *Nucleic acids research* **30**:1255-1261.
45. **Benos PV, Bulyk ML, Stormo GD.** 2002. Additivity in protein-DNA interactions: how good an approximation is it? *Nucleic acids research* **30**:4442-4451.
46. **Fawcett TW, Martindale JL, Guyton KZ, Hai T, Holbrook NJ.** 1999. Complexes containing activating transcription factor (ATF)/cAMP-responsive-element-binding protein (CREB) interact with the CCAAT/enhancer-binding protein (C/EBP)-ATF composite site to regulate Gadd153 expression during the stress response. *The Biochemical journal* **339 ( Pt 1)**:135-141.
47. **Oh M, Dey A, Gerard RD, Hill JA, Rothermel BA.** 2010. The CCAAT/enhancer binding protein beta (C/EBPbeta) cooperates with NFAT to control expression of the calcineurin regulatory protein RCAN1-4. *The Journal of biological chemistry* **285**:16623-16631.
48. **Niehof M, Kubicka S, Zender L, Manns MP, Trautwein C.** 2001. Autoregulation enables different pathways to control CCAAT/enhancer binding protein beta (C/EBP beta) transcription. *Journal of molecular biology* **309**:855-868.
49. **Catron KM, Brickwood JR, Shang C, Li Y, Shannon MF, Parks TP.** 1998. Cooperative binding and synergistic activation by RelA and C/EBPbeta on the intercellular adhesion molecule-1 promoter. *Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research* **9**:949-959.
50. **Shelest E, Kel AE, Goessling E, Wingender E.** 2003. Prediction of potential C/EBP/NF-kappaB composite elements using matrix-based search methods. *In silico biology* **3**:71-79.
51. **Messeguer X, Escudero R, Farre D, Nunez O, Martinez J, Alba MM.** 2002. PROMO: detection of known transcription regulatory elements using species-tailored searches. *Bioinformatics (Oxford, England)* **18**:333-334.

52. Farre D, Roset R, Huerta M, Adsuara JE, Rosello L, Alba MM, Messeguer X. 2003. Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN. *Nucleic acids research* **31**:3651-3653.
53. Cuthill S, Sibbet GJ, Campo MS. 1993. Characterization of a nuclear factor, papilloma enhancer binding factor-1, that binds the long control region of human papillomavirus type 16 and contributes to enhancer activity. *Molecular carcinogenesis* **8**:96-104.

#### 4.3.7 Figures and Tables

**Frequency-Based Position Weight Matrix Logo  
(Jaspar MA0466.1)**



**Energy-Based Position Weight Matrix logo**

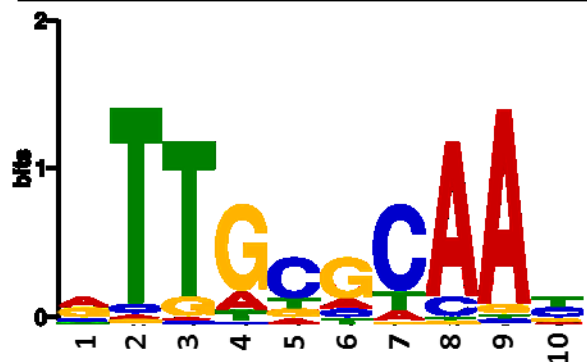
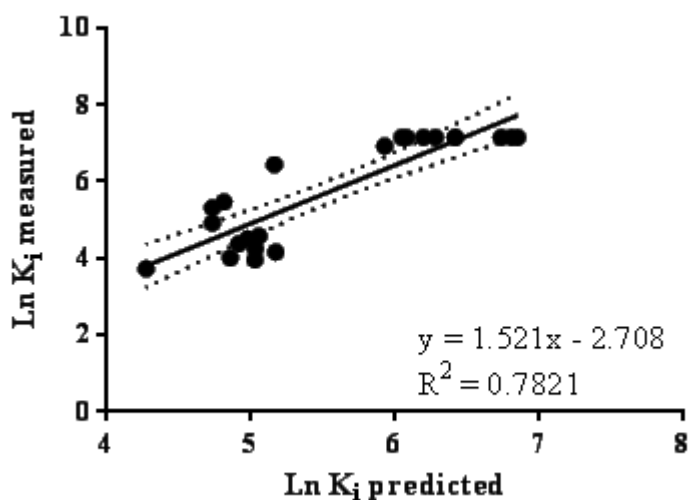
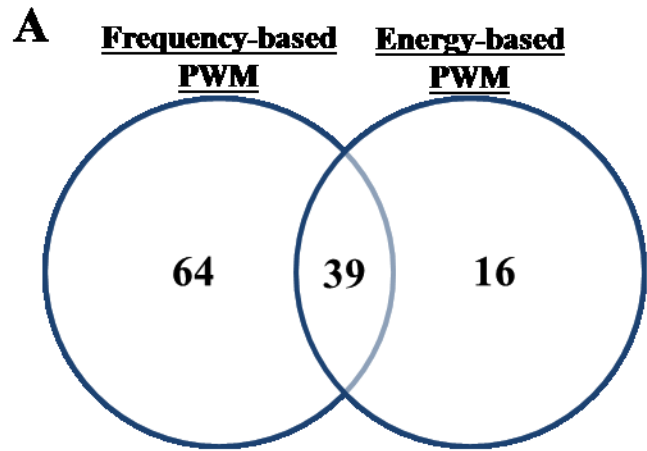


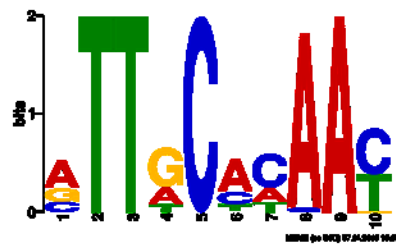
Figure 4.1. Sequence logo of the C/EBP $\beta$  consensus binding site as determined with the frequency- and energy-based position weight matrices.



**Figure 4.2. Correspondence of the measured and predicted binding affinities of C/EBP $\beta$  of C/EBP $\beta$  for different DNA sequences.** Each dot represents the values of the measured and predicted  $K_i$  values for one oligonucleotide. Permutations of every possible nucleotide were introduced in both half-site of the C/EBP $\beta$  consensus sequence (Der16 to 30). Previously characterized C/EBP $\beta$  sites in the LCR of HPV33 (33BS1, 33BS2 and 33BS3) and their mutants (33mBS1, T7791C, 33mBS2 and 33mBS3) were also included to the analysis as controls. The equation of the linear regression line and coefficient of determination ( $R^2$ ) are shown.



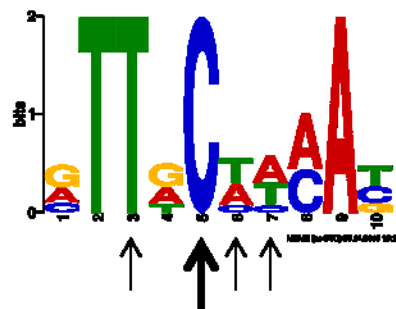
**B** Logo of the 39 common C/EBP $\beta$  sites



**C** Logo of the 16 Energy-based C/EBP $\beta$  sites



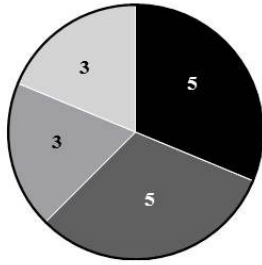
**D** Logo of the 64 Frequency-based (Jaspar) C/EBP $\beta$  sites



**Figure 4.3. Comparison of the sets of predicted C/EBP $\beta$  sites of frequency- and energy-based position weight matrixes.** A) Venn diagram showing the number of predicted C/EBP $\beta$  sites identified specifically with the frequency-based position weight matrix (PWM), the energy-based PWM or common to both sets. Sequence logo corresponding to the C/EBP $\beta$  sites predicted that are common (B) or specifically to the frequency-based PWM (C) or energy-based PWM (D). The arrows show the positions that are critical for binding of C/EBP $\beta$ .

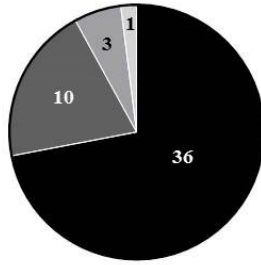
**A**

Mucosal HR-HPV types



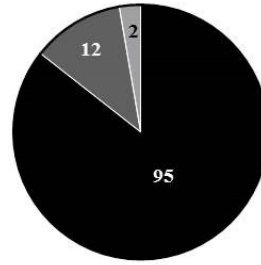
Total = 16 types

Mucosal LR-HPV types



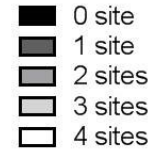
Total = 50 types

Cutaneous HPV types

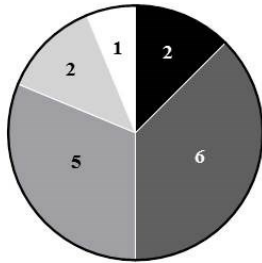


Total = 109 types

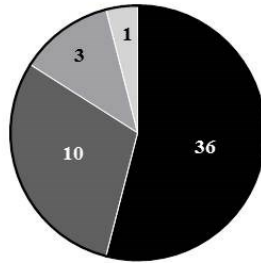
Energy-based prediction



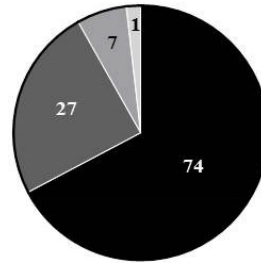
**B**



Total = 16 types

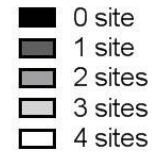


Total = 50 types

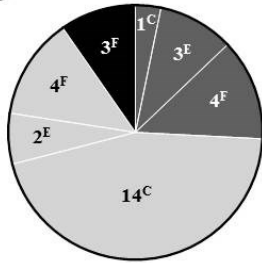


Total = 109 types

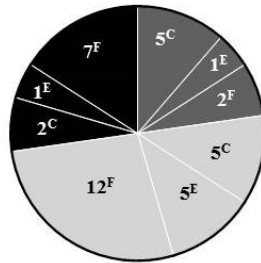
Frequency-based Prediction



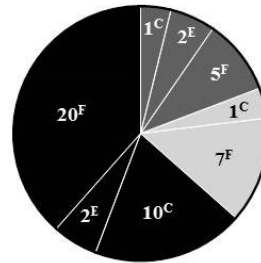
**C**



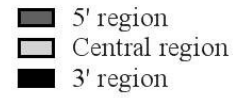
Total = 31 sites



Total = 40 sites



Total = 48 sites

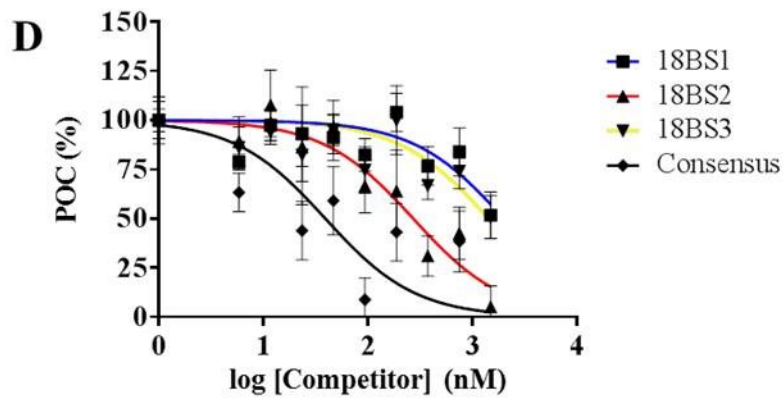
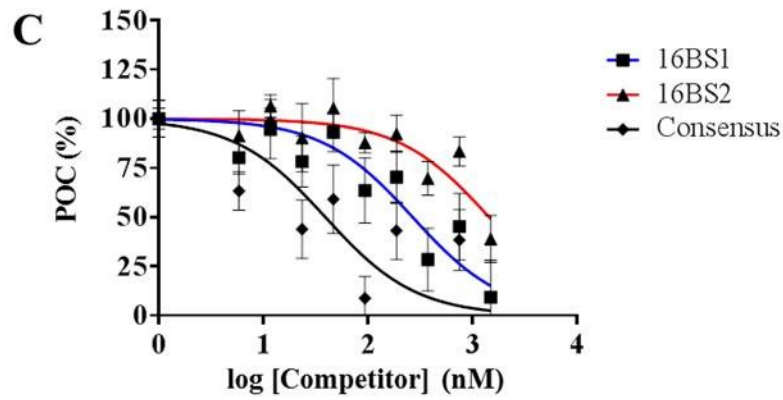
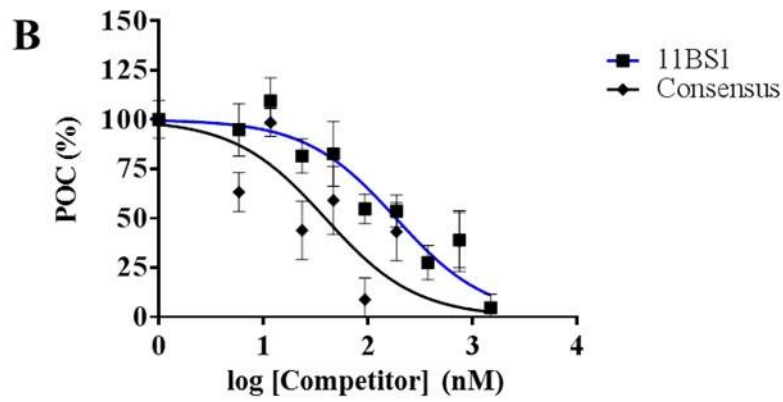
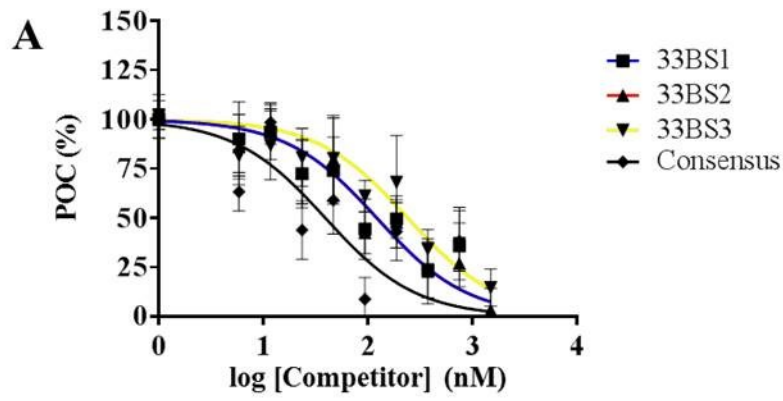


E: energy-PWM specific sites

F: frequency-PWM specific sites

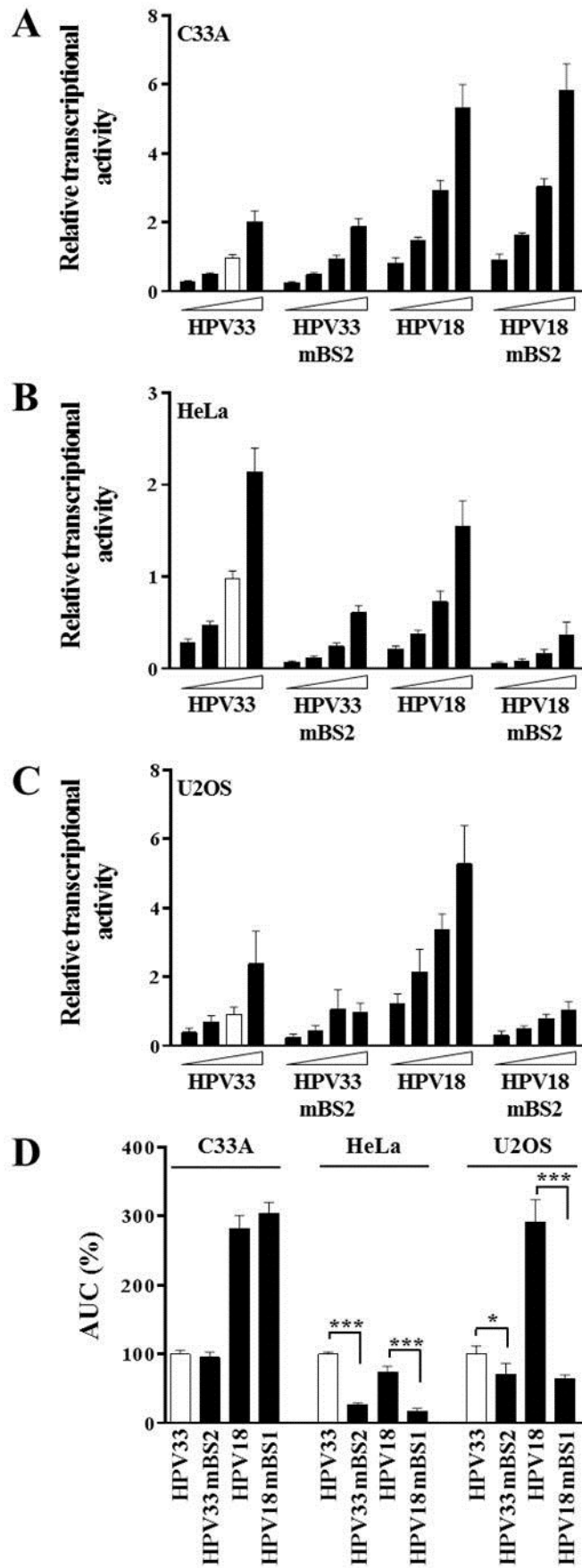
C: common sites

**Figure 4.4. Conservation of C/EBP $\beta$  binding sites in the LCR of known HPV types.** Quantification of the number of C/EBP $\beta$  binding sites found in the LCR of high-risk (HR), low-risk (LR) and cutaneous (Ct) HPVs as predicted with the frequency-bases PWM (A) or the energy-based PWM (B). The total number of HPV types classified in the HR, LR and Ct groups are indicated below each chart. C) Localization of C/EBP $\beta$  binding site in the LCR of different HPV types. The C/EBP $\beta$  binding site(s) of HPV types categorized in the HR, LR and Ct groups were further characterized according to their position in the LCR: the 5' region (delimitated from the beginning of the sequence to the first E2 binding site), the central region (between the first and the second E2 binding site) or in the 3' region (between the second E2 binding site to the end of the sequence). The number of predicted sites that are common (C), or specific to the frequency-based PWM (F) and the energy-based PWM (E) is indicated. The total number of sites is indicated below each chart.





**Figure 4.5. C/EBP $\beta$  binding sites in the LCR of HPV11, HPV16 and HPV18 show different degree of affinity for C/EBP $\beta$ .** Affinities of GST-C/EBP $\beta$ -BZ for different sites identified in the LCRs of HPV33 (A), HPV11 (BS), HPV16 (C) and HPV18 (D) as determined in competitive fluorescence polarization. Increasing amount (0, 5.9, 11.7, 23.4, 46.9, 93.8, 187.5, 375, 750 and 1500 nM) of unlabeled oligonucleotides (33BS1, 33BS2, 33BS3, 11BS1, 16BS1, 16BS2, 18BS1, 18BS2, 18BS3) were used to compete against a fluorescein-labeled 33BS1 probe (15 nM) for the binding of C/EBP $\beta$ -BZ (125 nM). The C/EBP $\beta$  binding site sequences of HPV33, HPV11, -16 and -18 are indicated in Table 5. Competitor concentration is indicated as the log<sub>10</sub> value. Polarization levels were normalized to the value at 0 ng of competitor which was set at 100% and are reported as a percent of control (POC). Each data point is the average of duplicates of at least 2 independent experiments. Error bars represent the standard deviation.



**Figure 4.6. The HPV18 BS2 is important for activation of the LCR in HeLa and U2OS cells.** The quadruple inactivating mutation (mBS2) was introduced into the HPV33 and HPV18 LCR-Rluc plasmids. The transcriptional activities of the resulting mutant LCRs were measured using increasing quantities of LCR-Rluc plasmid (25, 50, 100 and 200 ng) in C33A (A), HeLa (B) and U2OS (C) cells. The LCR-PT-Fluc plasmid (25 ng) was co-transfected as an internal control. Each bar represents the mean value obtained from a minimum of 4 independent experiments performed in duplicates. All values are presented relative to the level of transcriptional activity obtained with 100 ng of LCR-PT, which was used as the reference (white bar) and assigned a value of 1.0. Standard deviations are indicated by error bars. D) Integrated measure of LCR transcriptional activity using the area under the curve (AUC). AUC values were calculated from the dose-response curves presented in panels A-C and used as an integrated measure of LCR activity in C33A, HeLa and U2OS cells, as indicated. The AUC value of each LCR is reported as a percentage of the AUC value calculated for LCR-PT, which was set as 100. Statistical analysis of AUC values was performed using a one-way ANOVA followed by Dunnett's post-hoc analysis to compare the transcriptional activity of the prototypical LCR of HPV33 or HPV18 (white bar) with that of their respective mBS2 mutant. \* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$ .

## A BS2-like cluster

HPV180/7227-7227  
 HPV71/7818-7885  
 HPV90/8083-8149  
 HPV106/7726-7805  
 HPV83/7762-7829  
 HPV89/8215-8283  
 HPV61/7745-7810  
 HPV62/7751-7818  
 HPV81/7849-7915  
 HPV86/8126-8193  
 HPV84/7620-7687  
 HPV114/7950-8017  
 HPV42/7684-7747  
 HPV7/7805-7870  
 HPV67/7547-7598  
 HPV52/7682-7758  
 HPV33/7675-7730  
 HPV58/7591-7646  
 HPV44/7634-7679  
 HPV51/7534-7607  
 HPV82/7599-7675  
 HPV26/7602-7665  
 HPV39/7578-7655  
 HPV68/7462-7540  
 HPV85/7565-7641  
 HPV59/7594-7669  
 HPV97/7481-7557  
 HPV18/7599-7675  
 HPV45/7597-7673  
 HPV103/7003-7053  
 HPV65/7224-7271  
 HPV178/7154-7202  
 HPV8/7694-7736  
 HPV151/7478-7508  
 HPV22/7460-7490  
 HPV9/7529-7556  
 HPV108/7548-7592

## B BS1-like cluster

HPV180/7186-7227  
 HPV71/7885-7956  
 HPV90/8149-8220  
 HPV106/7805-7876  
 HPV83/7829-7898  
 HPV89/8283-8352  
 HPV61/7810-7880  
 HPV62/7818-7887  
 HPV81/7915-7984  
 HPV86/8193-8261  
 HPV84/7687-7755  
 HPV114/8017-8085  
 HPV42/7747-7818  
 HPV7/7870-7937  
 HPV67/7598-7668  
 HPV52/7758-7829  
 HPV33/7730-7800  
 HPV58/7646-7715  
 HPV44/7679-7750  
 HPV51/7607-7678  
 HPV82/7675-7746  
 HPV26/7665-7738  
 HPV39/7655-7724  
 HPV68/7540-7609  
 HPV85/7641-7711  
 HPV59/7669-7738  
 HPV97/7557-7628  
 HPV18/7675-7746  
 HPV45/7673-7744  
 HPV103/7053-7112  
 HPV65/7271-7338  
 HPV178/7202-7242  
 HPV8/7736-7765  
 HPV151/7508-7534  
 HPV22/7490-7516  
 HPV9/7556-7585  
 HPV108/7592-7651

**Figure 4.7. Distribution of C/EBP $\beta$  binding site(s) in the enhancer region of the LCR of different HPV types.** Representation of C/EBP $\beta$  binding site positions in the central region of the LCR of different HPV types according to the localization of the HPV33 BS2 (A) and HPV33 BS1 (B) sites. LCRs that contained at least one C/EBP $\beta$  binding site in the central region were aligned using Clustal Omega and Jalview. Nucleotides highlighted in green, purple and black represents the C/EBP $\beta$  binding sites that were predicted by the frequency-based PWM, the energy-based PWM or by both methods, respectively. Portions of the LCR that are represented are numbered according to the numbering in the PAVE database.

**Table 4.1. Effect of systematic permutations at each position of one half-site on the affinity of a C/EBP $\beta$  homodimer for its consensus binding sequence**

<b>Competitor DNA</b>	<b>Sequence (5'-3')</b>	<b>K<sub>i</sub> (nM)</b>	<b><math>\Delta\Delta G^*</math> (kcal/mol)</b>
<b>Consensus</b>	<b>ATTGCGCAAT</b>	38 $\pm$ 14	0
<b>Der1</b>	<b>C.....</b>	55 $\pm$ 9	0.939
<b>Der2</b>	<b>G.....</b>	40 $\pm$ 21	0.125
<b>Der3</b>	<b>T.....</b>	73 $\pm$ 7	1.612
<b>Der4</b>	<b>.C.....</b>	183 $\pm$ 34	3.879
<b>Der5</b>	<b>.G.....</b>	516 $\pm$ 120	6.446
<b>Der6</b>	<b>.A.....</b>	387 $\pm$ 178	5.736
<b>Der7</b>	<b>..C.....</b>	980 $\pm$ 425	8.036
<b>Der8</b>	<b>..G.....</b>	96 $\pm$ 11	2.289
<b>Der9</b>	<b>..A.....</b>	646 $\pm$ 222	7.006
<b>Der10</b>	<b>...C.....</b>	279 $\pm$ 144	4.926
<b>Der11</b>	<b>...T.....</b>	97 $\pm$ 9	2.307
<b>Der12</b>	<b>...A.....</b>	75 $\pm$ 45	1.689
<b>Der13</b>	<b>....G.....</b>	78 $\pm$ 2	1.785
<b>Der14</b>	<b>....T.....</b>	69 $\pm$ 13	1.489
<b>Der15</b>	<b>....A.....</b>	85 $\pm$ 7	2.005

Note. The K<sub>i</sub> values were determined as described in the Materials and Methods section and using 125 nM of a GST- C/EBP $\beta$ -BZ fusion protein.  $\Delta G(x)$  is the binding energy required for C/EBP $\beta$  to interact with a given DNA target. G(ref) is the binding energy required for C/EBP $\beta$  to interact with the consensus sequence. \* Relative to the consensus sequence which was set at 0.

**Table 4.2. Binding predictor matrix definition**

<b>Position</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>
<b>Consensus</b>	<b>A</b>	<b>T</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>T</b>
<b>A</b>	<b>0.00</b>	<b>5.74</b>	<b>7.01</b>	<b>1.69</b>	<b>2.00</b>	<b>1.49</b>	<b>2.31</b>	<b>0.00</b>	<b>0.00</b>	<b>1.61</b>
<b>C</b>	<b>0.94</b>	<b>3.88</b>	<b>8.04</b>	<b>4.93</b>	<b>0.00</b>	<b>1.78</b>	<b>0.00</b>	<b>2.29</b>	<b>6.45</b>	<b>0.13</b>
<b>G</b>	<b>0.13</b>	<b>6.45</b>	<b>2.29</b>	<b>0.00</b>	<b>1.78</b>	<b>0.00</b>	<b>4.93</b>	<b>8.04</b>	<b>3.88</b>	<b>0.94</b>
<b>T</b>	<b>1.61</b>	<b>0.00</b>	<b>0.00</b>	<b>2.31</b>	<b>1.49</b>	<b>2.00</b>	<b>1.69</b>	<b>7.01</b>	<b>5.74</b>	<b>0.00</b>

Note. The numbers refer to the  $\Delta\Delta G$  (kcal/mol) relative to the nucleotide of the consensus sequence and were determined as described in the Materials and Methods section.

**Table 4.3. Correspondence of measured and predicted  $K_i$  of different DNA sequences for C/EBP $\beta$**

Competitor	Sequence (5'-3')	$K_i$ measured (nM)	$K_i$ predicted (nM)	$\Delta\Delta G^*$ (kcal/mol)
Consensus	ATTGCGCAAT	35 ± 5	38	0
Der16	C . . . . . G	41 ± 29	72	0.392
Der17	G . . . . . C	52 ± 17	96	0.980
Der18	T . . . . . A	235 ± 213	123	4.718
Der19	.C . . . . .G.	> 1269	615	> 8.896
Der20	.G . . . . .C.	> 1269	493	> 8.896
Der21	.A . . . . .T.	> 1269	439	> 8.896
Der22	..C . . . .G..	> 1269	615	> 8.896
Der23	..G . . . .C..	624 ± 367	175	7.137
Der24	..A . . . .T..	> 1269	536	> 8.896
Der25	...C . .G...	> 1269	377	> 8.896
Der26	...T . .A...	70 ± 109	176	1.717
Der27	...A . .T...	55 ± 9	129	1.109
Der28	....GC....	79 ± 1	137	2.021
Der29	....TA....	136 ± 128	114	3.368
Der30	....AT....	70 ± 89	153	1.709
33BS1	C . .A .A . . .	97 ± 2	157	2.524
33BS2	....AA...	90 ± 26	145	2.352
33BS3	....TA....	203 ± 80	114	4.355
33mBS1	CCCA .A .CC .	> 1269	947	> 8.896
33mBS2	.CC . .AACC .	> 1269	912	> 8.896
33mBS3	.CC .TA .CC .	> 1269	904	> 8.896
T7791C	C .AA .A . . .	> 1269	425	> 8.896

Note. The measured  $K_i$  values were determined as described in the Materials and Methods section and using 125 nM of a GST- C/EBP $\beta$ -BZ fusion protein.  $\Delta G(x)$  is the binding energy required for C/EBP $\beta$  to interact with a given DNA target.  $G(\text{ref})$  is the binding energy required for C/EBP $\beta$  to interact with the consensus sequence. \* Relative to the consensus sequence which was set at 0.



**Table 4.4. Functions of known C/EBP $\beta$  sites**

HPV types	Sequence (5'-3')	Region	Functions	References
HPV33 BS1	CTTACACAAT	7789-7798	- Binding assessed in competition FPA - - Represses the LCR in C33A, HeLa and U2OS cells	Alvarez et al. (2016)
HPV33 BS2	ATTGCAAAAT	7700-7709	- Binding assessed in competition FPA - Activates the LCR in HeLa cells	Alvarez et al. (2016)
HPV33 BS3	ATTGTACAAT	7471-7480	- Weak binding in competition FPA - Not active in C33A, HeLa and U2OS cells	Alvarez et al. (2016)
HPV11	TTTGTGCAAT	7451-7459	- Binding assessed by EMSA - Represses the LCR through a promoter proximal YY1 site in HFKs	Auborn and Steinberg (1991) Ralph et al. (2006)
HPV16	nd	7285 to 7315	- Binding assessed by EMSA - Represses the LCR in HeLa and CaSki cells	Kyo et al. (1996)
HPV18 BS1 (switch region)	TTTACTTAAG	7709-7718	- Binding assess by EMSA - Activates the LCR through the formation of a YY1-C/EBP $\beta$ complex in HeLa cells	Bauknecht et al. (1996)

nd: specific site(s) not determined

FPA: fluorescence polarization assay

EMSA: electrophoretic mobility shift assay

**Table 4.5. Affinities ( $K_i$ ) of C/EBP $\beta$  homodimer for previously described binding sites**

<b>Competitor</b>	<b>References</b>	<b>Sequence (5'-3')</b>	<b><math>K_i</math> (nM)</b>
<b>Consensus</b>	Williams et al. (1995)	ATTGCGCAAT	$33 \pm 2$
<b>33BS1</b>	Alvarez et al. (2016)	CTTACACAAT	$103 \pm 26$
<b>33BS2</b>	Alvarez et al. (2016)	ATTGCAAAAT	$102 \pm 36$
<b>33BS3</b>	Alvarez et al. (2016)	ATTGTACAAT	$203 \pm 80$
<b>11BS1</b>	Ralph et al. (2006)	TTTGTGCAAT	$156 \pm 28$
<b>16BS1</b>		CTTGTGTAAC	$226 \pm 195$
<b>16BS2</b>		TTGTTTCAAC	$> 1269$
<b>18BS1</b>	Baucknecht et al. (1996)	TTTACTTAAG	$1208 \pm 163$
<b>18BS2</b>		CTTGCATAAC	$229 \pm 156$
<b>18BS3</b>		CTTGTACAAC	$> 1269$

Note. The measured  $K_i$  values were determined as described in the Materials and Methods section and using 125 nM of a GST- C/EBP $\beta$ -BZ fusion protein.

## Chapitre cinquième : Discussion

L'objectif général de ce projet de recherche était de caractériser les fonctions des polymorphismes du VPH33 sur l'activité transcriptionnelle du LCR. Plusieurs conclusions peuvent être formulées à partir des résultats présentés dans les chapitres précédents. Tout d'abord, les résultats du chapitre 2 indiquent que la phylogénie des variantes corrèle avec l'activité de transcription i.e. que les variantes au sein d'une même lignée présentent des activités transcriptionnelles similaires. Plus important encore, certaines variantes non-prototypiques, notamment celles de la sous-lignée A2, montrent une activité du LCR accrue dans les cellules C33A. L'augmentation de l'activité transcriptionnelle de ces variantes est attribuable à la combinaison de plusieurs variations, dont certaines sont caractéristiques de la sous-lignée A2. Fait intéressant, nos études antérieures ont indiqué qu'une de ces variations, C7732G, est associée à un risque plus élevé de développer un HSIL (72, 701). Une deuxième variation, soit une délétion de 79 pb, est associée à la persistance de l'infection (72, 701). Cette délétion, qui n'est pas spécifique à la sous-lignée A2, est également retrouvée dans les variantes de la lignée B, mais pas dans la sous-lignée prototypique A1. La délétion n'altère pas l'activité transcriptionnelle du LCR du VPH33 à elle seule, mais en combinaison avec C7732G, elle augmente l'activité promotrice de 50%. Toutefois, ces deux variations ne sont pas suffisantes pour augmenter l'activité du prototype au niveau de l'activité de la variante LCR6 de la sous-lignée A2. L'activité du LCR6 est obtenue dans le LCR-PT lorsque les variations C7732G, C7537A et A7879G et la délétion sont introduites conjointement. Cette étude indique que l'activité du LCR est modulée par plusieurs variations dont l'effet individuel est plutôt modeste.

Néanmoins, une variation dans un site critique pour la régulation de l'expression génique peut aussi exercer un impact significatif sur le LCR. Ainsi, les résultats du chapitre 3 indiquent que la variante rare de la lignée B, LCR10, montre une activité transcriptionnelle de 2-3 fois plus élevée que celle du prototype dans les cellules C33A, HeLa et U2OS. La variation T7791C spécifique à LCR10 cause la dérégulation du LCR en abolissant un site de liaison pour C/EBP $\beta$  (BS1), indiquant que C/EBP $\beta$  est un répresseur important de l'expression génique du VPH33. Dans cette même étude, nous avons identifié un deuxième site C/EBP $\beta$  (BS2) qui, cette fois-ci, est impliqué dans l'activation du LCR du VPH33. Ainsi, le VPH33 porte des sites activateurs et répresseurs sur la même région

promotrice. Alors que BS1 induit la répression du LCR dans les cellules C33A, HeLa et U2OS, BS2 n'est actif que chez ces deux dernières lignées, bien que l'effet dans les U2OS soit modeste. Ces résultats dénotent que les fonctions de C/EBP $\beta$  sont également dictées par le contexte cellulaire et donc par les différents interactomes des cellules étudiées.

Finalement, dans le quatrième chapitre de cette thèse, nous montrons que la modulation par C/EBP $\beta$  de l'expression des gènes viraux est un processus conservé chez la plupart des VPHs muqueux, en particulier les VPHs-HR. Effectivement, nous notons la présence de sites pour des homodimères de C/EBP $\beta$  qui sont similaires aux sites BS1 et BS2 retrouvés dans le LCR du VPH33. Alors que quelques études indiquent que C/EBP $\beta$  est également impliqué dans la régulation de l'expression génique du VPH11, 16, 18 et 31, nous montrons que C/EBP $\beta$  lie des sites différents que ceux préalablement identifiés dans le VPH18. Nos prédictions qui se basent sur l'utilisation de matrices des énergies de liaison et de fréquence permettent d'identifier les sites à haute affinité pour des homodimères de C/EBP $\beta$ .

## **5.1 Modulation de l'expression génique du VPH33 par les variations dans le LCR**

Notre étude est la première à caractériser l'effet individuel des variations du LCR du VPH33 sur la transcription virale par une approche de mutagenèse détaillée. Les quelques études qui se sont penchées sur la question ont utilisé des LCRs chimères de variantes ou des constructions contenant des délétions, ce qui rend difficile l'identification des polymorphismes qui sont associés à un changement de phénotype (711, 740). Nos résultats indiquent que certaines variations dans le LCR augmentent l'expression génique virale et pourraient donc altérer l'histoire naturelle de l'infection au VPH33. Ainsi, le potentiel oncogénique de certaines variantes du VPH33 serait plus grand que celui du prototype.

### **5.1.1 Le LCR est représentatif de la variabilité génomique du VPH33**

La grande hétérogénéité des VPHs complique la comparaison des études d'association des variantes avec un risque plus élevé de CC, d'autant plus qu'il n'existait

pas jusqu'à tout récemment un système standardisé pour la nomenclature des isolats d'un virus donné. L'établissement d'une classification des variantes en lignées facilite ce processus. Cette classification se base sur la comparaison des séquences nucléotidiques des génomes entiers des virus et non plus selon une région génomique virale en particulier (5). Des combinaisons de polymorphismes se fixent au cours de l'évolution et deviennent des caractéristiques propres à une lignée donnée. Ces variations ne sont pas distribuées également dans le génome viral, certaines régions étant plus variables (5). De par sa nature non-codante, le LCR, qui est moins contraint à la pression évolutive par rapport aux gènes, acquiert donc plus de modifications dans le temps.

Notre première étude se base sur la comparaison de la séquence nucléotidique du LCR de plusieurs isolats du VPH33. Malheureusement, au moment de nos analyses, le séquençage du génome entier n'avait pas été réalisé. Par contre, notre première étude indique que la séquence du LCR est suffisante pour attribuer chaque variante à une lignée. Ainsi, en n'utilisant que les séquences en ADN du LCR des variantes qui ont été identifiées par notre groupe et par d'autres, nous obtenons un arbre phylogénétique présentant les embranchements majeurs A et B, avec le premier groupe sous-divisé en deux sous-lignées A1 et A2 (chapitre 2, Fig. 2.1), tel que décrites par Chen et al. (5). Plus récemment, deux variantes rares de E6 du VPH33 ont été identifiées ayant 3 et 4 variations de plus que leurs variantes les plus proches et ont donc été classifiées dans la sous-lignée A3 et la lignée C, respectivement (745). Malheureusement, les LCRs de ces variantes n'ont pas été séquencés et n'ont donc pas été inclus dans notre analyse phylogénétique. Ces variantes sont retrouvées uniquement en Asie et en Afrique. Ainsi, la présence d'au moins trois polymorphismes additionnels dans une région donnée est parfois suffisante pour caractériser une nouvelle sous-lignée.

Parmi les variantes que nous avons caractérisées, le LCR10 diverge de trois nucléotides en plus d'être dépourvu du polymorphisme à la position 18, par rapport aux autres variantes de la lignée B (chapitre 3, Fig. 3.1). De plus, parmi les variantes de la lignée B identifiée ailleurs dans le Monde, LCR10 reste la plus polymorphique de toutes avec un total de 19 variations. Ainsi, l'accumulation de polymorphisme plus élevée de la variante LCR10 suggère une divergence plus récente de cette variante de l'ancêtre commun (*most recent common ancestor*) et serait possiblement un indicatif de la présence d'une

nouvelle sous-lignée dans le groupe B. Toutefois, il est important de noter que la variante LCR10 n'a été retrouvée que dans une seule patiente des 95 femmes recrutées dans l'étude. Bien que cette variante ait été isolée chez plusieurs spécimens consécutifs de cette femme d'origine haïtienne, l'infection par le VPH33 était résolue à la fin du suivi. Il est possible que cette variante soit défavorisée par rapport aux autres due à la variation qui affecte le site pour C/EBP $\beta$  ou bien qu'elle soit nouvellement émergente, ce qui expliquerait qu'elle ne soit pas encore répandue dans la population. Il serait donc intéressant d'un point de vue évolutif d'étudier la prévalence de cette variante dans d'autres populations.

### **5.1.2 Association entre les variantes non-prototypiques et le développement de cancer du col de l'utérus**

Une grande partie de la littérature indique que certaines variantes non-prototypiques, en particulier les variantes AA et NA du VPH16, sont plus oncogéniques de par leur association avec la progression vers le CC (58, 702, 716, 731, 733, 736). Le risque de progression pour les variantes moins prévalentes est moins connu, mais furent abordées de façon plus exhaustive dernièrement (72, 709, 740, 745, 753). Dans un article récent, le risque de CC pour les variantes de E6 et E7 du VPH33 a été évalué (745). Les variantes A2 sont rarement détectées dans des spécimens provenant de l'Afrique et de l'Amérique du Sud. Les variantes des lignées A3 et C sont davantage isolées en Asie tandis que celles de la sous-lignée A1 sont retrouvées partout dans le monde, avec une préférence en Amérique du Sud. De plus, les variantes de la sous-lignée A1 sont surreprésentées dans les cas de CC en Asie et en Afrique. Nos études dénotent que les variantes du VPH33 ne sont pas associées avec l'ethnicité des participantes, alors que les échantillons cliniques proviennent de Montréalaises et de Brésiliennes, ce qui est également observé dans d'autres études, en Slovénie notamment (754, 755). De plus, nos résultats suggèrent que ce sont plutôt les variantes de la sous-lignée A2 qui sont associées avec un risque plus élevé de progression. Ainsi, des études antérieures indiquent que les variantes contenant la variation C7732G et une délétion de 79-bp étaient associées à un plus grand risque de développer des lésions de haut-grade et d'infection persistante, respectivement (72, 701). Par contre, ces articles ont analysé des variations de façon individuelle. Nous savons maintenant que les variations

tendent à se fixer dans une lignée donnée et donc qu'une variation est associée avec d'autres polymorphismes (2). Malheureusement, dû au faible nombre de spécimens positifs pour le VPH33, des études d'association entre les différentes lignées du VPH33 n'ont pu être effectuées. Malgré tout, nos études indiquent que les variantes de la sous-lignée A2 ont un potentiel d'oncogénicité accru par rapport au prototype.

Les différences entre ces deux études épidémiologiques pourraient s'expliquer par les différentes populations étudiées. Notre étude a analysé des échantillons provenant de femmes Canadiennes et Brésiliennes tandis que la première étude a analysé des spécimens provenant d'une trentaine de pays, mais seulement six Brésiliennes et aucune de Canadienne ou Nord-Américaine (745). Il est difficile de faire des analyses d'association avec l'ethnicité brésilienne du fait de la grande hétérogénéité de cette population. Bien que le nombre de femmes infectées par le VPH33 dans les études soit faible, elles sont les études les plus puissantes à date portant sur le VPH33. D'autres études et possiblement une méta-analyse permettraient de confirmer l'association des variantes du VPH33 avec un risque accru de développer le CC dans d'autres populations. Toutefois, nos études suggèrent que les variantes de la sous-lignée A1 ont généralement une activité transcriptionnelle plus faible que celles de la lignée A2 dans les cellules C33A et HeLa à l'exception du prototype qui est plus actif dans le dernier type cellulaire.

La détection des variations caractéristiques des lignées ayant un potentiel oncogénique plus élevé pourrait servir comme outil de diagnostic pour identifier les patientes ayant un plus grand risque de progression vers le CC. Ainsi, dans un programme de dépistage du CC, il serait intéressant de suivre davantage la progression de l'infection chez les femmes infectées avec des variantes oncogéniques tels que les variantes non-prototypiques du VPH16 et VPH18 et des variantes de la sous-lignée A2 du VPH33 si l'association avec un risque accru de progression vers un HSIL s'avère fondée. Bien sûr, les prochaines études d'association devront se basées sur un plus grand nombre d'échantillonnages, mais devront davantage discriminer les variantes selon leur potentiel oncogénique.

### 5.1.3 Activité transcriptionnelle élevée des variantes non-prototypiques

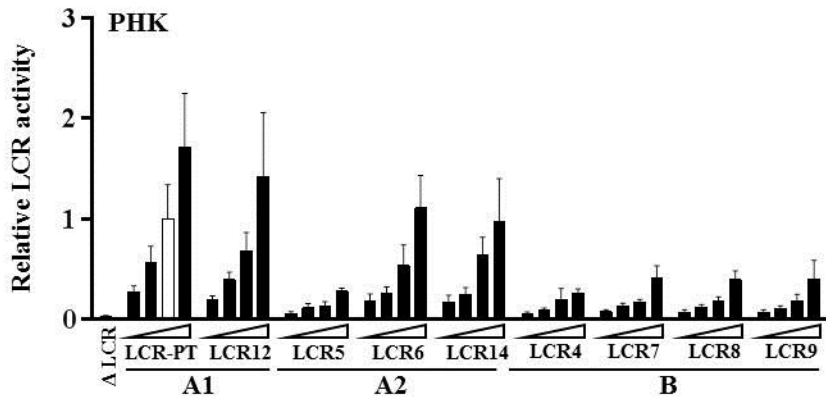
La majorité des études sur les variantes du VPH16 et VPH18 indiquent que les variantes non-prototypiques sont souvent associées à un risque plus grand de développer un CC et ceci corrèle assez bien avec des activités transcriptionnelles et répliquatives plus élevées (615, 711). Nos études antérieures suggèrent une association entre les variations définissant la sous-lignée A2 et un potentiel oncogénique accru. Nos études actuelles montrent également que l'activité transcriptionnelle de ces variantes est accrue par rapport au prototype (chapitre 2, Fig. 2.2). De plus, nous montrons que la variante LCR10 de la lignée B est également plus active autant dans les cellules C33A, HeLa et U2OS que le LCR-PT (chapitre 3, Fig. 3.1). Ainsi, nos études sont en accord avec la littérature i.e. que certaines variantes non-prototypiques du VPH33 ont une activité transcriptionnelle plus élevée que celle du prototype. L'activité de certains de ces LCR dépend du contexte cellulaire alors que les variantes A2 sont davantage activées dans les cellules C33A que dans les cellules HeLa. La plupart des études fonctionnelles sur les variantes n'ont étudié la transcription que dans un type cellulaire (604, 696, 697, 711, 717).

Nos études indiquent qu'il est important de mesurer l'activité du LCR dans plus d'une lignée cellulaire. Ainsi, l'effet de certaines variations, à première vue silencieuses dans un type de cellules, pourrait se manifester dans une lignée cellulaire distincte. Nous avons utilisés dans cette première étude des cellules C33A (VPH-négatives) et HeLa (VPH-positive). Ainsi, deux facteurs principaux peuvent moduler l'expression génique virale de façon différentielle dans les cellules C33A et HeLa. Tout d'abord, ces cellules sont dérivées d'un carcinome et d'un adénocarcinome du col de l'utérus, ce qui indique qu'elles ont acquis des mutations dans l'ADN génomique les immortalisant et les transformant. Alors que les cellules HeLa sont bien caractérisées, tel que montré par le répertoire large de publications utilisant cette lignée, moins est connu sur les caractéristiques génétiques des cellules C33A. Il est possible que certaines voies signalétiques soient davantage activées ou réprimées dans les cellules HeLa que dans les C33A. Dans cette optique, il serait intéressant de comparer le transcriptome de ces deux types cellulaires et de regarder si l'expression des gènes des facteurs de transcription, dont les sites potentiels sont affectés par les variations identifiées dans notre étude, est altérée. Deuxièmement, contrairement aux cellules C33A, les cellules HeLa ont été immortalisées par l'intégration du VPH18 et

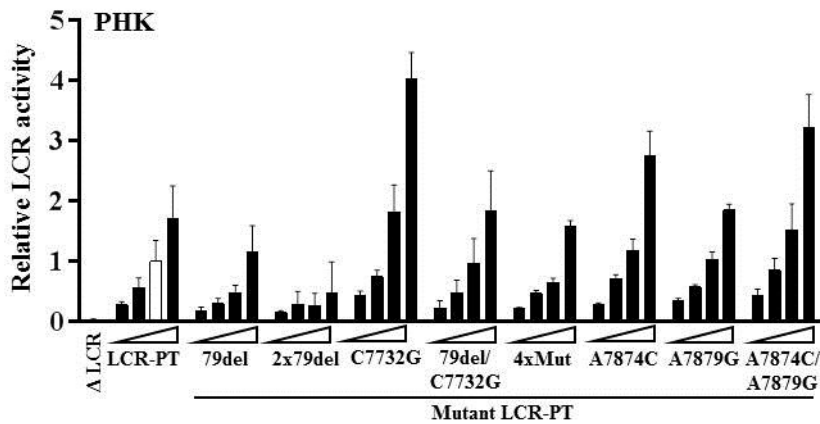


donc par la surexpression des oncogènes E6 et E7. Nous savons que les oncoprotéines encodées par ces gènes affectent plusieurs procédés cellulaires (voir section 1.2.4 de cette thèse). Ainsi, nous ne pouvons exclure un rôle additionnel de E6 et E7 dans la régulation de l'expression des gènes viraux dans les cellules HeLa comparé aux cellules C33A.

Toutefois, nous notons que les cellules C33A, dont le génome a acquis une mutation dans *TP53* entre autre, sont un bon modèle d'étude pour la réplication de l'ADN viral tel que montré dans certaines études dans notre laboratoire (230, 231, 385). La capacité de ces cellules à répliquer un plasmide contenant l'ori viral en cis indique que la machinerie de transcription des gènes viraux est fonctionnelle aussi. Néanmoins, il reste préférable de tester les LCRs dans un contexte plus pertinent pour l'étude de l'infection au VPH i.e dans des kératinocytes primaires. À la figure 5.1, nous montrons que bien que les variantes du LCR sont pour la plupart moins actifs que le prototype dans des kératinocytes primaires, les variantes LCR6 et LCR14 de la sous-lignée A2 sont plus actives que le LCR5. L'effet d'amplification de la variation C7732G sur l'activité du LCR est évident dans ces cellules en comparant la variante LCR6 à la variante LCR5 qui est dépourvu de la variation C7732G. De plus, l'introduction de variations, en particulier C7732G et A7874C, augmente l'activité transcriptionnelle du LCR-PT tel que montré à la figure 5.2.



**Figure 5.1** Activité transcriptionnelle du LCR prototypique et des variantes dans des kératinocytes primaires. L'activité transcriptionnelle du LCR prototypique et celle des variantes a été mesurée dans les kératinocytes primaires dans un essai de transcription transitoire tel que décrit dans (705).



**Figure 5.2** Influence des variations nucléotidiques sur l'activité transcriptionnelle du LCR du VPH33 dans des kératinocytes primaires. Les variations indiquées dans la figure ont été introduites dans le LCR prototype par mutagenèse dirigée. L'activité transcriptionnelle du LCR prototypique et celle des mutants ont été mesurées dans des kératinocytes primaires dans un essai de transcription transitoire tel que décrit dans (705). Le mutant 4xMut contient les variations C7537A, 79del, C7732G et A7879G.

Des études supplémentaires sont nécessaires afin de déterminer si la hausse de l'activité du LCR des variantes A2 se traduira en un accroissement de la réplication. Malheureusement, l'utilisation de ces LCRs dans un système de réplication de l'ADN transitoire ne démontra pas de différence significative entre les variantes et le prototype (données non montrées ici). Cet essai, qui a été développé à partir d'un protocole élaboré dans notre laboratoire (voir (231) et en annexe de cette thèse), ne permet que la réplication d'un plasmide contenant l'ori et non le génome au complet. L'expression *in trans* des protéines de réplication E1 et E2 est notamment contrôlée par le promoteur fort du CMV et non par le promoteur du VPH qui est d'emblée plus réprimé. Ainsi, une augmentation de l'expression de E1 et E2 due à une activité accrue du LCR ne serait observable que dans un contexte où le génome entier est utilisé tel que dans des cultures organotypiques en radeau de cellules épithéliales où la différenciation cellulaire peut être induite. Toutefois, des études effectuées sur les variantes NA et AA du VPH16 montrent qu'une augmentation de l'activité du LCR d'un ordre de 2-3 fois résulte en une hausse de la réplication virale de 20 fois par rapport au prototype (714), indiquant qu'une augmentation modeste des niveaux de transcription augmente de manière significative le nombre de copies de l'épisome viral. Ainsi, il est très probable que les variantes ayant une activité du LCR accrue par rapport au prototype, tel que les variantes LCR6, LCR14 et LCR10, corrèlent avec une réplication plus active due à l'augmentation des niveaux de protéines de réplication virales E1 et E2.

#### **5.1.4 Influence des variations nucléotidiques faibles sur l'activité transcriptionnelle du LCR**

La plupart des variations modulant l'activité transcriptionnelle des variantes du LCR de la sous-lignée A2 dans les cellules C33A affectent des sites d'interaction avec des FT dont le rôle dans la transcription virale n'a jamais été documenté. Ainsi, les polymorphismes C7537A et A7879G affectent des sites potentiels pour USF et pour NF-Y respectivement, et seraient responsable individuellement d'environ 25% de l'activité transcriptionnelle (chapitre 2, Fig. 2.5). Les fonctions de USF et NF-Y dans la modulation de l'expression des gènes viraux reste à être établie par d'autres analyses fonctionnelles et

essais de liaison à l'ADN. Toutefois, ces facteurs présentent souvent des rôles d'activateur et de répresseur de la transcription cellulaire (756-759). Les effets différents de la perte de ces sites dans les cellules C33A et HeLa reflètent la double fonction de ces FT qui serait déterminée par le contexte cellulaire. Ainsi, l'altération de sites d'interaction pour ces facteurs permettrait la modulation plus fine de l'expression génique. Nous notons que la variation A7879G peut également affecter un site Cre. Alors que l'effet de la mutation de ce site potentiel reste modeste dans les C33A, A7879G est responsable de la diminution considérable de l'activité de la variante LCR6 dans les cellules HeLa comparativement au niveau obtenu dans les cellules C33A (chapitre 2, Fig. 2.5). Par la nature répressive de cette variation, les variantes A2 sont moins actives dans les cellules HeLa que dans les cellules C33A. Toutefois, la réversion de cette variation à la séquence retrouvée dans le prototype rétablit les niveaux de transcription dans les cellules HeLa au-delà de ceux obtenus dans les C33A.

Individuellement, les variations dites faibles affectent peu l'activité transcriptionnelle du LCR tel qu'observé chez les variantes de la sous-lignée A2 dans les C33A. C'est plutôt la combinaison de ces polymorphismes qui impacte de façon significative l'activité transcriptionnelle des variantes (chapitre 2, Fig. 2.6). Plusieurs de ces variations, qui sont à première vue silencieuses, peuvent accentuer les effets des variations faibles quand elles sont combinées. Un exemple est la délétion de 79 pb qui, lorsqu'introduite dans le LCR prototypique, n'affecte pas l'activité transcriptionnelle, probablement par le rôle redondant de la deuxième région qui est intacte (chapitre 2, Fig. 2.3). Toutefois, lorsqu'elle est en combinaison avec la variation C7732G, la délétion potentialise l'activité de cette dernière. La délétion complète de la duplication de 79 pb réduit l'activité du LCR d'environ 20-30%, indiquant un rôle d'activateur pour cette région qui contient plusieurs sites de liaison pour NF1, NF-Y et Oct-1. La variation C7732G est suffisante pour augmenter l'activité du LCR de 25% en abolissant un site potentiel pour USF. Toutefois, le polymorphisme C7732T à la même position chez la variante LCR12 n'affecte pas l'activité transcriptionnelle du LCR (chapitre 2, Fig. 2.4), indiquant que seul un G à la position 7732 active la transcription virale.

Des études récentes montrent que les génomes viraux des VPH16 et 18 sont organisés sous forme de chromatine associée aux nucléosomes dont un est positionné sur la

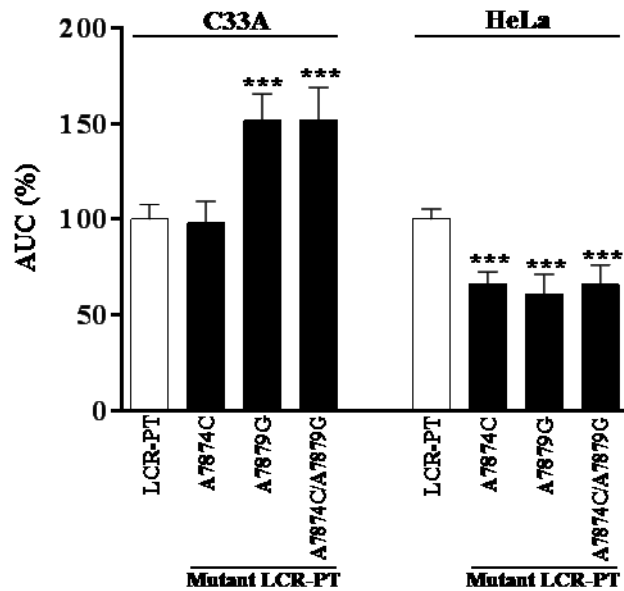
région centrale du LCR et serait conservé chez les PVs (176, 398, 511). Ce nucléosome serait présent dans une région similaire chez le VPH33 et chevaucherait la duplication de 79 pb tel que suggéré par le programme de prédiction du positionnement des nucléosomes (NuPop, <http://nucleosome.stats.northwestern.edu/>). Il est donc possible que la délétion de 79 pb altère le positionnement du nucléosome, permettant l'accès à des sites autrement fermés. Dans le cas des variantes de LCR de la lignée A2, l'altération du site pour USF par C7732G et la délétion de 79 bp pourrait permettre le recrutement d'un autre FT ou d'un coactivateur tel que CBP/p300, ce qui expliquerait l'effet additif des deux variations sur l'activité transcriptionnelle du LCR du VPH33. Dans l'ensemble, nos résultats montrent que la combinaison de variations faibles est suffisante pour augmenter significativement l'activité transcriptionnelle du LCR. Ces variations pour la plupart caractérisent la sous-lignée A2 et contribueraient donc à accroître le potentiel oncogénique de variantes.

### **5.1.5 Influence des variations nucléotidiques fortes sur l'activité transcriptionnelle du LCR**

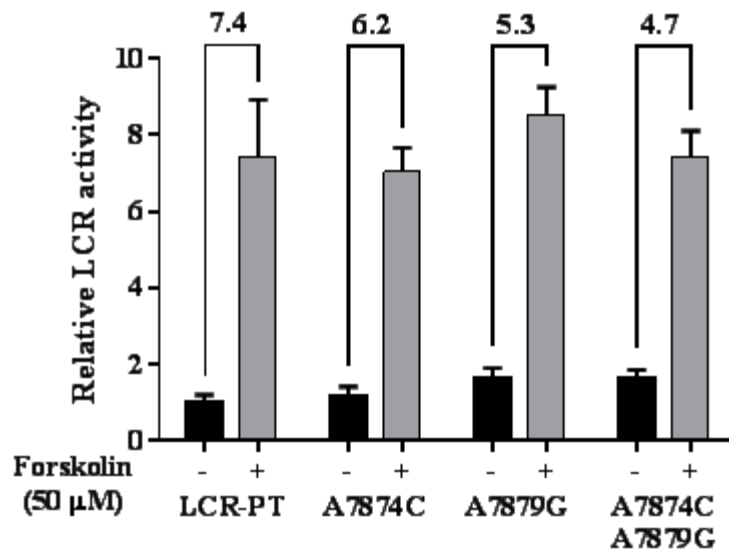
Nos études ont identifié trois variations dites fortes qui, individuellement, altèrent significativement l'activité transcriptionnelle du LCR. Il s'agit des polymorphismes T7791C du LCR10, A7874C du LCR12 et A7879G de la sous-lignée A2 (chapitre 2, Fig. 2.5; chapitre 3, Fig. 3.1). Ces deux dernières variations semblent avoir le même effet activateur dans les cellules HeLa. L'ajout de ces deux polymorphismes dans le LCR-PT ne potentialise pas cet effet, ce qui indique que ces deux variations affectent le même élément régulateur dans les cellules HeLa (voir Fig 5.3). Des analyses plus exhaustives de la séquence indiquent que ces deux variations peuvent affecter un site AP1, CRE et même la liaison avec E2. L'ajout du répresseur viral E2 et l'utilisation de TPA altèrent peu l'activité transcriptionnelle du LCR-PT et celle d'un mutant LCR-PT où A7874C et A7879G ont été introduits seuls ou en combinaison (donnée non montrée ici) suggérant que ces deux variations n'affectent pas un site AP1 ou pour E2. L'utilisation de la forskolin, de roflumilast et de rolipram (des inducteurs de l'AMPC) stimule davantage le LCR-PT que le double mutant (A7874C/A7879G) alors que les inhibiteurs (H89 et KT5720) n'affectent pas l'activité du LCR dans les cellules C33A, ce qui supporte l'hypothèse que les deux

variations affectent probablement un site Cre (voir Fig. 5.4 et 5.5). De plus, il semblerait que la voie de l'AMPc ne soit pas *de facto* active dans les cellules C33A puisque les inhibiteurs n'ont pas d'effet sur le LCR. Cela n'a pas encore été entièrement testé, mais nous pensons que contrairement aux cellules C33A, cette voie est activée dans les cellules HeLa du fait que la forskolin active peu le LCR (données non montrées ici). L'utilisation d'inhibiteurs de la voie de l'AMPc devrait se révéler plus efficace dans cette lignée cellulaire.

En utilisant cette même approche, nous avons identifié un nouveau site Cre potentiel qui serait altéré par la variation C7537A qui semble augmentée l'effet des stimulateurs de l'AMPc sur l'activité du LCR lorsque combinée avec la délétion de 79 pb (Fig. 5.5). Bien sûr, d'autres essais autant de liaisons à l'ADN que fonctionnels sont nécessaires pour démontrer que les variations nucléotidiques peuvent altérer l'activité du LCR du VPH33 via le voie de l'AMPc. Toutefois, nos données préliminaires indiquent que ce seraient le cas. La modulation de la transcription virale par l'AMPc a également été observée chez VPH16 et 18 pour qui la surexpression de CBP/p300 augmente l'activité transcriptionnelle (543, 760). CBP/p300 est un important coactivateur de la transcription cellulaire et virale et est une composante de l'amplificateur du promoteur de l'IFN- $\gamma$  et du VPH18. CBP/p300 est notamment recruté sur l'ADN par sa liaison avec CREB qui a son tour interagit avec la région promotrice au niveau des sites CRE (revu dans (761)). La possible perte de cet élément dans les variantes du LCR a notamment un effet majeur dans les cellules HeLa, indiquant que l'activation du LCR du VPH33 est dépendant des fonctions de CBP/p300 dans ces cellules et moins dans les C33A (voir section 5.1.3). Ainsi, la voie dépendante de l'AMPc serait beaucoup plus active dans les cellules HeLa que dans les C33A. Il serait intéressant d'étudier l'implication de la voie dépendante de l'AMPc dans la modulation du LCR dans des kératinocytes primaires.

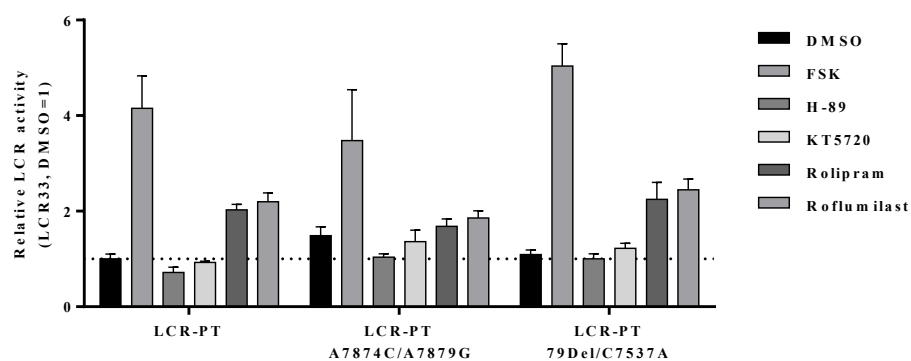


**Figure 5.3 Effets de l'introduction des variations A7874C et A7879G sur l'activité transcriptionnelle du LCR prototypique.** Les variations A7874C et A7879G ont été introduites dans le LCR-PT individuellement ou en combinaison. L'activité transcriptionnelle de ces LCRs a été mesurée dans les cellules C33A et HeLa dans un essai de transcription transitoire tel que décrit dans (705). Uniquement les aires sous la courbe (AUC) rapportés en pourcentage sont présentées. Une ANOVA unidirectionnelle a été appliquée afin de tester la différence statistique entre le prototype et les LCR mutants. \*\*\* =  $p < 0.001$ .



**Figure 5.4 Effet de la forskolin sur l'activité transcriptionnelle du LCR prototypique et mutants.** Les variations A7874C et A7879G ont été introduites dans le LCR-PT individuellement ou en combinaison. L'activité transcriptionnelle de ces LCRs (100 ng) a été mesurée 24h post-transfection dans les cellules C33A dans un essai de transcription transitoire décrit dans (705) en présence ou non de 50 $\mu$ M de forskolin. Les valeurs sont rapportées relatif à l'activité du LCR-PT qui a été assignée une valeur de 1. Les niveaux d'activation sont indiqués en haut de chaque crochet. Notez que cette figure a été produite par mon collègue Dr David Gagnon.





**Figure 5.5 Effet d’inhibiteurs et d’activateurs de la voie dépendante de l’AMPC sur l’activité transcriptionnelle du LCR.** Les variations A7874C et A7879G et la délétion de 79bp et C7537A ont été introduites dans le LCR-PT en combinaison, respectivement. L’activité transcriptionnelle de ces LCRs (100 ng) a été mesurée 24h post-transfection dans les cellules C33A dans un essai de transcription transitoire décrit dans (705) en présence ou non composés chimiques inhibiteurs (H89 et KT5720) ou activateurs (forskolin ou FSK, rolipram et roflumilast) de l’AMPC. Les valeurs sont rapportées relatif à l’activité du LCR-PT qui a été assignée une valeur de 1 en présence de DMSO (ligne pointillée). Notez que cette figure a été produite par mon collègue Dr David Gagnon.

Quant à la variation T7791C, nos études montrent qu’elle abolit un site répresseur pour C/EBP $\beta$  (chapitre 3, Fig. 3.3). Quelques études ont rapporté des mutations au niveau des sites C/EBP dans le LCR des VPHs 16, 31, 52 et 58, mais aucune n’a déterminé l’implication réelle de l’altération de ces sites sur la modulation de la transcription virale (719, 747, 748, 755). Nous sommes les premiers à démontrer que l’altération de la séquence de C/EBP $\beta$  par une variation naturelle unique affecte significativement les niveaux de transcription dans les cellules C33A, HeLa et U2OS (chapitre 3, Fig. 3.2). Une analogie peut être faite avec le FT YY1, dont les sites sont fréquemment altérés ou même supprimés dans le LCR du VPH16 et 18 (603, 604). Ces sites agissent comme répresseurs de l’activité transcriptionnelle alors que leurs mutations sont associées à l’augmentation des fonctions promotrices virales (603, 604). Ainsi, comme la mutation des sites YY1, l’abolition des sites C/EBP $\beta$  représente une stratégie pour lever la répression du LCR. La

répression par le site BS1 du VPH33 serait une fonction basale puisque contrairement aux autres facteurs identifiés dans la variante LCR12 et les variantes de la sous-lignée A2, C/EBP $\beta$  est un répresseur du LCR autant dans les cellules C33A, HeLa et U2OS, et ce, avec la même amplitude. Finalement, nous montrons qu'il est également important d'investiguer l'effet de variation unique qui altère la liaison pour des régulateurs importants de la transcription virale, dans ce cas-ci C/EBP $\beta$ . Les variations dans les sites C/EBP $\beta$  devront être davantage répertoriées dans les prochaines études épidémiologiques afin de déterminer si l'affectation de ces sites serait impliquée dans un processus de carcinogénèse plus agressive chez les individus infectés par le VPH.

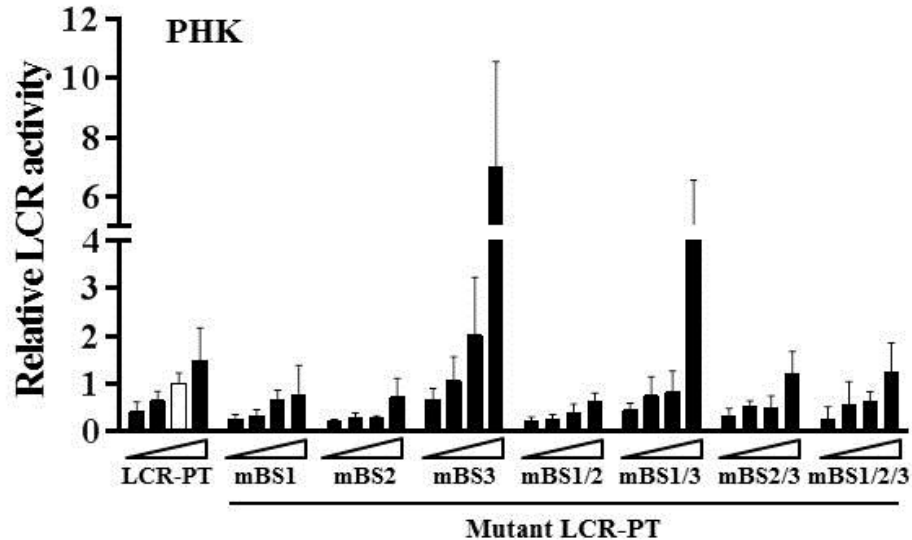
## **5.2 Modulation de l'expression des gènes viraux par C/EBP $\beta$**

Le facteur C/EBP $\beta$  est un régulateur de l'expression des gènes du VPH11, 16 et 18 (584, 609, 635-638). Nos études sont les premières à montrer que C/EBP $\beta$  est également impliqué dans la modulation de l'activité transcriptionnelle du LCR du VPH33 via un site répresseur (33BS1) et un site activateur (33BS2).

### **5.2.1 Bi-fonctionnalité de C/EBP $\beta$ au sein du même promoteur**

Les sites de C/EBP $\beta$  BS1 et BS2 dans le LCR du VPH33 sont impliqués dans la répression et l'activation de l'expression des gènes viraux (chapitre 3, Fig. 3.5). Alors que l'activité de répression de BS1 est conservée dans les cellules C33A, HeLa et U2OS, les fonctions activatrices de BS2 sont clairement dominantes dans les cellules HeLa où le LCR est davantage actif que dans les C33A. Ce site pourrait possiblement avoir un rôle d'activateur moins apparent dans les cellules U2OS. Fait intéressant, alors que BS1 et BS2 sont des sites répresseurs et activateurs dans les PHKs, BS3 semble être impliqué aussi dans l'activation du LCR du VPH33 (voir Fig. 5.6). Dans les kératinocytes primaires, C/EBP $\beta$  jouerait un rôle plutôt d'activateur alors que les activités de BS2 et BS3 semblent dominer sur celle de BS1 tel qu'observé chez un triple mutant. Un effet de BS3 n'a pas été observé chez aucune des lignées cellulaires immortalisées (C33A, HeLa et U2OS). Nos études montrent que BS3 est un site de faible affinité comparé à BS1 et BS2. Il est donc

possible que dans les PHKs, la syntaxe optimale, ou l'organisation des sites de liaison autant dans l'espace et au niveau de l'orientation, autour de BS3 permette l'activation du LCR en compensant pour la faible affinité de ce site. Effectivement, un article récent montre l'importance de la syntaxe de la région amplificatrice sur l'expression génique chez *Ciona intestinalis* lorsque des sites à faible affinité pour des FTs sont impliqués (762)



**Figure 5.6 Effets de la mutation des sites C/EBP $\beta$  l'activité transcriptionnelle du LCR prototypique dans des kératinocytes primaires humains.** Les sites C/EBP $\beta$  BS1, BS2 et BS3 ont été inactivés par l'introduction de quatre mutations soit individuellement ou en combinaison. L'activité transcriptionnelle de ces LCRs a été mesurée dans les kératinocytes primaires humains (PHKs) dans un essai de transcription transitoire tel que décrit dans (705).

Les fonctions de C/EBP $\beta$  sont également déterminées en partie par l'expression différentielle des isoformes. LIP endogène est mieux exprimé que LAP (ratio LAP/LIP  $\approx$  0.5) dans les cellules C33A et U2OS alors que dans les cellules HeLa, LAP est plutôt la forme dominante (ratio LAP/LIP  $\approx$  2). Ainsi, la différence entre le ratio LAP et LIP dans les cellules C33A et U2OS comparativement aux cellules HeLa pourrait expliquer pourquoi le site BS2 est davantage un activateur dans ces dernières cellules. La dualité dans les

fonctions de C/EBP $\beta$  a également été rapportée chez le VPH18 et 11. C/EBP $\beta$  peut activer ou réprimer le LCR selon le contexte cellulaire (602, 609). Ainsi, l'activation du LCR serait dépendante de la formation du complexe YY1-C/EBP $\beta$  dans les cellules HeLa alors qu'il est absent dans les cellules HepG2 où le LCR n'est pas actif. Ce complexe lie spécifiquement la séquence 18-BS1, mais non le site consensus à grande affinité de C/EBP $\beta$  retrouvé dans la région promotrice de IL-6. Ainsi, ce site dans le LCR du VPH18 est un élément composite. La surexpression de C/EBP $\beta$  cause la répression du LCR du VPH11 dans les HFK, mais active la transcription virale dans les cellules C33A même lorsque les sites C/EBP (deux sites C/EBP $\alpha$  et un site C/EBP $\beta$ ) sont mutés (562). La répression par C/EBP $\beta$  impliquerait un site YY1 situé à plus de 300 pb en amont. C/EBP $\beta$  pourrait donc avoir un rôle indépendant de sa liaison à l'ADN possiblement en séquestrant certains partenaires d'interaction tels que YY1. Dans le cas du VPH33, la surexpression de LAP active plutôt le LCR dans les cellules C33A, HeLa et U2OS (chapitre 3, Fig. 3.7). La surexpression de LIP réprime le LCR particulièrement dans les cellules HeLa alors qu'il est déjà peu actif dans les C33A et U2OS. Fait intéressant, nos expériences de surexpression de LAP et LIP indiquent que les fonctions intrinsèques de ces sites peuvent être changées selon les niveaux de LAP et LIP produits (chapitre 3, Fig. 3.7). Ainsi, BS1 et BS2 ont le potentiel autant d'activer que de réprimer le LCR, ce qui expliquerait les résultats contradictoires que certaines études ont montré lorsque C/EBP $\beta$  est surexprimé.

Par contre, l'expression différentielle de LAP et LIP n'explique pas comment les fonctions basales de BS1 et BS2 sont déterminées au sein d'une même région promotrice i.e pourquoi BS1 agit principalement comme un répresseur et BS2 comme un activateur du LCR. Une possibilité est que LIP lie préférentiellement le site BS1. Par contre, nos résultats de FP montrent que ces deux sites ont des affinités similaires pour un homodimère de C/EBP $\beta$  (chapitre 3, Fig. 3.4 et Tableau 3.1). Toutefois, nous utilisons une fusion GST qui ne code que le DBD et le bZIP dans nos essais de liaison à l'ADN in vitro. Il est connu que le TAD de C/EBP $\beta$  interagit avec les complexes CBP/p300 et SWI/SNF ((763, 764) et revu dans (618)). La phosphorylation du TAD permet non seulement de réguler la transactivation par C/EBP $\beta$ , mais également sa liaison avec d'autres régulateurs transcriptionnels (618). Ainsi, il est possible que LAP soit préférentiellement recruté au site BS2 par l'intermédiaire d'une interaction entre son TAD et un autre FT dont son site serait avoisinant. Une analyse des séquences adjacentes à BS1 et BS2 montrent que le contexte de ces sites est différent.

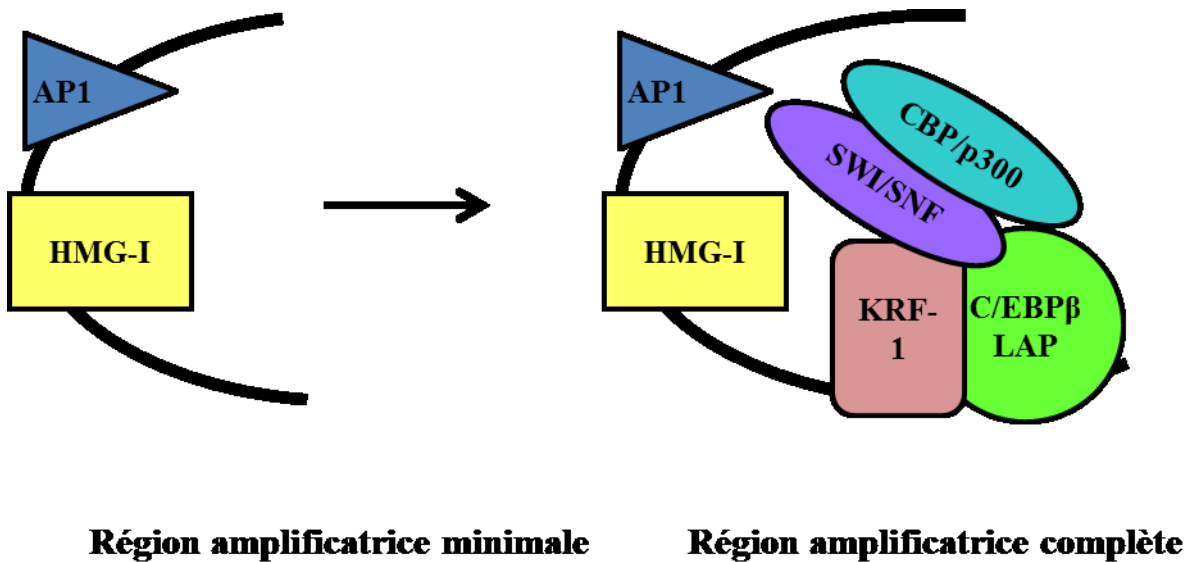
Effectivement, un site NF1 est immédiatement adjacent à BS1 et l'occupation de ce dernier par C/EBP $\beta$  pourrait donc empêcher la liaison de NF1 sur l'ADN (voir alignement au chapitre 4, Fig. 4.7). Quant au site BS2, aucun site pour un TF particulier ne semble chevaucher ce site, alors que les éléments les plus proches sont situés à une vingtaine de pb en amont (AP1) et en aval (USF). Ainsi, il est possible que l'occupation de BS2 par C/EBP $\beta$  facilite le recrutement de FT et coactivateurs.

### **5.2.2 La formation de l'amplificateur à la base de l'activation du LCR dans les cellules HeLa?**

Des analyses d'empreinte indiquent la présence de sites pour HMG-I(Y), Oct-1, NF1, AP1 et KRF-1 dans la région amplificatrice minimale du LCR du VPH18 qu (530, 541, 543). Fait intéressant, ces sites sont pour la plupart nécessaires pour la formation de l'amplificateur du VPH18, en particulier le site AP1 (voir Fig. 5.7) (541, 580). Nos prédictions ont identifié un site similaire à BS2 dans la région protégée identifiée pour VPH18 (chapitre 4). Il est difficile de déterminer si C/EBP $\beta$  est effectivement le facteur qui lie la région de KRF-1 ou si son site chevauche le site KRF-1. Toutefois, nos études fonctionnelles indiquent que le site C/EBP $\beta$  du VPH18 est également un activateur du LCR et est homologue au site 33BS2 autant au niveau de sa localisation dans le LCR et que dans ses fonctions activatrices de la transcription (chapitre 4, Fig. 4.6). Nous pouvons donc spéculer que le site C/EBP $\beta$  est une composante importante pour la formation de l'amplificateur sur le LCR du VPH18 et 33.

Plusieurs facteurs sont recrutés au niveau de l'amplificateur, que ce soit des protéines architecturale ou des complexes de remodelage de la chromatine tels HMG-I et SWI/SNF pour stimuler l'assemblage du PIC (541, 543, 765-768). Le recrutement de C/EBP $\beta$  serait possible non seulement à cause de la présence de son site de liaison, mais aussi par l'intermédiaire de son interaction avec SWI/SNF et/ou le complexe médiateur via son TAD. Ainsi, la forme activatrice LAP serait préférablement recrutée sur l'amplificateur du fait que la forme inhibitrice est dépourvue de N-terminal (763). La composition de l'amplificateur serait probablement différente pour chaque type, mais les éléments communs seraient organisés autour du site AP1 et contiendrait NF1, Oct1, KRF-1 et LAP.

De ce fait, AP1 est exprimé dans les cellules basales et suprabasales dans l'épithélium normal et les LSIL, mais dans toutes les strates de l'épithélium dans les lésions de haut-grade et CC (579, 581). L'expression de C/EBP $\beta$  est similaire à AP1 i.e que ce FT est exprimé dans les couches basales et suprabasales dans un épithélium normal, mais est retrouvé de façon ubiquitaire dans les tissus infectés (632). Ainsi, la formation de l'amplificateur sur le LCR serait favorisée dans les cellules exprimant E6 et E7 et corrèle donc avec l'expression spatiale de AP1 et de C/EBP $\beta$  dans l'épithélium infecté.

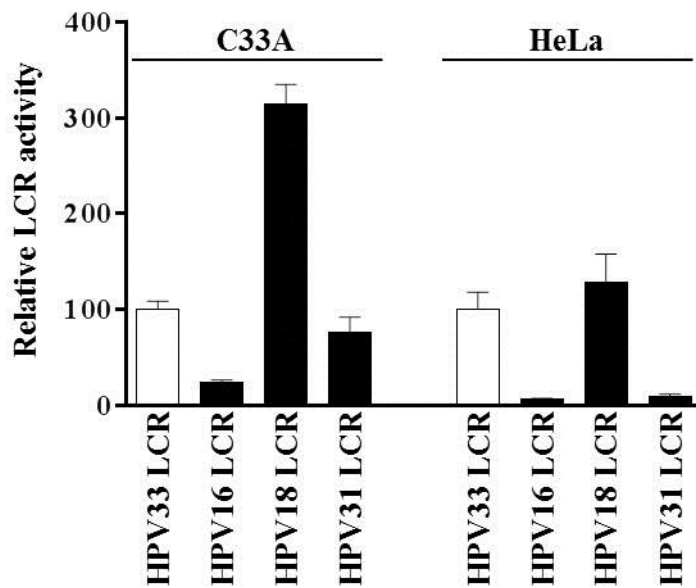


**Figure 5.7** Modèle de la formation de l'amplificateur sur le LCR du VPH18.

La région amplificatrice minimale du LCR du VPH18 (7566–7674) contient des sites pour AP1 et de HMG-I. La liaison de ces facteurs permet le recrutement séquentiel d'autre FT, dont possiblement KRF-1 et C/EBP $\beta$ , pour la formation d'un complexe d'ordre supérieur et l'activation maximale du LCR. La formation de l'amplificateur sur le LCR du VPH33 se ferait de manière similaire.

Alors que la formation d'une structure amplificatrice sur le LCR du VPH18 dans les cellules HeLa est documentée (541, 543), il n'y a aucune étude qui indique que ce complexe soit formé dans les C33A ou dans les U2OS. Toutefois, l'expression de LAP est

considérablement réduite dans les cellules C33A et U2OS (chapitre 3, Fig. 3.6), indiquant que la formation de l'amplificateur n'est pas optimale dans les cellules C33A et U2OS. De plus, les C33A sont déficientes de BRG1, une composante de SWI/SNF qui est requise pour l'activation du LCR (769). Nos prédictions indiquent que le LCR du VPH16 et 31 ne contiennent pas de site pour des homodimères de C/EBP $\beta$  dans la région minimale amplificatrice (chapitre 4). Ces deux types ne sont pas actifs dans les cellules HeLa, contrairement au VPH33 et VPH18 (voir Fig. 5.8). Ainsi, en absence d'un site C/EBP $\beta$ , l'amplificateur sur le LCR du VPH16 et VPH31 ne se formerait pas dans les cellules HeLa.



**Figure 5.8** Activité transcriptionnelle du LCR pour différents types de VPHs.

Les LCRs des VPHs 16, 18 et 31 ont été clonés et leur activité de transcription a été mesurée dans les cellules C33A et HeLa tel que décrit dans (705). Uniquement les aires sous la courbe (AUC) exprimées en pourcentage du LCR du VPH33 sont présentées. Notez que l'activité du LCR du VPH33 est 3-4 fois plus élevée dans les cellules HeLa que dans les cellules C33A.

### 5.3. Conservation des sites C/EBP $\beta$ des VPHs muqueux

La prédiction de sites C/EBP $\beta$  de 175 types de VPH à partir d'une matrice d'énergie a permis d'identifier plusieurs sites dans le LCR, notamment ceux qui ont été caractérisés pour le VPH11 et VPH33 (chapitre 4) (544, 562). L'identification de ces sites est basée sur des analyses *in silico* et il est donc important qu'un plus grand nombre de sites C/EBP $\beta$  soient validés autant *in vitro* que *in vivo*. Néanmoins, la correspondance entre les affinités prédites et mesurées pour certains sites que nous avons évalués est élevée, ce qui indique que notre prédiction est assez juste (chapitre 4, Fig. 4.2). De plus, près de 70% des sites que nous avons prédits par la matrice énergétique sont également identifiés par la matrice de fréquence, ce qui augmente la probabilité que ces sites soient des sites authentiques pour C/EBP $\beta$  (chapitre 4, Fig. 4.3).

Notre analyse montre que de 70-85% des types muqueux possèdent au moins un site C/EBP $\beta$  dans le LCR dont une grande proportion a été identifié autant par la matrice énergétique que de fréquence (chapitre 4, Fig. 4.4). La grande majorité des types cutanés ne contiennent pas de site C/EBP $\beta$ , suggérant ainsi que C/EBP $\beta$  joue un rôle dans la détermination du tropisme viral. Effectivement, le LCR des VPHs muqueux est habituellement plus actif dans les cellules épithéliales d'origine muqueuse que dans d'autres types cellulaires (544, 545, 608). Ainsi, le LCR du VPH18 et 31 ne sont pas actifs dans les cellules rénales HepG2 qui n'expriment pas C/EBP $\beta$ , alors que la surexpression de ce FT dans ces cellules réactive le LCR (602, 609). Les expériences réalisées dans notre laboratoire ainsi que les études par d'autres indiquent que les cellules U2OS dérivées d'un ostéosarcome supportent l'expression génique et la réplication virale (voir chapitre 3 et 4) (544, 770-772). L'activité basale du LCR dans ces cellules est plutôt réprimée et corrèle avec les niveaux faibles de C/EBP $\beta$  comparativement aux cellules HeLa. La surexpression de C/EBP $\beta$  dans ces cellules augmente l'activité transcriptionnelle du LCR. Il est donc possible que le degré d'activation du LCR dépend en partie des niveaux de LAP endogène dans les cellules. De ce fait, il serait intéressant de déterminer si l'activation du LCR par la surexpression de C/EBP $\beta$  résulterait en une augmentation de la réplication virale.

La plupart des études sur la liaison de C/EBP $\beta$  sur le LCR des VPHs se basent sur l'utilisation d'essais d'empreintes ou de retard sur gel (562, 609, 635, 638). Ces études ne permettent pas nécessairement de déterminer si C/EBP $\beta$  lie directement le LCR via un site



homodimérique ou si l'hétérodimérisation avec un autre facteur est requis pour son interaction sur l'ADN. Notre prédiction de sites en utilisant une matrice d'énergie complétée avec les résultats obtenus à partir de la matrice de fréquence permet d'identifier les sites pour un homodimère de C/EBP $\beta$  et possiblement les sites hétérodimériques (chapitre 4). Ainsi, nous montrons que certains sites dans le VPH33, 11, 18 et 16 sont liés par des homodimères de C/EBP $\beta$ . Les sites identifiés par la matrice de fréquence, mais non par la matrice énergétique pourraient représenter des éléments composites pour des hétérodimères de C/EBP $\beta$ . Ces sites pourraient être distingués des éléments homodimériques selon l'identité des nucléotides aux positions 5, 6 et 7 de la séquence consensus 5'-ATTGCGCAAT-3' (voir logo au chapitre 4, Fig. 4.3). De ce fait, la liaison d'homodimères requiert que les deux demi-sites soit relativement conservés, ce qui est souvent observé pour les sites homodimériques 33BS1 (CTTACACAAT) et 33BS2 (ATTGCAAAT). Pour les sites d'hétérodimères, la conservation d'un seul demi-site serait suffisante pour la liaison de C/EBP $\beta$  alors que la spécificité de l'interaction à l'ADN est également fournie par l'autre demi-site pour un autre FT. Les séquences étant très hétérogènes, il est difficile de déterminer la prévalence d'un site composite pour un hétérodimère particulier parmi les éléments identifiés par la matrice des fréquences. Cette hétérogénéité reflète donc les nombreuses combinaisons que peut former C/EBP $\beta$ . Des études sont nécessaires pour identifier ces différents complexes, mais surtout leurs effets sur la modulation de l'expression génique du VPH. Il est très probable que la diversité des sites composites permette la modulation de la transcription de manière différente dans les types de VPHs. Dans ce sens, nos résultats montrent que certains VPHs contiennent des sites C/EBP $\beta$  au niveau de la région centrale du LCR tandis que d'autres sont dépourvus de ces sites (chapitre 4, Fig. 4.7), ce qui indique que ces derniers sont régulés différemment. Le LCR du VPH16 qui est dépourvu de site C/EBP $\beta$  au niveau de la région amplificatrice possède un site GRE qui ne semble pas conservé chez les VPHs ayant un site C/EBP $\beta$  (531, 570, 773).

Les études indiquent que C/EBP $\beta$  peut se dimériser avec AP1, ATF, NF- $\kappa$ B et CREB et lier plusieurs séquences sur les promoteurs cellulaires (774-776). Fait intéressant, l'activation de la voie dépendante de l'AMPc par des hormones ou facteurs de croissance régule positivement l'activité transcriptionnelle de CREB, qui a son tour régule l'expression de C/EBP $\beta$  (619, 777-779). De plus, les kinases PKA et PKC impliquées dans

cette voie signalétique modulent les fonctions de C/EBP $\beta$  en phosphorylant certains résidus dans le TAD ou dans le DBD (617, 780, 781). Ainsi, l'activation de la voie dépendante de l'AMPc peut potentialiser l'effet de C/EBP $\beta$  sur l'expression génique des VPHs en réponse à plusieurs stimuli. La formation d'hétérodimères permettrait la liaison de C/EBP $\beta$  sur des sites non-canoniques et diversifierait encore plus ses activités régulatrices.

D'autres études montrent que la liaison d'hétérodimères c-Jun-C/EBP $\beta$  sur les promoteurs cellulaires est associée à la répression de l'expression génique tandis que les sites homodimériques C/EBP $\beta$  sont associés à l'activation (764, 775, 782). La liaison de ces complexes sur l'ADN varie en fonction de la différenciation cellulaire, indiquant que les sites hétérodimériques et homodimériques sont impliqués dans la répression de l'expression des gènes de prolifération et dans l'activation des gènes de différenciation, respectivement. Ainsi, il est envisageable que l'expression différentielle de C/EBP $\beta$  et les différentes combinaisons avec d'autres FT modulent l'activité transcriptionnelle du LCR durant le cycle viral permettant la transition de l'expression des gènes précoces et tardifs.

## Chapitre sixième : Conclusions et perspectives

Le but premier du projet de recherche détaillé dans cette thèse était d'étudier la modulation de l'expression génique virale par des variations naturelles contenues dans le LCR du VPH33. Dans le chapitre 2, nous avons montré que c'était la contribution de quelques variations qui augmente l'activité transcriptionnelle des variantes de la sous-lignée A2, expliquant leur association avec les lésions de haut-grade par rapport au prototype. Nous avons également validé l'hypothèse que les variations dans un élément régulateur changent l'affinité de liaison pour un FT donné, qui dans le cas du VPH33, résulte en l'augmentation de l'activité transcriptionnelle du LCR. Ainsi, nous avons démontré que la variation T7791C dans la variante LCR10, dont sa caractérisation a été présentée dans le chapitre 3 de cette thèse, abolissait un site répresseur pour le FT C/EBP $\beta$  et était donc responsable de l'activité transcriptionnelle élevée de cette variante. Outre son rôle dans la répression du LCR, nous avons également montré que C/EBP $\beta$  est un activateur important de l'expression des gènes du VPH33. L'expression différentielle des isoformes LAP et LIP détermine les fonctions de C/EBP $\beta$  sur la modulation du LCR notamment. L'analyse de la prévalence des sites C/EBP $\beta$  dans le LCR de 175 types de VPHs, présentée dans le chapitre 4, a révélé l'enrichissement des sites C/EBP $\beta$  dans le LCR des VPHs oncogéniques, dont plusieurs contiennent des éléments similaires aux sites répresseurs et activateurs du VPH33. De plus, nos résultats suggèrent que la diversité des sites liant des homo- et des hétérodimères de C/EBP $\beta$  augmente l'étendue de l'action de ce FT sur l'activité du LCR.

Dans l'ensemble, nos résultats corroborent le concept que des variations de la séquence nucléotidique du LCR altèrent la liaison de facteurs de transcription sur l'ADN, ce qui induit des changements dans la modulation de l'expression génique virale. Nos résultats démontrent aussi que certaines variantes non-prototypiques associées à des lésions de haut-grade ont une activité transcriptionnelle accrue. De plus, nos études soulignent que C/EBP $\beta$  est un régulateur important de l'expression des gènes viraux. La perte d'un site répresseur C/EBP $\beta$  induit l'activation du LCR du VPH33 et pourrait donc être associée à une oncogénicité accrue du virus. Bien que nous ayons atteints nos objectifs généraux de caractérisation des variantes, plusieurs autres notions sur la régulation de l'activité transcriptionnelle restent à explorer.

## **6.1 Transition précoce/tardive de l'expression génique virale**

Des études montrent que certains éléments dans le LCR modulent l'activité du promoteur tardif. Alors que nos études ont porté sur l'analyse de la transcription des gènes viraux précoces, la régulation de l'expression tardive n'a pas été évaluée. Dans ce sens, il serait intéressant de déterminer si les variations dans le LCR peuvent également altérer l'expression des gènes tardifs dont le promoteur est retrouvé dans le cadre de lecture de E7. Notre essai de transcription pourrait être utilisé à cette fin pour mesurer l'activité transcriptionnelle du promoteur tardif dans un contexte où le LCR est présent en *cis*. Un candidat potentiel est certainement le FT C/EBP $\beta$  qui semblerait également être impliqué dans la modulation de l'expression des gènes tardifs, du moins pour VPH31 (176, 628). Il serait donc intéressant d'étudier le rôle de C/EBP $\beta$  dans la transition précoce/tardive, alors que peu d'études se sont penchées sur la question. Tout d'abord, il serait important de déterminer la prévalence des sites C/EBP $\beta$  dans la région promotrice tardive en particulier pour VPH33. Par la suite, des études de mutagenèses détaillées couplées à des essais fonctionnels et de liaison à l'ADN seront nécessaires pour déterminer l'occupation des régions promotrices par C/EBP $\beta$  selon l'état de différenciation cellulaire. La variation T7791C pourrait être introduite afin de voir son effet dans les phases tardives du cycle viral.

## **6.2 Effet de la co-ségrégation des variations ailleurs dans le génome viral**

Comme nous le montrons dans le chapitre 2 de cette thèse, la combinaison de variations altère l'expression des gènes viraux. D'autres variations ailleurs dans le génome sont vraisemblablement reliées aux polymorphismes retrouvés dans le LCR. La co-ségrégation de ces variations pourrait donc avoir un impact plus grand sur le cycle viral en général. Bien que l'effet de variations dans E6 et E7 ait été examiné, peu d'études ont analysé l'impact de variations dans E2 sur la répression du LCR ainsi que dans la réplication virale. Des polymorphismes dans E1 pourraient également influencer les

niveaux de réplication de l'épisome viral. Les effets de ces variations devront par ailleurs être étudiés dans un contexte de génome complet. Ainsi, les séquences nucléotidiques des variantes directement isolées de spécimens cliniques devront être clonées pour ensuite être introduites dans des kératinocytes primaires où l'expression des gènes et la réplication virale seront évaluées. Des études similaires ont testé soit des génomes prototypiques, dont certains éléments ou régions ont été artificiellement mutés ou supprimés (654), ou des génomes chimères où le LCR a été introduits dans le contexte d'un génome prototypique (615, 714), mais aucune à partir de véritable génome de variantes.

### **6.3 Détermination des voies de signalisation cellulaires régulant l'expression génique virale**

L'utilisation d'inhibiteurs ou même de composés induisant la différenciation cellulaire pourrait identifier les voies signalétiques impliquées dans la régulation différentielle des différents types et variantes de VPH, ce qui améliorerait notre compréhension du cycle viral. Ainsi, les voies signalétiques impliquées pourront être ciblées pour inhiber la transcription des gènes viraux et de ce fait, la réplication virale aussi. Dans une telle optique, il serait intéressant d'analyser plus en profondeur l'implication des voies des Map kinases et de PI3K/Akt sur l'expression génique virale. Il est connu que les oncoprotéines du VPH modulent positivement l'activité de ERK1/2 et de PI3K et que les voies impliquant ces kinases sont souvent activées dans les CIN, CC et dans les cancers oropharyngés (783-787). Ces voies induisent notamment l'expression de certains FT tels que NF- $\kappa$ B et AP1 ou régulent directement l'activité d'autres FT tels que C/EBP $\beta$  qui, à leurs tours, sont impliqués dans la modulation de la transcription des gènes du VPH (voir section 1.4.5).

Par ailleurs, il est bien connu que les oncoprotéines E6 et E7 ciblent plusieurs protéines cellulaires, mais leurs rôles dans la régulation des gènes viraux sont moins connus. Ainsi, il serait intéressant de déterminer l'effet de la surexpression de E6 et E7 sur l'activité du LCR prototypique et des variantes. Notre essai de gène rapporteur est facilement adaptable pour tester divers inhibiteurs des voies Map kinases et PI3K/Akt et la surexpression de E6 et E7 dans un essai de criblage. Étant à haut débit, une multitude de

LCRs, provenant de différents types et variantes de VPH, pourront être testés simultanément.

## **6.4 Étude de l'amplificateur du VPH33**

Les études sur l'activation du LCR du VPH18, couplées à nos propres analyses, indiquent que la liaison de C/EBP $\beta$  dans la région centrale du LCR serait impliquée dans la formation de l'amplificateur. Pour confirmer ceci, des études seront nécessaires pour, dans un premier temps, identifier la région amplificatrice minimale chez le VPH33 et, dans un second temps, identifier les complexes liant cette région par des essais de retard sur gel. Encore une fois, l'utilisation de composés activateurs ou inhibiteurs de l'expression génique pourrait nous permettre de déterminer quelles sont les conditions cellulaires qui favorisent la formation de l'amplificateur.

Ensemble, les études qui sont proposées ici permettront de mieux comprendre les mécanismes qui modulent l'expression des gènes viraux pendant le cycle viral et enfin de déterminer les éléments-clés pour ce processus. Le ciblage de ces régulateurs pourrait représenter une bonne stratégie thérapeutique à être exploitée pour la confection d'un antiviral à utiliser dans les stades précoces de l'infection au VPH.

## Bibliographie

1. **Bosch FX, Lorincz A, Munoz N, Meijer CJ, Shah KV.** 2002. The causal relation between human papillomavirus and cervical cancer. *Journal of clinical pathology* **55**:244-265.
2. **Burk RD, Harari A, Chen Z.** 2013. Human papillomavirus genome variants. *Virology* **445**:232-243.
3. **de Villiers EM.** 2013. Cross-roads in the classification of papillomaviruses. *Virology* **445**:2-10.
4. **de Villiers EM, Fauquet C, Broker TR, Bernard HU, zur Hausen H.** 2004. Classification of papillomaviruses. *Virology* **324**:17-27.
5. **Chen Z, Schiffman M, Herrero R, Desalle R, Anastos K, Segondy M, Sahasrabudde VV, Gravitt PE, Hsing AW, Burk RD.** 2011. Evolution and taxonomic classification of human papillomavirus 16 (HPV16)-related variant genomes: HPV31, HPV33, HPV35, HPV52, HPV58 and HPV67. *PloS one* **6**:e20183.
6. **Bernard HU, Burk RD, Chen Z, van Doorslaer K, zur Hausen H, de Villiers EM.** 2010. Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology* **401**:70-79.
7. **Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, Snijders PJ, Peto J, Meijer CJ, Munoz N.** 1999. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *The Journal of pathology* **189**:12-19.
8. **Lacey CJ, Lowndes CM, Shah KV.** 2006. Chapter 4: Burden and management of non-cancerous HPV-related conditions: HPV-6/11 disease. *Vaccine* **24 Suppl 3**:S3/35-41.
9. **Munoz N.** 2000. Human papillomavirus and cancer: the epidemiological evidence. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology* **19**:1-5.
10. **Bruni L, Diaz M, Castellsague X, Ferrer E, Bosch FX, de Sanjose S.** 2010. Cervical human papillomavirus prevalence in 5 continents: meta-analysis of 1 million women with normal cytological findings. *The Journal of infectious diseases* **202**:1789-1799.
11. **Dunne EF, Unger ER, Sternberg M, McQuillan G, Swan DC, Patel SS, Markowitz LE.** 2007. Prevalence of HPV infection among females in the United States. *JAMA : the journal of the American Medical Association* **297**:813-819.
12. **Franceschi S, Herrero R, Clifford GM, Snijders PJ, Arslan A, Anh PT, Bosch FX, Ferreccio C, Hieu NT, Lazcano-Ponce E, Matos E, Molano M, Qiao YL, Rajkumar R, Ronco G, de Sanjose S, Shin HR, Sukvirach S, Thomas JO, Meijer CJ, Munoz N.** 2006. Variations in the age-specific curves of human papillomavirus prevalence in women worldwide. *International journal of cancer. Journal international du cancer* **119**:2677-2684.
13. **Smith JS, Melendy A, Rana RK, Pimenta JM.** 2008. Age-specific prevalence of infection with human papillomavirus in females: a global review. *The Journal of adolescent health : official publication of the Society for Adolescent Medicine* **43**:S5-25, S25.e21-41.

14. **Palefsky JM, Minkoff H, Kalish LA, Levine A, Sacks HS, Garcia P, Young M, Melnick S, Miotti P, Burk R.** 1999. Cervicovaginal human papillomavirus infection in human immunodeficiency virus-1 (HIV)-positive and high-risk HIV-negative women. *Journal of the National Cancer Institute* **91**:226-236.
15. **Massad LS, Riestler KA, Anastos KM, Fruchter RG, Palefsky JM, Burk RD, Burns D, Greenblatt RM, Muderspach LI, Miotti P.** 1999. Prevalence and predictors of squamous cell abnormalities in Papanicolaou smears from women infected with HIV-1. Women's Interagency HIV Study Group. *Journal of acquired immune deficiency syndromes (1999)* **21**:33-41.
16. **Valdespino VM, Valdespino VE.** 2006. Cervical cancer screening: state of the art. *Current opinion in obstetrics & gynecology* **18**:35-40.
17. **Waxman AG, Zsemlye MM.** 2008. Preventing cervical cancer: the Pap test and the HPV vaccine. *The Medical clinics of North America* **92**:1059-1082, x.
18. **Canada ScdceIndcd.** 2008. Statistiques canadiennes sur le cancer.
19. **Statistics CCSsSCoC.** 2012. Canadian Cancer Statistics 2012. Canadian Cancer Society, Toronto, ON.
20. **Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A.** 2015. Global cancer statistics, 2012. *CA: a cancer journal for clinicians*.
21. **Hariri S, Bennett NM, Niccolai LM, Schafer S, Park IU, Bloch KC, Unger ER, Whitney E, Julian P, Scahill MW, Abdullah N, Levine D, Johnson ML, Steinau M, Markowitz LE.** 2015. Reduction in HPV 16/18-associated high grade cervical lesions following HPV vaccine introduction in the United States - 2008-2012. *Vaccine*.
22. **International Agency for Research on Cancer WHO.** 2007. *Cancer Incidence in Five Continents, Vol. IX*.
23. **De Vuyst H, Clifford GM, Nascimento MC, Madeleine MM, Franceschi S.** 2009. Prevalence and type distribution of human papillomavirus in carcinoma and intraepithelial neoplasia of the vulva, vagina and anus: a meta-analysis. *International journal of cancer. Journal international du cancer* **124**:1626-1636.
24. **de Sanjose S, Alemany L, Ordi J, Tous S, Alejo M, Bigby SM, Joura EA, Maldonado P, Laco J, Bravo IG, Vidal A, Guimera N, Cross P, Wain GV, Petry KU, Mariani L, Bergeron C, Mandys V, Sica AR, Felix A, Usubutun A, Seoud M, Hernandez-Suarez G, Nowakowski AM, Wilson G, Dalstein V, Hampl M, Kasamatsu ES, Lombardi LE, Tinoco L, Alvarado-Cabrero I, Perrotta M, Bhatla N, Agorastos T, Lynch CF, Goodman MT, Shin HR, Viarheichyk H, Jach R, Cruz MO, Velasco J, Molina C, Bornstein J, Ferrera A, Domingo EJ, Chou CY, Banjo AF, Castellsague X, Pawlita M, Lloveras B, Quint WG, Munoz N, Bosch FX.** 2013. Worldwide human papillomavirus genotype attribution in over 2000 cases of intraepithelial and invasive lesions of the vulva. *European journal of cancer (Oxford, England : 1990)* **49**:3450-3461.
25. **Daling JR, Madeleine MM, Schwartz SM, Shera KA, Carter JJ, McKnight B, Porter PL, Galloway DA, McDougall JK, Tamimi H.** 2002. A population-based study of squamous cell vaginal cancer: HPV and cofactors. *Gynecologic oncology* **84**:263-270.
26. **Wakeham K, Kavanagh K.** 2014. The burden of HPV-associated anogenital cancers. *Current oncology reports* **16**:402.
27. **Goodman MT, Shvetsov YB, McDuffie K, Wilkens LR, Zhu X, Thompson PJ, Ning L, Killeen J, Kamemoto L, Hernandez BY.** 2010. Sequential acquisition of



- human papillomavirus (HPV) infection of the anus and cervix: the Hawaii HPV Cohort Study. *The Journal of infectious diseases* **201**:1331-1339.
28. **Hernandez BY, Wilkens LR, Zhu X, Thompson P, McDuffie K, Shvetsov YB, Kamemoto LE, Killeen J, Ning L, Goodman MT.** 2008. Transmission of human papillomavirus in heterosexual couples. *Emerging infectious diseases* **14**:888-894.
  29. **Crawford R, Grignon AL, Kitson S, Winder DM, Ball SL, Vaughan K, Stanley MA, Sterling JC, Goon PK.** 2011. High prevalence of HPV in non-cervical sites of women with abnormal cervical cytology. *BMC cancer* **11**:473.
  30. **Daling JR, Weiss NS, Hislop TG, Maden C, Coates RJ, Sherman KJ, Ashley RL, Beagrie M, Ryan JA, Corey L.** 1987. Sexual practices, sexually transmitted diseases, and the incidence of anal cancer. *The New England journal of medicine* **317**:973-977.
  31. **Palefsky JM, Holly EA, Efirdc JT, Da Costa M, Jay N, Berry JM, Darragh TM.** 2005. Anal intraepithelial neoplasia in the highly active antiretroviral therapy era among HIV-positive men who have sex with men. *AIDS (London, England)* **19**:1407-1414.
  32. **de Sanjose S, Bruni L, Alemany L.** 2014. HPV in genital cancers (at the exception of cervical cancer) and anal cancers. *Presse medicale (Paris, France : 1983)* **43**:e423-e428.
  33. **Dayyani F, Etzel CJ, Liu M, Ho CH, Lippman SM, Tsao AS.** 2010. Meta-analysis of the impact of human papillomavirus (HPV) on cancer risk and overall survival in head and neck squamous cell carcinomas (HNSCC). *Head & neck oncology* **2**:15.
  34. **Mork J, Lie AK, Glatre E, Hallmans G, Jellum E, Koskela P, Moller B, Pukkala E, Schiller JT, Youngman L, Lehtinen M, Dillner J.** 2001. Human papillomavirus infection as a risk factor for squamous-cell carcinoma of the head and neck. *The New England journal of medicine* **344**:1125-1131.
  35. **Kreimer AR, Clifford GM, Boyle P, Franceschi S.** 2005. Human papillomavirus types in head and neck squamous cell carcinomas worldwide: a systematic review. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **14**:467-475.
  36. **Thibaudeau E, Fortin B, Coutlee F, Nguyen-Tan P, Weng X, Audet ML, Abboud O, Guertin L, Christopoulos A, Tabet J, Soulieres D.** 2013. HPV Prevalence and Prognostic Value in a Prospective Cohort of 255 Patients with Locally Advanced HNSCC: A Single-Centre Experience. *International journal of otolaryngology* **2013**:437815.
  37. **Syrjanen S.** 2010. The role of human papillomavirus infection in head and neck cancers. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* **21 Suppl 7**:vii243-245.
  38. **Chang F, Syrjanen S, Shen Q, Cintorino M, Santopietro R, Tosi P, Syrjanen K.** 2000. Human papillomavirus involvement in esophageal carcinogenesis in the high-incidence area of China. A study of 700 cases by screening and type-specific in situ hybridization. *Scandinavian journal of gastroenterology* **35**:123-130.
  39. **Taylor DC, Falconer MA, Bruton CJ, Corsellis JA.** 1971. Focal dysplasia of the cerebral cortex in epilepsy. *Journal of neurology, neurosurgery, and psychiatry* **34**:369-387.

40. **Chen J, Tsai V, Parker WE, Aronica E, Baybis M, Crino PB.** 2012. Detection of human papillomavirus in human focal cortical dysplasia type IIB. *Annals of neurology* **72**:881-892.
41. **Liu S, Lu L, Cheng X, Xu G, Yang H.** 2014. Viral infection and focal cortical dysplasia. *Annals of neurology* **75**:614-616.
42. **Coras R, Korn K, Bien CG, Kalbhenn T, Rossler K, Kobow K, Giedl J, Fleckenstein B, Blumcke I.** 2015. No evidence for human papillomavirus infection in focal cortical dysplasia IIB. *Annals of neurology* **77**:312-319.
43. **Thom M, Liu J, Reeves C, Stopps V, Sisodiya SM.** 2015. A cautionary note in the interpretation of human papillomavirus E6 immunohistochemistry in focal cortical dysplasia. *Annals of neurology* **77**:352-353; discussion 353.
44. **Tao M, Kruhlak M, Xia S, Androphy E, Zheng ZM.** 2003. Signals that dictate nuclear localization of human papillomavirus type 16 oncoprotein E6 in living cells. *Journal of virology* **77**:13232-13247.
45. **Moscicki AB, Schiffman M, Burchell A, Albero G, Giuliano AR, Goodman MT, Kjaer SK, Palefsky J.** 2012. Updating the natural history of human papillomavirus and anogenital cancers. *Vaccine* **30 Suppl 5**:F24-33.
46. **Woodman CB, Collins SI, Young LS.** 2007. The natural history of cervical HPV infection: unresolved issues. *Nature reviews. Cancer* **7**:11-22.
47. **Rodriguez AC, Schiffman M, Herrero R, Hildesheim A, Bratti C, Sherman ME, Solomon D, Guillen D, Alfaro M, Morales J, Hutchinson M, Katki H, Cheung L, Wacholder S, Burk RD.** 2010. Longitudinal study of human papillomavirus persistence and cervical intraepithelial neoplasia grade 2/3: critical role of duration of infection. *Journal of the National Cancer Institute* **102**:315-324.
48. **Winer RL, Hughes JP, Feng Q, Xi LF, Cherne S, O'Reilly S, Kiviat NB, Koutsky LA.** 2011. Early natural history of incident, type-specific human papillomavirus infections in newly sexually active young women. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **20**:699-707.
49. **Gravitt PE.** 2011. The known unknowns of HPV natural history. *The Journal of clinical investigation* **121**:4593-4599.
50. **Winer RL, Kiviat NB, Hughes JP, Adam DE, Lee SK, Kuypers JM, Koutsky LA.** 2005. Development and duration of human papillomavirus lesions, after initial infection. *The Journal of infectious diseases* **191**:731-738.
51. **Holowaty P, Miller AB, Rohan T, To T.** 1999. Natural history of dysplasia of the uterine cervix. *Journal of the National Cancer Institute* **91**:252-258.
52. **Moscicki AB, Ma Y, Wibbelsman C, Darragh TM, Powers A, Farhat S, Shiboski S.** 2010. Rate of and risks for regression of cervical intraepithelial neoplasia 2 in adolescents and young women. *Obstetrics and gynecology* **116**:1373-1380.
53. **Castle PE, Schiffman M, Wheeler CM, Solomon D.** 2009. Evidence for frequent regression of cervical intraepithelial neoplasia-grade 2. *Obstetrics and gynecology* **113**:18-25.
54. **Matsumoto K, Oki A, Furuta R, Maeda H, Yasugi T, Takatsuka N, Mitsuhashi A, Fujii T, Hirai Y, Iwasaka T, Yaegashi N, Watanabe Y, Nagai Y, Kitagawa T, Yoshikawa H.** 2011. Predicting the progression of cervical precursor lesions by

- human papillomavirus genotyping: a prospective cohort study. *International journal of cancer. Journal international du cancer* **128**:2898-2910.
55. **McCredie MR, Sharples KJ, Paul C, Baranyai J, Medley G, Jones RW, Skegg DC.** 2008. Natural history of cervical neoplasia and risk of invasive cancer in women with cervical intraepithelial neoplasia 3: a retrospective cohort study. *The Lancet. Oncology* **9**:425-434.
  56. **Schlecht NF, Platt RW, Duarte-Franco E, Costa MC, Sobrinho JP, Prado JC, Ferenczy A, Rohan TE, Villa LL, Franco EL.** 2003. Human papillomavirus infection and time to progression and regression of cervical intraepithelial neoplasia. *Journal of the National Cancer Institute* **95**:1336-1343.
  57. **Kjaer SK, Frederiksen K, Munk C, Iftner T.** 2010. Long-term absolute risk of cervical intraepithelial neoplasia grade 3 or worse following human papillomavirus infection: role of persistence. *Journal of the National Cancer Institute* **102**:1478-1488.
  58. **Freitas LB, Chen Z, Muqui EF, Boldrini NA, Miranda AE, Spano LC, Burk RD.** 2014. Human papillomavirus 16 non-European variants are preferentially associated with high-grade cervical lesions. *PloS one* **9**:e100746.
  59. **Castellsague X.** 2008. Natural history and epidemiology of HPV infection and cervical cancer. *Gynecologic oncology* **110**:S4-7.
  60. **Ho GY, Burk RD, Klein S, Kadish AS, Chang CJ, Palan P, Basu J, Tachezy R, Lewis R, Romney S.** 1995. Persistent genital human papillomavirus infection as a risk factor for persistent cervical dysplasia. *Journal of the National Cancer Institute* **87**:1365-1371.
  61. **Chen HC, Schiffman M, Lin CY, Pan MH, You SL, Chuang LC, Hsieh CY, Liaw KL, Hsing AW, Chen CJ.** 2011. Persistence of type-specific human papillomavirus infection and increased long-term risk of cervical cancer. *Journal of the National Cancer Institute* **103**:1387-1396.
  62. **Koshiol J, Lindsay L, Pimenta JM, Poole C, Jenkins D, Smith JS.** 2008. Persistent human papillomavirus infection and cervical neoplasia: a systematic review and meta-analysis. *American journal of epidemiology* **168**:123-137.
  63. **Schlecht NF, Kulaga S, Robitaille J, Ferreira S, Santos M, Miyamura RA, Duarte-Franco E, Rohan TE, Ferenczy A, Villa LL, Franco EL.** 2001. Persistent human papillomavirus infection as a predictor of cervical intraepithelial neoplasia. *JAMA : the journal of the American Medical Association* **286**:3106-3114.
  64. **Syrjanen K, Shabalova I, Naud P, Kozachenko V, Derchain S, Zakharchenko S, Roteli-Martins C, Nerovjna R, Longatto-Filho A, Kljukina L, Tatti S, Branovskaja M, Hammes LS, Branca M, Grunjberga V, Erzen M, Sarian LO, Juschenko A, Costa S, Podistov J, Syrjanen S.** 2009. Persistent high-risk human papillomavirus infections and other end-point markers of progressive cervical disease among women prospectively followed up in the New Independent States of the Former Soviet Union and the Latin American Screening study cohorts. *International journal of gynecological cancer : official journal of the International Gynecological Cancer Society* **19**:934-942.
  65. **Pett M, Coleman N.** 2007. Integration of high-risk human papillomavirus: a key event in cervical carcinogenesis? *The Journal of pathology* **212**:356-367.
  66. **Corden SA, Sant-Cassia LJ, Easton AJ, Morris AG.** 1999. The integration of HPV-18 DNA in cervical carcinoma. *Molecular pathology : MP* **52**:275-282.

67. **Cullen AP, Reid R, Champion M, Lorincz AT.** 1991. Analysis of the physical state of different human papillomavirus DNAs in intraepithelial and invasive cervical neoplasm. *Journal of virology* **65**:606-612.
68. **Kulmala SM, Syrjanen SM, Gyllensten UB, Shabalova IP, Petrovichev N, Tosi P, Syrjanen KJ, Johansson BC.** 2006. Early integration of high copy HPV16 detectable in women with normal and low grade cervical cytology and histology. *Journal of clinical pathology* **59**:513-517.
69. **Arias-Pulido H, Peyton CL, Joste NE, Vargas H, Wheeler CM.** 2006. Human papillomavirus type 16 integration in cervical carcinoma in situ and in invasive cervical cancer. *Journal of clinical microbiology* **44**:1755-1762.
70. **Klaes R, Woerner SM, Ridder R, Wentzensen N, Duerst M, Schneider A, Lotz B, Melsheimer P, von Knebel Doeberitz M.** 1999. Detection of high-risk cervical intraepithelial neoplasia and cervical cancer by amplification of transcripts derived from integrated papillomavirus oncogenes. *Cancer research* **59**:6132-6136.
71. **Peitsaro P, Johansson B, Syrjanen S.** 2002. Integrated human papillomavirus type 16 is frequently found in cervical cancer precursors as demonstrated by a novel quantitative real-time PCR technique. *Journal of clinical microbiology* **40**:886-891.
72. **Khouadri S, Villa LL, Gagnon S, Koushik A, Richardson H, Ferreira S, Tellier P, Simao J, Matlashewski G, Roger M, Franco EL, Coutlee F.** 2006. Human papillomavirus type 33 polymorphisms and high-grade squamous intraepithelial lesions of the uterine cervix. *The Journal of infectious diseases* **194**:886-894.
73. **Tonon SA, Picconi MA, Bos PD, Zinovich JB, Galuppo J, Alonio LV, Teysse AR.** 2001. Physical status of the E2 human papilloma virus 16 viral gene in cervical preneoplastic and neoplastic lesions. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology* **21**:129-134.
74. **Hopman AH, Smedts F, Dignef W, Ummelen M, Sonke G, Mravunac M, Vooijs GP, Speel EJ, Ramaekers FC.** 2004. Transition of high-grade cervical intraepithelial neoplasia to micro-invasive carcinoma is characterized by integration of HPV 16/18 and numerical chromosome abnormalities. *The Journal of pathology* **202**:23-33.
75. **Vernon SD, Unger ER, Miller DL, Lee DR, Reeves WC.** 1997. Association of human papillomavirus type 16 integration in the E2 gene with poor disease-free survival from cervical cancer. *International journal of cancer. Journal international du cancer* **74**:50-56.
76. **Kalantari M, Karlsten F, Kristensen G, Holm R, Hagmar B, Johansson B.** 1998. Disruption of the E1 and E2 reading frames of HPV 16 in cervical carcinoma is associated with poor prognosis. *International journal of gynecological pathology : official journal of the International Society of Gynecological Pathologists* **17**:146-153.
77. **Hudelist G, Manavi M, Pischinger KI, Watkins-Riedel T, Singer CF, Kubista E, Czerwenka KF.** 2004. Physical state and expression of HPV DNA in benign and dysplastic cervical tissue: different levels of viral integration are correlated with lesion grade. *Gynecologic oncology* **92**:873-880.
78. **Pirami L, Giache V, Becciolini A.** 1997. Analysis of HPV16, 18, 31, and 35 DNA in pre-invasive and invasive lesions of the uterine cervix. *Journal of clinical pathology* **50**:600-604.
79. **Chaturvedi AK, Katki HA, Hildesheim A, Rodriguez AC, Quint W, Schiffman M, Van Doorn LJ, Porras C, Wacholder S, Gonzalez P, Sherman ME, Herrero**

- R. 2011. Human papillomavirus infection with multiple types: pattern of coinfection and risk of cervical disease. *The Journal of infectious diseases* **203**:910-920.
80. **Lee SA, Kang D, Seo SS, Jeong JK, Yoo KY, Jeon YT, Kim JW, Park NH, Kang SB, Lee HP, Song YS.** 2003. Multiple HPV infection in cervical cancer screened by HPV DNA chip. *Cancer letters* **198**:187-192.
81. **Nielsen A, Kjaer SK, Munk C, Iftner T.** 2008. Type-specific HPV infection and multiple HPV types: prevalence and risk factor profile in nearly 12,000 younger and older Danish women. *Sexually transmitted diseases* **35**:276-282.
82. **Bello BD, Spinillo A, Alberizzi P, Cesari S, Gardella B, D'Ambrosio G, Roccio M, Silini EM.** 2009. Cervical infections by multiple human papillomavirus (HPV) genotypes: Prevalence and impact on the risk of precancerous epithelial lesions. *Journal of medical virology* **81**:703-712.
83. **Levi JE, Kleter B, Quint WG, Fink MC, Canto CL, Matsubara R, Linhares I, Segurado A, Vanderborght B, Neto JE, Van Doorn LJ.** 2002. High prevalence of human papillomavirus (HPV) infections and high frequency of multiple HPV genotypes in human immunodeficiency virus-infected women in Brazil. *Journal of clinical microbiology* **40**:3341-3345.
84. **Trottier H, Mahmud S, Costa MC, Sobrinho JP, Duarte-Franco E, Rohan TE, Ferenczy A, Villa LL, Franco EL.** 2006. Human papillomavirus infections with multiple types and risk of cervical neoplasia. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **15**:1274-1280.
85. **Herrero R, Hildesheim A, Bratti C, Sherman ME, Hutchinson M, Morales J, Balmaceda I, Greenberg MD, Alfaro M, Burk RD, Wacholder S, Plummer M, Schiffman M.** 2000. Population-based study of human papillomavirus infection and cervical neoplasia in rural Costa Rica. *Journal of the National Cancer Institute* **92**:464-474.
86. **Cuschieri KS, Cubie HA, Whitley MW, Seagar AL, Arends MJ, Moore C, Gilkisson G, McGoogan E.** 2004. Multiple high risk HPV infections are common in cervical neoplasia and young women in a cervical screening population. *Journal of clinical pathology* **57**:68-72.
87. **Fontaine J, Hankins C, Mayrand MH, Lefevre J, Money D, Gagnon S, Rachlis A, Poursaux K, Ferenczy A, Coutlee F.** 2005. High levels of HPV-16 DNA are associated with high-grade cervical lesions in women at risk or infected with HIV. *AIDS (London, England)* **19**:785-794.
88. **Josefsson AM, Magnusson PK, Ylitalo N, Sorensen P, Qvarforth-Tubbin P, Andersen PK, Melbye M, Adami HO, Gyllenstein UB.** 2000. Viral load of human papilloma virus 16 as a determinant for development of cervical carcinoma in situ: a nested case-control study. *Lancet* **355**:2189-2193.
89. **van Duin M, Snijders PJ, Schrijnemakers HF, Voorhorst FJ, Rozendaal L, Nobbenhuis MA, van den Brule AJ, Verheijen RH, Helmerhorst TJ, Meijer CJ.** 2002. Human papillomavirus 16 load in normal and abnormal cervical scrapes: an indicator of CIN II/III and viral clearance. *International journal of cancer. Journal international du cancer* **98**:590-595.
90. **Lefevre J, Hankins C, Money D, Rachlis A, Poursaux K, Coutlee F.** 2004. Human papillomavirus type 16 viral load is higher in human immunodeficiency

- virus-seropositive women with high-grade squamous intraepithelial lesions than in those with normal cytology smears. *Journal of clinical microbiology* **42**:2212-2215.
91. **Monnier-Benoit S, Dalstein V, Riethmuller D, Lalaoui N, Mougin C, Pretet JL.** 2006. Dynamics of HPV16 DNA load reflect the natural history of cervical HPV-associated lesions. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology* **35**:270-277.
  92. **Wu Y, Chen Y, Li L, Yu G, Zhang Y, He Y.** 2006. Associations of high-risk HPV types and viral load with cervical cancer in China. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology* **35**:264-269.
  93. **Flores R, Papenfuss M, Klimecki WT, Giuliano AR.** 2006. Cross-sectional analysis of oncogenic HPV viral load and cervical intraepithelial neoplasia. *International journal of cancer. Journal international du cancer* **118**:1187-1193.
  94. **Moberg M, Gustavsson I, Gyllensten U.** 2004. Type-specific associations of human papillomavirus load with risk of developing cervical carcinoma in situ. *International journal of cancer. Journal international du cancer* **112**:854-859.
  95. **Azizi N, Brazete J, Hankins C, Money D, Fontaine J, Koushik A, Rachlis A, Pourreaux K, Ferenczy A, Franco E, Coutlee F.** 2008. Influence of human papillomavirus type 16 (HPV-16) E2 polymorphism on quantification of HPV-16 episomal and integrated DNA in cervicovaginal lavages from women with cervical intraepithelial neoplasia. *The Journal of general virology* **89**:1716-1728.
  96. **Alvarez J, de Pokomandy A, Rouleau D, Ghattas G, Vezina S, Cote P, Allaire G, Hadjeres R, Franco EL, Coutlee F.** 2010. Episomal and integrated human papillomavirus type 16 loads and anal intraepithelial neoplasia in HIV-seropositive men. *AIDS (London, England)* **24**:2355-2363.
  97. **Cheung JL, Cheung TH, Ng CW, Yu MY, Wong MC, Siu SS, Yim SF, Chan PK.** 2009. Analysis of human papillomavirus type 18 load and integration status from low-grade cervical lesion to invasive cervical cancer. *Journal of clinical microbiology* **47**:287-293.
  98. **Ylitalo N, Sorensen P, Josefsson AM, Magnusson PK, Andersen PK, Ponten J, Adami HO, Gyllensten UB, Melbye M.** 2000. Consistent high viral load of human papillomavirus 16 and risk of cervical carcinoma in situ: a nested case-control study. *Lancet* **355**:2194-2198.
  99. **Plummer M, Herrero R, Franceschi S, Meijer CJ, Snijders P, Bosch FX, de Sanjose S, Munoz N.** 2003. Smoking and cervical cancer: pooled analysis of the IARC multi-centric case--control study. *Cancer causes & control : CCC* **14**:805-814.
  100. **Castellsague X, Munoz N.** 2003. Chapter 3: Cofactors in human papillomavirus carcinogenesis--role of parity, oral contraceptives, and tobacco smoking. *Journal of the National Cancer Institute. Monographs*:20-28.
  101. **Syrjanen K, Shabalova I, Petrovichev N, Kozachenko V, Zakharova T, Pajanidi J, Podistov J, Chemeris G, Sozaeva L, Lipova E, Tsidaeva I, Ivanchenko O, Pshepurko A, Zakharenko S, Nerovjna R, Kljukina L, Erokhina O, Branovskaja M, Nikitina M, Grunberga V, Grunberg A, Juschenko A, Santopietro R, Cintonino M, Tosi P, Syrjanen S.** 2007. Smoking is an independent risk factor for oncogenic human papillomavirus (HPV) infections but not for high-grade CIN. *European journal of epidemiology* **22**:723-735.

102. **Harris TG, Kulasingam SL, Kiviat NB, Mao C, Agoff SN, Feng Q, Koutsky LA.** 2004. Cigarette smoking, oncogenic human papillomavirus, Ki-67 antigen, and cervical intraepithelial neoplasia. *American journal of epidemiology* **159**:834-842.
103. **Collins S, Rollason TP, Young LS, Woodman CB.** 2010. Cigarette smoking is an independent risk factor for cervical intraepithelial neoplasia in young women: a longitudinal study. *European journal of cancer (Oxford, England : 1990)* **46**:405-411.
104. 2004. Tobacco smoke and involuntary smoking. IARC monographs on the evaluation of carcinogenic risks to humans / World Health Organization, International Agency for Research on Cancer **83**:1-1438.
105. **Gunnell AS, Tran TN, Torrang A, Dickman PW, Sparen P, Palmgren J, Ylitalo N.** 2006. Synergy between cigarette smoking and human papillomavirus type 16 in cervical cancer in situ development. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **15**:2141-2147.
106. **Roura E, Castellsague X, Pawlita M, Travier N, Waterboer T, Margall N, Bosch FX, de Sanjose S, Dillner J, Gram IT, Tjonneland A, Munk C, Pala V, Palli D, Khaw KT, Barnabas RV, Overvad K, Clavel-Chapelon F, Boutron-Ruault MC, Fagherazzi G, Kaaks R, Lukanova A, Steffen A, Trichopoulou A, Trichopoulos D, Klinaki E, Tumino R, Sacerdote C, Panico S, Bueno-de-Mesquita HB, Peeters PH, Lund E, Weiderpass E, Redondo ML, Sanchez MJ, Tormo MJ, Barricarte A, Larranaga N, Ekstrom J, Hortlund M, Lindquist D, Wareham N, Travis RC, Rinaldi S, Tommasino M, Franceschi S, Riboli E.** 2014. Smoking as a major risk factor for cervical cancer and pre-cancer: results from the EPIC cohort. *International journal of cancer. Journal international du cancer* **135**:453-466.
107. **Luhn P, Walker J, Schiffman M, Zuna RE, Dunn ST, Gold MA, Smith K, Mathews C, Allen RA, Zhang R, Wang S, Wentzensen N.** 2013. The role of cofactors in the progression from human papillomavirus infection to cervical cancer. *Gynecologic oncology* **128**:265-270.
108. **Ribeiro AA, Costa MC, Alves RR, Villa LL, Saddi VA, Carneiro MA, Zeferino LC, Rabelo-Santos SH.** 2015. HPV infection and cervical neoplasia: associated risk factors. *Infectious agents and cancer* **10**:16.
109. 2009. Cervical carcinoma and sexual behavior: collaborative reanalysis of individual data on 15,461 women with cervical carcinoma and 29,164 women without cervical carcinoma from 21 epidemiological studies. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **18**:1060-1069.
110. **Plummer M, Peto J, Franceschi S.** 2012. Time since first sexual intercourse and the risk of cervical cancer. *International journal of cancer. Journal international du cancer* **130**:2638-2644.
111. **Burkett BJ, Peterson CM, Birch LM, Brennan C, Nuckols ML, Ward BE, Crum CP.** 1992. The relationship between contraceptives, sexual practices, and cervical human papillomavirus infection among a college population. *Journal of clinical epidemiology* **45**:1295-1302.
112. **Deacon JM, Evans CD, Yule R, Desai M, Binns W, Taylor C, Peto J.** 2000. Sexual behaviour and smoking as determinants of cervical HPV infection and of

- CIN3 among those infected: a case-control study nested within the Manchester cohort. *British journal of cancer* **83**:1565-1572.
113. **Moscicki AB, Hills NK, Shiboski S, Darragh TM, Jay N, Powell K, Hanson E, Miller SB, Farhat S, Palefsky J.** 1999. Risk factors for abnormal anal cytology in young heterosexual women. *Cancer epidemiology, biomarkers & prevention* : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology **8**:173-178.
  114. **Dixon AR, Pringle JH, Holmes JT, Watkin DF.** 1991. Cervical intraepithelial neoplasia and squamous cell carcinoma of the anus in sexually active women. *Postgraduate medical journal* **67**:557-559.
  115. **Sastre-Garau X, Loste MN, Vincent-Salomon A, Favre M, Mouret E, de la Rochefordiere A, Durand JC, Tartour E, Lepage V, Charron D.** 1996. Decreased frequency of HLA-DRB1 13 alleles in Frenchwomen with HPV-positive carcinoma of the cervix. *International journal of cancer. Journal international du cancer* **69**:159-164.
  116. **Cuzick J, Terry G, Ho L, Monaghan J, Lopes A, Clarkson P, Duncan I.** 2000. Association between high-risk HPV types, HLA DRB1\* and DQB1\* alleles and cervical cancer in British women. *British journal of cancer* **82**:1348-1352.
  117. **Beskow AH, Josefsson AM, Gyllensten UB.** 2001. HLA class II alleles associated with infection by HPV16 in cervical cancer in situ. *International journal of cancer. Journal international du cancer* **93**:817-822.
  118. **Apple RJ, Erlich HA, Klitz W, Manos MM, Becker TM, Wheeler CM.** 1994. HLA DR-DQ associations with cervical carcinoma show papillomavirus-type specificity. *Nature genetics* **6**:157-162.
  119. **Hu X, Zhang Z, Ma D, Huettner PC, Massad LS, Nguyen L, Borecki I, Rader JS.** 2010. TP53, MDM2, NQO1, and susceptibility to cervical cancer. *Cancer epidemiology, biomarkers & prevention* : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology **19**:755-761.
  120. **Hildesheim A, Wang SS.** 2002. Host and viral genetics and risk of cervical cancer: a review. *Virus research* **89**:229-240.
  121. **Stanczuk GA, Sibanda EN, Perrey C, Chirara M, Pravica V, Hutchinson IV, Tswana SA.** 2001. Cancer of the uterine cervix may be significantly associated with a gene polymorphism coding for increased IL-10 production. *International journal of cancer. Journal international du cancer* **94**:792-794.
  122. **Goodman MT, McDuffie K, Hernandez B, Wilkens LR, Bertram CC, Killeen J, Le Marchand L, Selhub J, Murphy S, Donlon TA.** 2001. Association of methylenetetrahydrofolate reductase polymorphism C677T and dietary folate with the risk of cervical dysplasia. *Cancer epidemiology, biomarkers & prevention* : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology **10**:1275-1280.
  123. **Nunobiki O, Ueda M, Toji E, Yamamoto M, Akashi K, Sato N, Izuma S, Torii K, Tanaka I, Okamoto Y, Noda S.** 2011. Genetic Polymorphism of Cancer Susceptibility Genes and HPV Infection in Cervical Carcinogenesis. *Pathology research international* **2011**:364069.
  124. **Zhen S, Hu CM, Bian LH.** 2013. Glutathione S-transferase polymorphism interactions with smoking status and HPV infection in cervical cancer risk: an evidence-based meta-analysis. *PloS one* **8**:e83497.



125. **Sharma A, Sharma JK, Murthy NS, Mitra AB.** 2004. Polymorphisms at GSTM1 and GSTT1 gene loci and susceptibility to cervical cancer in Indian population. *Neoplasma* **51**:12-16.
126. **Amaral CM, Cetkovska K, Gurgel AP, Cardoso MV, Chagas BS, Paiva Junior SS, de Lima Rde C, Silva-Neto JC, Silva LA, Muniz MT, Balbino VQ, Freitas AC.** 2014. MDM2 polymorphism associated with the development of cervical lesions in women infected with Human papillomavirus and using of oral contraceptives. *Infectious agents and cancer* **9**:24.
127. **Kordi Tamandani DM, Sobti RC, Shekari M.** 2008. Association of Fas-670 gene polymorphism with risk of cervical cancer in North Indian population. *Clinical and experimental obstetrics & gynecology* **35**:183-186.
128. **Lai HC, Sytwu HK, Sun CA, Yu MH, Yu CP, Liu HS, Chang CC, Chu TY.** 2003. Single nucleotide polymorphism at Fas promoter is associated with cervical carcinogenesis. *International journal of cancer. Journal international du cancer* **103**:221-225.
129. **Storey A, Thomas M, Kalita A, Harwood C, Gardiol D, Mantovani F, Breuer J, Leigh IM, Matlashewski G, Banks L.** 1998. Role of a p53 polymorphism in the development of human papillomavirus-associated cancer. *Nature* **393**:229-234.
130. **Hildesheim A, Schiffman M, Brinton LA, Fraumeni JF, Jr., Herrero R, Bratti MC, Schwartz P, Mortel R, Barnes W, Greenberg M, McGowan L, Scott DR, Martin M, Herrera JE, Carrington M.** 1998. p53 polymorphism and risk of cervical cancer. *Nature* **396**:531-532.
131. **Josefsson AM, Magnusson PK, Ylitalo N, Quarforth-Tubbin P, Ponten J, Adami HO, Gyllensten UB.** 1998. p53 polymorphism and risk of cervical cancer. *Nature* **396**:531; author reply 532.
132. **Rosenthal AN, Ryan A, Al-Jehani RM, Storey A, Harwood CA, Jacobs IJ.** 1998. p53 codon 72 polymorphism and risk of cervical cancer in UK. *Lancet* **352**:871-872.
133. **Helland A, Langerod A, Johnsen H, Olsen AO, Skovlund E, Borresen-Dale AL.** 1998. p53 polymorphism and risk of cervical cancer. *Nature* **396**:530-531; author reply 532.
134. **Gober MD, Rady PL, He Q, Tucker SB, Tyring SK, Gaspari AA.** 2007. Novel homozygous frameshift mutation of EVER1 gene in an epidermodysplasia verruciformis patient. *The Journal of investigative dermatology* **127**:817-820.
135. **Tate G, Suzuki T, Kishimoto K, Mitsuya T.** 2004. Novel mutations of EVER1/TMC6 gene in a Japanese patient with epidermodysplasia verruciformis. *Journal of human genetics* **49**:223-225.
136. **Sun XK, Chen JF, Xu AE.** 2005. A homozygous nonsense mutation in the EVER2 gene leads to epidermodysplasia verruciformis. *Clinical and experimental dermatology* **30**:573-574.
137. **Dugue PA, Rebolj M, Garred P, Lyng E.** 2013. Immunosuppression and risk of cervical cancer. *Expert review of anticancer therapy* **13**:29-42.
138. **Veroux M, Corona D, Scalia G, Garozzo V, Gagliano M, Giuffrida G, Costanzo CM, Giaquinta A, Palermo I, Zappala D, Tallarita T, Zerbo D, Russo R, Cappellani A, Franchina C, Scriffignano V, Veroux P.** 2009. Surveillance of human papilloma virus infection and cervical cancer in kidney transplant recipients: preliminary data. *Transplantation proceedings* **41**:1191-1194.

139. **Castellsague X, Bosch FX, Munoz N.** 2002. Environmental co-factors in HPV carcinogenesis. *Virus research* **89**:191-199.
140. **Hawes SE, Critchlow CW, Faye Niang MA, Diouf MB, Diop A, Toure P, Aziz Kasse A, Dembele B, Salif Sow P, Coll-Seck AM, Kuypers JM, Kiviat NB.** 2003. Increased risk of high-grade cervical squamous intraepithelial lesions and invasive cervical cancer among African women with human immunodeficiency virus type 1 and 2 infections. *The Journal of infectious diseases* **188**:555-563.
141. **Holmes RS, Hawes SE, Toure P, Dem A, Feng Q, Weiss NS, Kiviat NB.** 2009. HIV infection as a risk factor for cervical cancer and cervical intraepithelial neoplasia in Senegal. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **18**:2442-2446.
142. **Smith JS, Bosetti C, Munoz N, Herrero R, Bosch FX, Eluf-Neto J, Meijer CJ, Van Den Brule AJ, Franceschi S, Peeling RW.** 2004. Chlamydia trachomatis and invasive cervical cancer: a pooled analysis of the IARC multicentric case-control study. *International journal of cancer. Journal international du cancer* **111**:431-439.
143. **Smith JS, Herrero R, Bosetti C, Munoz N, Bosch FX, Eluf-Neto J, Castellsague X, Meijer CJ, Van den Brule AJ, Franceschi S, Ashley R.** 2002. Herpes simplex virus-2 as a human papillomavirus cofactor in the etiology of invasive cervical cancer. *Journal of the National Cancer Institute* **94**:1604-1613.
144. **Koskela P, Anttila T, Bjorge T, Brunsvig A, Dillner J, Hakama M, Hakulinen T, Jellum E, Lehtinen M, Lenner P, Luostarinen T, Pukkala E, Saikku P, Thoresen S, Youngman L, Paavonen J.** 2000. Chlamydia trachomatis infection as a risk factor for invasive cervical cancer. *International journal of cancer. Journal international du cancer* **85**:35-39.
145. **Zhao Y, Cao X, Zheng Y, Tang J, Cai W, Wang H, Gao Y, Wang Y.** 2012. Relationship between cervical disease and infection with human papillomavirus types 16 and 18, and herpes simplex virus 1 and 2. *Journal of medical virology* **84**:1920-1927.
146. **Lehtinen M, Koskela P, Jellum E, Bloigu A, Anttila T, Hallmans G, Luukkaala T, Thoresen S, Youngman L, Dillner J, Hakama M.** 2002. Herpes simplex virus and risk of cervical cancer: a longitudinal, nested case-control study in the nordic countries. *American journal of epidemiology* **156**:687-692.
147. **Naucler P, Chen HC, Persson K, You SL, Hsieh CY, Sun CA, Dillner J, Chen CJ.** 2007. Seroprevalence of human papillomaviruses and Chlamydia trachomatis and cervical cancer risk: nested case-control study. *The Journal of general virology* **88**:814-822.
148. **Jones C.** 1995. Cervical cancer: is herpes simplex virus type II a cofactor? *Clinical microbiology reviews* **8**:549-556.
149. **Moreno V, Bosch FX, Munoz N, Meijer CJ, Shah KV, Walboomers JM, Herrero R, Franceschi S.** 2002. Effect of oral contraceptives on risk of cervical cancer in women with human papillomavirus infection: the IARC multicentric case-control study. *Lancet* **359**:1085-1092.
150. **Negrini BP, Schiffman MH, Kurman RJ, Barnes W, Lannom L, Malley K, Brinton LA, Delgado G, Jones S, Tchabo JG, et al.** 1990. Oral contraceptive use, human papillomavirus infection, and risk of early cytological abnormalities of the cervix. *Cancer research* **50**:4670-4675.

151. **Jensen KE, Schmiedel S, Norrild B, Frederiksen K, Iftner T, Kjaer SK.** 2013. Parity as a cofactor for high-grade cervical disease among women with persistent human papillomavirus infection: a 13-year follow-up. *British journal of cancer* **108**:234-239.
152. **Elson DA, Riley RR, Lacey A, Thordarson G, Talamantes FJ, Arbeit JM.** 2000. Sensitivity of the cervical transformation zone to estrogen-induced squamous carcinogenesis. *Cancer research* **60**:1267-1275.
153. **Kim CJ, Um SJ, Kim TY, Kim EJ, Park TC, Kim SJ, Namkoong SE, Park JS.** 2000. Regulation of cell growth and HPV genes by exogenous estrogen in cervical cancer cells. *International journal of gynecological cancer : official journal of the International Gynecological Cancer Society* **10**:157-164.
154. **Fisher B, Costantino JP, Wickerham DL, Redmond CK, Kavanah M, Cronin WM, Vogel V, Robidoux A, Dimitrov N, Atkins J, Daly M, Wieand S, Tan-Chiu E, Ford L, Wolmark N.** 1998. Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *Journal of the National Cancer Institute* **90**:1371-1388.
155. **Fisher B, Costantino JP, Redmond CK, Fisher ER, Wickerham DL, Cronin WM.** 1994. Endometrial cancer in tamoxifen-treated breast cancer patients: findings from the National Surgical Adjuvant Breast and Bowel Project (NSABP) B-14. *Journal of the National Cancer Institute* **86**:527-537.
156. **Groves IJ, Coleman N.** 2014. Pathogenesis of human papillomavirus-associated mucosal disease. *The Journal of pathology*.
157. **Lehoux M, D'Abramo CM, Archambault J.** 2009. Molecular mechanisms of human papillomavirus-induced carcinogenesis. *Public health genomics* **12**:268-280.
158. **Doorbar J.** 2006. Molecular biology of human papillomavirus infection and cervical cancer. *Clinical science (London, England : 1979)* **110**:525-541.
159. **Hebner CM, Laimins LA.** 2006. Human papillomaviruses: basic mechanisms of pathogenesis and oncogenicity. *Reviews in medical virology* **16**:83-97.
160. **Schiffman MH, Brinton LA.** 1995. The epidemiology of cervical carcinogenesis. *Cancer* **76**:1888-1901.
161. **Doorbar J.** 2013. Latent papillomavirus infections and their regulation. *Current opinion in virology* **3**:416-421.
162. **Vande Pol SB, Klingelutz AJ.** 2013. Papillomavirus E6 oncoproteins. *Virology* **445**:115-137.
163. **Roman A, Munger K.** 2013. The papillomavirus E7 proteins. *Virology* **445**:138-168.
164. **Shin HJ, Joo J, Yoon JH, Yoo CW, Kim JY.** 2014. Physical status of human papillomavirus integration in cervical cancer is associated with treatment outcome of the patients treated with radiotherapy. *PloS one* **9**:e78995.
165. **Kristiansen E, Jenkins A, Holm R.** 1994. Coexistence of episomal and integrated HPV16 DNA in squamous cell carcinoma of the cervix. *Journal of clinical pathology* **47**:253-256.
166. **Park JS, Hwang ES, Park SN, Ahn HK, Um SJ, Kim CJ, Kim SJ, Namkoong SE.** 1997. Physical status and expression of HPV genes in cervical cancers. *Gynecologic oncology* **65**:121-129.
167. **Cheung JL, Cheung TH, Yu MY, Chan PK.** 2013. Virological characteristics of cervical cancers carrying pure episomal form of HPV16 genome. *Gynecologic oncology* **131**:374-379.

168. **Smith PP, Friedman CL, Bryant EM, McDougall JK.** 1992. Viral integration and fragile sites in human papillomavirus-immortalized human keratinocyte cell lines. *Genes, chromosomes & cancer* **5**:150-157.
169. **Thorland EC, Myers SL, Gostout BS, Smith DI.** 2003. Common fragile sites are preferential targets for HPV16 integrations in cervical tumors. *Oncogene* **22**:1225-1237.
170. **Demeret C, Desaintes C, Yaniv M, Thierry F.** 1997. Different mechanisms contribute to the E2-mediated transcriptional repression of human papillomavirus type 18 viral oncogenes. *Journal of virology* **71**:9343-9349.
171. **Badal V, Chuang LS, Tan EH, Badal S, Villa LL, Wheeler CM, Li BF, Bernard HU.** 2003. CpG methylation of human papillomavirus type 16 DNA in cervical cancer cell lines and in clinical specimens: genomic hypomethylation correlates with carcinogenic progression. *Journal of virology* **77**:6227-6234.
172. **Reuschenbach M, Huebbers CU, Prigge ES, Bermejo JL, Kalteis MS, Preuss SF, Seuthe IM, Kolligs J, Speel EJ, Olthof N, Kremer B, Wagner S, Klussmann JP, Vinokurova S, von Knebel Doeberitz M.** 2015. Methylation status of HPV16 E2-binding sites classifies subtypes of HPV-associated oropharyngeal cancers. *Cancer* **121**:1966-1976.
173. **Chaiwongkot A, Vinokurova S, Pientong C, Ekalaksananan T, Kongyingyoes B, Kleebkaow P, Chumworathayi B, Patarapadungkit N, Reuschenbach M, von Knebel Doeberitz M.** 2013. Differential methylation of E2 binding sites in episomal and integrated HPV 16 genomes in preinvasive and invasive cervical lesions. *International journal of cancer. Journal international du cancer* **132**:2087-2094.
174. **Thain A, Jenkins O, Clarke AR, Gaston K.** 1996. CpG methylation directly inhibits binding of the human papillomavirus type 16 E2 protein to specific DNA sequences. *Journal of virology* **70**:7233-7235.
175. **Sanchez IE, Dellarole M, Gaston K, de Prat Gay G.** 2008. Comprehensive comparison of the interaction of the E2 master regulator with its cognate target DNA sites in 73 human papillomavirus types by sequence statistics. *Nucleic acids research* **36**:756-769.
176. **Wooldridge TR, Laimins LA.** 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.
177. **del Mar Pena LM, Laimins LA.** 2001. Differentiation-dependent chromatin rearrangement coincides with activation of human papillomavirus type 31 late gene expression. *Journal of virology* **75**:10005-10013.
178. **De-Castro Arce J, Gockel-Krzikalla E, Rosl F.** 2012. Silencing of multi-copy HPV16 by viral self-methylation and chromatin occlusion: a model for epigenetic virus-host interaction. *Human molecular genetics* **21**:1693-1705.
179. **Beglin M, Melar-New M, Laimins L.** 2009. Human papillomaviruses and the interferon response. *Journal of interferon & cytokine research : the official journal of the International Society for Interferon and Cytokine Research* **29**:629-635.
180. **Stanley MA.** 2012. Epithelial cell responses to infection with human papillomavirus. *Clinical microbiology reviews* **25**:215-222.
181. **Braun L, Durst M, Mikumo R, Crowley A, Robinson M.** 1992. Regulation of growth and gene expression in human papillomavirus-transformed keratinocytes by

- transforming growth factor-beta: implications for the control of papillomavirus infection. *Molecular carcinogenesis* **6**:100-111.
182. **Malejczyk J, Malejczyk M, Kock A, Urbanski A, Majewski S, Hunzelmann N, Jablonska S, Orth G, Luger TA.** 1992. Autocrine growth limitation of human papillomavirus type 16-harboring keratinocytes by constitutively released tumor necrosis factor-alpha. *Journal of immunology (Baltimore, Md. : 1950)* **149**:2702-2708.
  183. **Fensterl V, Sen GC.** 2009. Interferons and viral infections. *BioFactors (Oxford, England)* **35**:14-20.
  184. **Nees M, Geoghegan JM, Hyman T, Frank S, Miller L, Woodworth CD.** 2001. Papillomavirus type 16 oncogenes downregulate expression of interferon-responsive genes and upregulate proliferation-associated and NF-kappaB-responsive genes in cervical keratinocytes. *Journal of virology* **75**:4283-4296.
  185. **Barnard P, Payne E, McMillan NA.** 2000. The human papillomavirus E7 protein is able to inhibit the antiviral and anti-growth functions of interferon-alpha. *Virology* **277**:411-419.
  186. **Romagnani S.** 2000. T-cell subsets (Th1 versus Th2). *Annals of allergy, asthma & immunology : official publication of the American College of Allergy, Asthma, & Immunology* **85**:9-18; quiz 18, 21.
  187. **Giannini SL, Hanon E, Moris P, Van Mechelen M, Morel S, Dessy F, Fourneau MA, Colau B, Suzich J, Losonksy G, Martin MT, Dubin G, Wettendorff MA.** 2006. Enhanced humoral and memory B cellular immunity using HPV16/18 L1 VLP vaccine formulated with the MPL/aluminium salt combination (AS04) compared to aluminium salt only. *Vaccine* **24**:5937-5949.
  188. **Scott M, Nakagawa M, Moscicki AB.** 2001. Cell-mediated immune response to human papillomavirus infection. *Clinical and diagnostic laboratory immunology* **8**:209-220.
  189. **Nakagawa M, Stites DP, Farhat S, Sisler JR, Moss B, Kong F, Moscicki AB, Palefsky JM.** 1997. Cytotoxic T lymphocyte responses to E6 and E7 proteins of human papillomavirus type 16: relationship to cervical intraepithelial neoplasia. *The Journal of infectious diseases* **175**:927-931.
  190. **Bourgault Villada I, Moyal Barracco M, Ziol M, Chaboissier A, Barget N, Berville S, Paniel B, Jullian E, Clerici T, Maillere B, Guillet JG.** 2004. Spontaneous regression of grade 3 vulvar intraepithelial neoplasia associated with human papillomavirus-16-specific CD4(+) and CD8(+) T-cell responses. *Cancer research* **64**:8761-8766.
  191. **Monnier-Benoit S, Mauny F, Riethmuller D, Guerrini JS, Capilna M, Felix S, Seilles E, Mouglin C, Pretet JL.** 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. *Gynecologic oncology* **102**:22-31.
  192. **de Jong A, van Poelgeest MI, van der Hulst JM, Drijfhout JW, Fleuren GJ, Melief CJ, Kenter G, Offringa R, van der Burg SH.** 2004. Human papillomavirus type 16-positive cervical cancer is associated with impaired CD4+ T-cell immunity against early antigens E2 and E6. *Cancer research* **64**:5449-5455.
  193. **Xu Y, Zhu KJ, Zhu N, Jiang DH, Chen XZ, Cheng H.** 2009. Expression of Foxp3+CD4+CD25+ regulatory T cells and Th1/Th2, Tc1/Tc2 profiles in the peripheral blood of patients with condyloma acuminatum. *Clinical and experimental dermatology* **34**:229-235.

194. **Bais AG, Beckmann I, Lindemans J, Ewing PC, Meijer CJ, Snijders PJ, Helmerhorst TJ.** 2005. A shift to a peripheral Th2-type cytokine pattern during the carcinogenesis of cervical cancer becomes manifest in CIN III lesions. *Journal of clinical pathology* **58**:1096-1100.
195. **Shah W, Yan X, Jing L, Zhou Y, Chen H, Wang Y.** 2011. A reversed CD4/CD8 ratio of tumor-infiltrating lymphocytes and a high percentage of CD4(+)FOXP3(+) regulatory T cells are significantly associated with clinical outcome in squamous cell carcinoma of the cervix. *Cellular & molecular immunology* **8**:59-66.
196. **Torres LM, Cabrera T, Concha A, Oliva MR, Ruiz-Cabello F, Garrido F.** 1993. HLA class I expression and HPV-16 sequences in premalignant and malignant lesions of the cervix. *Tissue antigens* **41**:65-71.
197. **Vambutas A, Bonagura VR, Steinberg BM.** 2000. Altered expression of TAP-1 and major histocompatibility complex class I in laryngeal papillomatosis: correlation of TAP-1 with disease. *Clinical and diagnostic laboratory immunology* **7**:79-85.
198. **Québec Indspd.** 2011. Lignes directrices sur le dépistage du cancer du col utérin au Québec.
199. **Negri G, Menia E, Egarter-Vigl E, Vittadello F, Mian C.** 2003. ThinPrep versus conventional Papanicolaou smear in the cytologic follow-up of women with equivocal cervical smears. *Cancer* **99**:342-345.
200. **Peyton CL, Schiffman M, Lorincz AT, Hunt WC, Mielzynska I, Bratti C, Eaton S, Hildesheim A, Morera LA, Rodriguez AC, Herrero R, Sherman ME, Wheeler CM.** 1998. Comparison of PCR- and hybrid capture-based human papillomavirus detection systems using multiple cervical specimen collection strategies. *Journal of clinical microbiology* **36**:3248-3254.
201. **Herrero R.** 2009. Human papillomavirus (HPV) vaccines: limited cross-protection against additional HPV types. *The Journal of infectious diseases* **199**:919-922.
202. **Zhai L, Tumban E.** 2016. Gardasil-9: A global survey of projected efficacy. *Antiviral research* **130**:101-109.
203. **cancer Scd** 2016, posting date. Traitement du cancer du col de l'utérus. [Online.]
204. **Gadducci A, Tana R, Cosio S, Cionini L.** 2010. Treatment options in recurrent cervical cancer (Review). *Oncology letters* **1**:3-11.
205. **Siddik ZH.** 2003. Cisplatin: mode of cytotoxic action and molecular basis of resistance. *Oncogene* **22**:7265-7279.
206. **Tewari KS, Sill MW, Long HJ, 3rd, Penson RT, Huang H, Ramondetta LM, Landrum LM, Oaknin A, Reid TJ, Leitao MM, Michael HE, Monk BJ.** 2014. Improved survival with bevacizumab in advanced cervical cancer. *The New England journal of medicine* **370**:734-743.
207. **Lace MJ, Anson JR, Klingelutz AJ, Harada H, Taniguchi T, Bossler AD, Haugen TH, Turek LP.** 2009. Interferon-beta treatment increases human papillomavirus early gene transcription and viral plasmid genome replication by activating interferon regulatory factor (IRF)-1. *Carcinogenesis* **30**:1336-1344.
208. **Park JS, Kim EJ, Kwon HJ, Hwang ES, Namkoong SE, Um SJ.** 2000. Inactivation of interferon regulatory factor-1 tumor suppressor protein by HPV E7 oncoprotein. Implication for the E7-mediated immune evasion mechanism in cervical carcinogenesis. *The Journal of biological chemistry* **275**:6764-6769.

209. **Ronco LV, Karpova AY, Vidal M, Howley PM.** 1998. Human papillomavirus 16 E6 oncoprotein binds to interferon regulatory factor-3 and inhibits its transcriptional activity. *Genes & development* **12**:2061-2072.
210. **Hong S, Mehta KP, Laimins LA.** 2011. Suppression of STAT-1 expression by human papillomaviruses is necessary for differentiation-dependent genome amplification and plasmid maintenance. *Journal of virology* **85**:9486-9494.
211. **Herdman MT, Pett MR, Roberts I, Alazawi WO, Teschendorff AE, Zhang XY, Stanley MA, Coleman N.** 2006. Interferon-beta treatment of cervical keratinocytes naturally infected with human papillomavirus 16 episomes promotes rapid reduction in episome numbers and emergence of latent integrants. *Carcinogenesis* **27**:2341-2353.
212. **Kim KY, Blatt L, Taylor MW.** 2000. The effects of interferon on the expression of human papillomavirus oncogenes. *The Journal of general virology* **81**:695-700.
213. **Singhania R, Khairuddin N, Clarke D, McMillan NA.** 2012. RNA interference for the treatment of papillomavirus disease. *The open virology journal* **6**:204-215.
214. **Jiang M, Milner J.** 2002. Selective silencing of viral gene expression in HPV-positive human cervical carcinoma cells treated with siRNA, a primer of RNA interference. *Oncogene* **21**:6041-6048.
215. **Jung HS, Erkin OC, Kwon MJ, Kim SH, Jung JI, Oh YK, Her SW, Ju W, Choi YL, Song SY, Kim JK, Kim YD, Shim GY, Shin YK.** 2012. The synergistic therapeutic effect of cisplatin with Human papillomavirus E6/E7 short interfering RNA on cervical cancer cell lines in vitro and in vivo. *International journal of cancer. Journal international du cancer* **130**:1925-1936.
216. **Putral LN, Bywater MJ, Gu W, Saunders NA, Gabrielli BG, Leggatt GR, McMillan NA.** 2005. RNA interference against human papillomavirus oncogenes in cervical cancer cells results in increased sensitivity to cisplatin. *Molecular pharmacology* **68**:1311-1319.
217. **Hall AH, Alexander KA.** 2003. RNA interference of human papillomavirus type 18 E6 and E7 induces senescence in HeLa cells. *Journal of virology* **77**:6066-6069.
218. **Butz K, Ristriani T, Hengstermann A, Denk C, Scheffner M, Hoppe-Seyler F.** 2003. siRNA targeting of the viral E6 oncogene efficiently kills human papillomavirus-positive cancer cells. *Oncogene* **22**:5938-5945.
219. **Yamato K, Fen J, Kobuchi H, Nasu Y, Yamada T, Nishihara T, Ikeda Y, Kizaki M, Yoshinouchi M.** 2006. Induction of cell death in human papillomavirus 18-positive cervical cancer cells by E6 siRNA. *Cancer gene therapy* **13**:234-241.
220. **Faucher AM, White PW, Brochu C, Grand-Maitre C, Rancourt J, Fazal G.** 2004. Discovery of small-molecule inhibitors of the ATPase activity of human papillomavirus E1 helicase. *Journal of medicinal chemistry* **47**:18-21.
221. **White PW, Faucher AM, Massariol MJ, Welchner E, Rancourt J, Cartier M, Archambault J.** 2005. Biphenylsulfonacetic acid inhibitors of the human papillomavirus type 6 E1 helicase inhibit ATP hydrolysis by an allosteric mechanism involving tyrosine 486. *Antimicrobial agents and chemotherapy* **49**:4834-4842.
222. **Bergvall M, Melendy T, Archambault J.** 2013. The E1 proteins. *Virology* **445**:35-56.
223. **McBride AA.** 2013. The papillomavirus E2 proteins. *Virology* **445**:57-79.
224. **Ilves I, Kivi S, Ustav M.** 1999. Long-term episomal maintenance of bovine papillomavirus type 1 plasmids is determined by attachment to host chromosomes,

- which is mediated by the viral E2 protein and its binding sites. *Journal of virology* **73**:4404-4412.
225. **You J, Croyle JL, Nishimura A, Ozato K, Howley PM.** 2004. Interaction of the bovine papillomavirus E2 protein with Brd4 tethers the viral DNA to host mitotic chromosomes. *Cell* **117**:349-360.
226. **Yoakim C, Ogilvie WW, Goudreau N, Naud J, Hache B, O'Meara JA, Cordingley MG, Archambault J, White PW.** 2003. Discovery of the first series of inhibitors of human papillomavirus type 11: inhibition of the assembly of the E1-E2-Origin DNA complex. *Bioorganic & medicinal chemistry letters* **13**:2539-2541.
227. **Wang Y, Coulombe R, Cameron DR, Thauvette L, Massariol MJ, Amon LM, Fink D, Titolo S, Welchner E, Yoakim C, Archambault J, White PW.** 2004. Crystal structure of the E2 transactivation domain of human papillomavirus type 11 bound to a protein interaction inhibitor. *The Journal of biological chemistry* **279**:6976-6985.
228. **White PW, Titolo S, Brault K, Thauvette L, Pelletier A, Welchner E, Bourgon L, Doyon L, Ogilvie WW, Yoakim C, Cordingley MG, Archambault J.** 2003. Inhibition of human papillomavirus DNA replication by small molecule antagonists of the E1-E2 protein interaction. *The Journal of biological chemistry* **278**:26765-26772.
229. **Cote-Martin A, Moody C, Fradet-Turcotte A, D'Abramo CM, Lehoux M, Joubert S, Poirier GG, Coulombe B, Laimins LA, Archambault J.** 2008. Human papillomavirus E1 helicase interacts with the WD repeat protein p80 to promote maintenance of the viral genome in keratinocytes. *Journal of virology* **82**:1271-1283.
230. **Lehoux M, Fradet-Turcotte A, Lussier-Price M, Omichinski JG, Archambault J.** 2012. Inhibition of human papillomavirus DNA replication by an E1-derived p80/UAF1-binding peptide. *Journal of virology* **86**:3486-3500.
231. **Fradet-Turcotte A, Morin G, Lehoux M, Bullock PA, Archambault J.** 2010. Development of quantitative and high-throughput assays of polyomavirus and papillomavirus DNA replication. *Virology* **399**:65-76.
232. **Baleja JD, Cherry JJ, Liu Z, Gao H, Nicklaus MC, Voigt JH, Chen JJ, Androphy EJ.** 2006. Identification of inhibitors to papillomavirus type 16 E6 protein based on three-dimensional structures of interacting proteins. *Antiviral research* **72**:49-59.
233. **Malecka KA, Fera D, Schultz DC, Hodawadekar S, Reichman M, Donover PS, Murphy ME, Marmorstein R.** 2014. Identification and characterization of small molecule human papillomavirus E6 inhibitors. *ACS chemical biology* **9**:1603-1612.
234. **Trimble CL, Peng S, Kos F, Gravitt P, Viscidi R, Sugar E, Pardoll D, Wu TC.** 2009. A phase I trial of a human papillomavirus DNA vaccine for HPV16+ cervical intraepithelial neoplasia 2/3. *Clinical cancer research : an official journal of the American Association for Cancer Research* **15**:361-367.
235. **Kenter GG, Welters MJ, Valentijn AR, Lowik MJ, Berends-van der Meer DM, Vloon AP, Essahsah F, Fathers LM, Offringa R, Drijfhout JW, Wafelman AR, Oostendorp J, Fleuren GJ, van der Burg SH, Melief CJ.** 2009. Vaccination against HPV-16 oncoproteins for vulvar intraepithelial neoplasia. *The New England journal of medicine* **361**:1838-1847.
236. **de Vos van Steenwijk PJ, Ramwadhoebe TH, Lowik MJ, van der Minne CE, Berends-van der Meer DM, Fathers LM, Valentijn AR, Oostendorp J, Fleuren**



- GJ, Hellebrekers BW, Welters MJ, van Poelgeest MI, Melief CJ, Kenter GG, van der Burg SH.** 2012. A placebo-controlled randomized HPV16 synthetic long-peptide vaccination study in women with high-grade cervical squamous intraepithelial lesions. *Cancer immunology, immunotherapy* : CII **61**:1485-1492.
237. **Kaufmann AM, Nieland JD, Jochmus I, Baur S, Friese K, Gabelsberger J, Giesecking F, Gissmann L, Glasschroder B, Grubert T, Hillemanns P, Hopfl R, Ikenberg H, Schwarz J, Karrasch M, Knoll A, Kuppers V, Lechmann M, Lelle RJ, Meissner H, Muller RT, Pawlita M, Petry KU, Pilch H, Walek E, Schneider A.** 2007. Vaccination trial with HPV16 L1E7 chimeric virus-like particles in women suffering from high grade cervical intraepithelial neoplasia (CIN 2/3). *International journal of cancer. Journal international du cancer* **121**:2794-2800.
238. **Einstein MH, Kadish AS, Burk RD, Kim MY, Wadler S, Streicher H, Goldberg GL, Runowicz CD.** 2007. Heat shock fusion protein-based immunotherapy for treatment of cervical intraepithelial neoplasia III. *Gynecologic oncology* **106**:453-460.
239. **Roman LD, Wilczynski S, Muderspach LI, Burnett AF, O'Meara A, Brinkman JA, Kast WM, Facio G, Felix JC, Aldana M, Weber JS.** 2007. A phase II study of Hsp-7 (SGN-00101) in women with high-grade cervical intraepithelial neoplasia. *Gynecologic oncology* **106**:558-566.
240. **Corona Gutierrez CM, Tinoco A, Navarro T, Contreras ML, Cortes RR, Calzado P, Reyes L, Posternak R, Morosoli G, Verde ML, Rosales R.** 2004. Therapeutic vaccination with MVA E2 can eliminate precancerous lesions (CIN 1, CIN 2, and CIN 3) associated with infection by oncogenic human papillomavirus. *Human gene therapy* **15**:421-431.
241. **Garcia-Hernandez E, Gonzalez-Sanchez JL, Andrade-Manzano A, Contreras ML, Padilla S, Guzman CC, Jimenez R, Reyes L, Morosoli G, Verde ML, Rosales R.** 2006. Regression of papilloma high-grade lesions (CIN 2 and CIN 3) is stimulated by therapeutic vaccination with MVA E2 recombinant vaccine. *Cancer gene therapy* **13**:592-597.
242. **Rosales R, Lopez-Contreras M, Rosales C, Magallanes-Molina JR, Gonzalez-Vergara R, Arroyo-Cazarez JM, Ricardez-Arenas A, Del Follo-Valencia A, Padilla-Arriaga S, Guerrero MV, Pirez MA, Arellano-Fiore C, Villarreal F.** 2014. Regression of human papillomavirus intraepithelial lesions is induced by MVA E2 therapeutic vaccine. *Human gene therapy* **25**:1035-1049.
243. **Del Vecchio AM, Romanczuk H, Howley PM, Baker CC.** 1992. Transient replication of human papillomavirus DNAs. *Journal of virology* **66**:5949-5958.
244. **Ustav M, Stenlund A.** 1991. Transient replication of BPV-1 requires two viral polypeptides encoded by the E1 and E2 open reading frames. *The EMBO journal* **10**:449-457.
245. **Kadaja M, Silla T, Ustav E, Ustav M.** 2009. Papillomavirus DNA replication - from initiation to genomic instability. *Virology* **384**:360-368.
246. **Fuchs E.** 1990. Epidermal differentiation: the bare essentials. *The Journal of cell biology* **111**:2807-2814.
247. **Giroglou T, Florin L, Schafer F, Streeck RE, Sapp M.** 2001. Human papillomavirus infection requires cell surface heparan sulfate. *Journal of virology* **75**:1565-1570.
248. **Joyce JG, Tung JS, Przysiecki CT, Cook JC, Lehman ED, Sands JA, Jansen KU, Keller PM.** 1999. The L1 major capsid protein of human papillomavirus type

- 11 recombinant virus-like particles interacts with heparin and cell-surface glycosaminoglycans on human keratinocytes. *The Journal of biological chemistry* **274**:5810-5822.
249. **Culp TD, Budgeon LR, Marinkovich MP, Meneguzzi G, Christensen ND.** 2006. Keratinocyte-secreted laminin 5 can function as a transient receptor for human papillomaviruses by binding virions and transferring them to adjacent cells. *Journal of virology* **80**:8940-8950.
  250. **Kines RC, Thompson CD, Lowy DR, Schiller JT, Day PM.** 2009. The initial steps leading to papillomavirus infection occur on the basement membrane prior to cell surface binding. *Proceedings of the National Academy of Sciences of the United States of America* **106**:20458-20463.
  251. **Richards RM, Lowy DR, Schiller JT, Day PM.** 2006. Cleavage of the papillomavirus minor capsid protein, L2, at a furin consensus site is necessary for infection. *Proceedings of the National Academy of Sciences of the United States of America* **103**:1522-1527.
  252. **Culp TD, Christensen ND.** 2004. Kinetics of in vitro adsorption and entry of papillomavirus virions. *Virology* **319**:152-161.
  253. **Day PM, Lowy DR, Schiller JT.** 2003. Papillomaviruses infect cells via a clathrin-dependent pathway. *Virology* **307**:1-11.
  254. **Selinka HC, Giroglou T, Sapp M.** 2002. Analysis of the infectious entry pathway of human papillomavirus type 33 pseudovirions. *Virology* **299**:279-287.
  255. **Smith JL, Campos SK, Ozbun MA.** 2007. Human papillomavirus type 31 uses a caveolin 1- and dynamin 2-mediated entry pathway for infection of human keratinocytes. *Journal of virology* **81**:9922-9931.
  256. **Smith JL, Campos SK, Wandinger-Ness A, Ozbun MA.** 2008. Caveolin-1-dependent infectious entry of human papillomavirus type 31 in human keratinocytes proceeds to the endosomal pathway for pH-dependent uncoating. *Journal of virology* **82**:9505-9512.
  257. **Duckert P, Brunak S, Blom N.** 2004. Prediction of proprotein convertase cleavage sites. *Protein engineering, design & selection : PEDS* **17**:107-112.
  258. **Day PM, Baker CC, Lowy DR, Schiller JT.** 2004. Establishment of papillomavirus infection is enhanced by promyelocytic leukemia protein (PML) expression. *Proceedings of the National Academy of Sciences of the United States of America* **101**:14252-14257.
  259. **Pyeon D, Pearce SM, Lank SM, Ahlquist P, Lambert PF.** 2009. Establishment of human papillomavirus infection requires cell cycle progression. *PLoS pathogens* **5**:e1000318.
  260. **Stanley MA, Browne HM, Appleby M, Minson AC.** 1989. Properties of a non-tumorigenic human cervical keratinocyte cell line. *International journal of cancer. Journal international du cancer* **43**:672-676.
  261. **Schmitt A, Rochat A, Zeltner R, Borenstein L, Barrandon Y, Wettstein FO, Iftner T.** 1996. The primary target cells of the high-risk cottontail rabbit papillomavirus colocalize with hair follicle stem cells. *Journal of virology* **70**:1912-1922.
  262. **Egawa K.** 2003. Do human papillomaviruses target epidermal stem cells? *Dermatology (Basel, Switzerland)* **207**:251-254.

263. **McPhillips MG, Oliveira JG, Spindler JE, Mitra R, McBride AA.** 2006. Brd4 is required for e2-mediated transcriptional activation but not genome partitioning of all papillomaviruses. *Journal of virology* **80**:9530-9543.
264. **Piirsoo M, Ustav E, Mandel T, Stenlund A, Ustav M.** 1996. Cis and trans requirements for stable episomal maintenance of the BPV-1 replicator. *The EMBO journal* **15**:1-11.
265. **Van Tine BA, Dao LD, Wu SY, Sonbuchner TM, Lin BY, Zou N, Chiang CM, Broker TR, Chow LT.** 2004. Human papillomavirus (HPV) origin-binding protein associates with mitotic spindles to enable viral DNA partitioning. *Proceedings of the National Academy of Sciences of the United States of America* **101**:4030-4035.
266. **Hoffmann R, Hirt B, Bechtold V, Beard P, Raj K.** 2006. Different modes of human papillomavirus DNA replication during maintenance. *Journal of virology* **80**:4431-4439.
267. **Kim K, Angeletti PC, Hassebroek EC, Lambert PF.** 2005. Identification of cis-acting elements that mediate the replication and maintenance of human papillomavirus type 16 genomes in *Saccharomyces cerevisiae*. *Journal of virology* **79**:5933-5942.
268. **Ravnan JB, Gilbert DM, Ten Hagen KG, Cohen SN.** 1992. Random-choice replication of extrachromosomal bovine papillomavirus (BPV) molecules in heterogeneous, clonally derived BPV-infected cell lines. *Journal of virology* **66**:6946-6952.
269. **Frattini MG, Lim HB, Laimins LA.** 1996. In vitro synthesis of oncogenic human papillomaviruses requires episomal genomes for differentiation-dependent late expression. *Proceedings of the National Academy of Sciences of the United States of America* **93**:3062-3067.
270. **Ruesch MN, Laimins LA.** 1998. Human papillomavirus oncoproteins alter differentiation-dependent cell cycle exit on suspension in semisolid medium. *Virology* **250**:19-29.
271. **Dollard SC, Wilson JL, Demeter LM, Bonnez W, Reichman RC, Broker TR, Chow LT.** 1992. Production of human papillomavirus and modulation of the infectious program in epithelial raft cultures. *OFF. Genes & development* **6**:1131-1142.
272. **Middleton K, Peh W, Southern S, Griffin H, Sotlar K, Nakahara T, El-Sherif A, Morris L, Seth R, Hibma M, Jenkins D, Lambert P, Coleman N, Doorbar J.** 2003. Organization of human papillomavirus productive cycle during neoplastic progression provides a basis for selection of diagnostic markers. *Journal of virology* **77**:10186-10201.
273. **Fehrmann F, Klumpp DJ, Laimins LA.** 2003. Human papillomavirus type 31 E5 protein supports cell cycle progression and activates late viral functions upon epithelial differentiation. *Journal of virology* **77**:2819-2831.
274. **Genther SM, Sterling S, Duensing S, Munger K, Sattler C, Lambert PF.** 2003. Quantitative role of the human papillomavirus type 16 E5 gene during the productive stage of the viral life cycle. *Journal of virology* **77**:2832-2842.
275. **Doorbar J, Foo C, Coleman N, Medcalf L, Hartley O, Prospero T, Naphine S, Sterling J, Winter G, Griffin H.** 1997. Characterization of events during the late stages of HPV16 infection in vivo using high-affinity synthetic Fabs to E4. *Virology* **238**:40-52.

276. **Peh WL, Middleton K, Christensen N, Nicholls P, Egawa K, Sotlar K, Brandsma J, Percival A, Lewis J, Liu WJ, Doorbar J.** 2002. Life cycle heterogeneity in animal models of human papillomavirus-associated disease. *Journal of virology* **76**:10401-10416.
277. **Bedell MA, Hudson JB, Golub TR, Turyk ME, Hosken M, Wilbanks GD, Laimins LA.** 1991. Amplification of human papillomavirus genomes in vitro is dependent on epithelial differentiation. *Journal of virology* **65**:2254-2260.
278. **Conway MJ, Meyers C.** 2009. Replication and assembly of human papillomaviruses. *Journal of dental research* **88**:307-317.
279. **Zhou J, Sun XY, Stenzel DJ, Frazer IH.** 1991. Expression of vaccinia recombinant HPV 16 L1 and L2 ORF proteins in epithelial cells is sufficient for assembly of HPV virion-like particles. *Virology* **185**:251-257.
280. **Holmgren SC, Patterson NA, Ozbun MA, Lambert PF.** 2005. The minor capsid protein L2 contributes to two steps in the human papillomavirus type 31 life cycle. *Journal of virology* **79**:3938-3948.
281. **Zhou J, Sun XY, Louis K, Frazer IH.** 1994. Interaction of human papillomavirus (HPV) type 16 capsid proteins with HPV DNA requires an intact L2 N-terminal sequence. *Journal of virology* **68**:619-625.
282. **Buck CB, Thompson CD, Pang YY, Lowy DR, Schiller JT.** 2005. Maturation of papillomavirus capsids. *Journal of virology* **79**:2839-2846.
283. **Modis Y, Trus BL, Harrison SC.** 2002. Atomic model of the papillomavirus capsid. *The EMBO journal* **21**:4754-4762.
284. **Cardone G, Moyer AL, Cheng N, Thompson CD, Dvoretzky I, Lowy DR, Schiller JT, Steven AC, Buck CB, Trus BL.** 2014. Maturation of the human papillomavirus 16 capsid. *mBio* **5**:e01104-01114.
285. **Florin L, Sapp C, Streeck RE, Sapp M.** 2002. Assembly and translocation of papillomavirus capsid proteins. *Journal of virology* **76**:10009-10014.
286. **Bryan JT, Brown DR.** 2000. Association of the human papillomavirus type 11 E1(E4) protein with cornified cell envelopes derived from infected genital epithelium. *Virology* **277**:262-269.
287. **Brown DR, Kitchin D, Qadadri B, Neptune N, Batteiger T, Ermel A.** 2006. The human papillomavirus type 11 E1--E4 protein is a transglutaminase 3 substrate and induces abnormalities of the cornified cell envelope. *Virology* **345**:290-298.
288. **Bryan JT, Brown DR.** 2001. Transmission of human papillomavirus type 11 infection by desquamated cornified cells. *Virology* **281**:35-42.
289. **Wentzensen N, Vinokurova S, von Knebel Doeberitz M.** 2004. Systematic review of genomic integration sites of human papillomavirus genomes in epithelial dysplasia and invasive cancer of the female lower genital tract. *Cancer research* **64**:3878-3884.
290. **Baker TS, Newcomb WW, Olson NH, Cowser LM, Olson C, Brown JC.** 1991. Structures of bovine and human papillomaviruses. Analysis by cryoelectron microscopy and three-dimensional image reconstruction. *Biophysical journal* **60**:1445-1456.
291. **Finnen RL, Erickson KD, Chen XS, Garcea RL.** 2003. Interactions between papillomavirus L1 and L2 capsid proteins. *Journal of virology* **77**:4818-4826.
292. **Liddington RC, Yan Y, Moulai J, Sahli R, Benjamin TL, Harrison SC.** 1991. Structure of simian virus 40 at 3.8-Å resolution. *Nature* **354**:278-284.

293. **Favre M, Breitburd F, Croissant O, Orth G.** 1977. Chromatin-like structures obtained after alkaline disruption of bovine and human papillomaviruses. *Journal of virology* **21**:1205-1209.
294. **Lambert PF, Spalholz BA, Howley PM.** 1987. A transcriptional repressor encoded by BPV-1 shares a common carboxy-terminal domain with the E2 transactivator. *Cell* **50**:69-78.
295. **Stubenrauch F, Hummel M, Iftner T, Laimins LA.** 2000. The E8E2C protein, a negative regulator of viral transcription and replication, is required for extrachromosomal maintenance of human papillomavirus type 31 in keratinocytes. *Journal of virology* **74**:1178-1186.
296. **Lace MJ, Anson JR, Thomas GS, Turek LP, Haugen TH.** 2008. The E8--E2 gene product of human papillomavirus type 16 represses early transcription and replication but is dispensable for viral plasmid persistence in keratinocytes. *Journal of virology* **82**:10841-10853.
297. **Han R, Cladel NM, Reed CA, Christensen ND.** 1998. Characterization of transformation function of cottontail rabbit papillomavirus E5 and E8 genes. *Virology* **251**:253-263.
298. **Doorbar J, Quint W, Banks L, Bravo IG, Stoler M, Broker TR, Stanley MA.** 2012. The biology and life-cycle of human papillomaviruses. *Vaccine* **30 Suppl 5**:F55-70.
299. **Zanier K, Charbonnier S, Sidi AO, McEwen AG, Ferrario MG, Poussin-Courmontagne P, Cura V, Brimer N, Babah KO, Ansari T, Muller I, Stote RH, Cavarelli J, Vande Pol S, Trave G.** 2013. Structural basis for hijacking of cellular LxxLL motifs by papillomavirus E6 oncoproteins. *Science (New York, N.Y.)* **339**:694-698.
300. **Chen JJ, Hong Y, Rustamzadeh E, Baleja JD, Androphy EJ.** 1998. Identification of an alpha helical motif sufficient for association with papillomavirus E6. *The Journal of biological chemistry* **273**:13537-13544.
301. **Park RB, Androphy EJ.** 2002. Genetic analysis of high-risk e6 in episomal maintenance of human papillomavirus genomes in primary human keratinocytes. *Journal of virology* **76**:11359-11364.
302. **Thomas JT, Hubert WG, Ruesch MN, Laimins LA.** 1999. Human papillomavirus type 31 oncoproteins E6 and E7 are required for the maintenance of episomes during the viral life cycle in normal human keratinocytes. *Proceedings of the National Academy of Sciences of the United States of America* **96**:8449-8454.
303. **Storey A, Banks L.** 1993. Human papillomavirus type 16 E6 gene cooperates with EJ-ras to immortalize primary mouse cells. *Oncogene* **8**:919-924.
304. **Pim D, Banks L.** 2010. Interaction of viral oncoproteins with cellular target molecules: infection with high-risk vs low-risk human papillomaviruses. *APMIS : acta pathologica, microbiologica, et immunologica Scandinavica* **118**:471-493.
305. **Munger K, Phelps WC, Bubb V, Howley PM, Schlegel R.** 1989. The E6 and E7 genes of the human papillomavirus type 16 together are necessary and sufficient for transformation of primary human keratinocytes. *Journal of virology* **63**:4417-4421.
306. **Hawley-Nelson P, Vousden KH, Hubbert NL, Lowy DR, Schiller JT.** 1989. HPV16 E6 and E7 proteins cooperate to immortalize human foreskin keratinocytes. *The EMBO journal* **8**:3905-3910.
307. **Moody CA, Laimins LA.** 2010. Human papillomavirus oncoproteins: pathways to transformation. *Nature reviews. Cancer* **10**:550-560.

308. **Zilfou JT, Lowe SW.** 2009. Tumor suppressive functions of p53. *Cold Spring Harbor perspectives in biology* **1**:a001883.
309. **Tomaic V, Pim D, Banks L.** 2009. The stability of the human papillomavirus E6 oncoprotein is E6AP dependent. *Virology* **393**:7-10.
310. **Scheffner M, Huibregtse JM, Vierstra RD, Howley PM.** 1993. The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. *Cell* **75**:495-505.
311. **Cooper B, Schneider S, Bohl J, Jiang Y, Beaudet A, Vande Pol S.** 2003. Requirement of E6AP and the features of human papillomavirus E6 necessary to support degradation of p53. *Virology* **306**:87-99.
312. **Lechner MS, Laimins LA.** 1994. Inhibition of p53 DNA binding by human papillomavirus E6 proteins. *Journal of virology* **68**:4262-4273.
313. **Patel D, Huang SM, Baglia LA, McCance DJ.** 1999. The E6 protein of human papillomavirus type 16 binds to and inhibits co-activation by CBP and p300. *The EMBO journal* **18**:5061-5072.
314. **Zimmermann H, Degenkolbe R, Bernard HU, O'Connor MJ.** 1999. The human papillomavirus type 16 E6 oncoprotein can down-regulate p53 activity by targeting the transcriptional coactivator CBP/p300. *Journal of virology* **73**:6209-6219.
315. **Heaton PR, Deyrieux AF, Bian XL, Wilson VG.** 2011. HPV E6 proteins target Ubc9, the SUMO conjugating enzyme. *Virus research* **158**:199-208.
316. **Alfandari J, Shnitman Magal S, Jackman A, Schlegel R, Gonen P, Sherman L.** 1999. HPV16 E6 oncoprotein inhibits apoptosis induced during serum-calcium differentiation of foreskin human keratinocytes. *Virology* **257**:383-396.
317. **Magal SS, Jackman A, Ish-Shalom S, Botzer LE, Gonen P, Schlegel R, Sherman L.** 2005. Downregulation of Bax mRNA expression and protein stability by the E6 protein of human papillomavirus 16. *The Journal of general virology* **86**:611-621.
318. **Thomas M, Banks L.** 1998. Inhibition of Bak-induced apoptosis by HPV-18 E6. *Oncogene* **17**:2943-2954.
319. **Thomas M, Banks L.** 1999. Human papillomavirus (HPV) E6 interactions with Bak are conserved amongst E6 proteins from high and low risk HPV types. *The Journal of general virology* **80 ( Pt 6)**:1513-1517.
320. **Underbrink MP, Howie HL, Bedard KM, Koop JI, Galloway DA.** 2008. E6 proteins from multiple human betapapillomavirus types degrade Bak and protect keratinocytes from apoptosis after UVB irradiation. *Journal of virology* **82**:10408-10417.
321. **Du J, Chen GG, Vlantis AC, Chan PK, Tsang RK, van Hasselt CA.** 2004. Resistance to apoptosis of HPV 16-infected laryngeal cancer cells is associated with decreased Bak and increased Bcl-2 expression. *Cancer letters* **205**:81-88.
322. **Yuan CH, Filippova M, Duerksen-Hughes P.** 2012. Modulation of apoptotic pathways by human papillomaviruses (HPV): mechanisms and implications for therapy. *Viruses* **4**:3831-3850.
323. **Filippova M, Parkhurst L, Duerksen-Hughes PJ.** 2004. The human papillomavirus 16 E6 protein binds to Fas-associated death domain and protects cells from Fas-triggered apoptosis. *The Journal of biological chemistry* **279**:25729-25744.

324. **Garnett TO, Filippova M, Duerksen-Hughes PJ.** 2006. Accelerated degradation of FADD and procaspase 8 in cells expressing human papilloma virus 16 E6 impairs TRAIL-mediated apoptosis. *Cell death and differentiation* **13**:1915-1926.
325. **Pim D, Bergant M, Boon SS, Ganti K, Kranjec C, Massimi P, Subbaiah VK, Thomas M, Tomaic V, Banks L.** 2012. Human papillomaviruses and the specificity of PDZ domain targeting. *The FEBS journal* **279**:3530-3537.
326. **Ganti K, Broniarczyk J, Manoubi W, Massimi P, Mittal S, Pim D, Szalmas A, Thatte J, Thomas M, Tomaic V, Banks L.** 2015. The Human Papillomavirus E6 PDZ Binding Motif: From Life Cycle to Malignancy. *Viruses* **7**:3530-3551.
327. **Gardiol D, Kuhne C, Glaunsinger B, Lee SS, Javier R, Banks L.** 1999. Oncogenic human papillomavirus E6 proteins target the discs large tumour suppressor for proteasome-mediated degradation. *Oncogene* **18**:5487-5496.
328. **Woods DF, Hough C, Peel D, Callaini G, Bryant PJ.** 1996. Dlg protein is required for junction structure, cell polarity, and proliferation control in *Drosophila* epithelia. *The Journal of cell biology* **134**:1469-1482.
329. **Humbert P, Russell S, Richardson H.** 2003. Dlg, Scribble and Lgl in cell polarity, cell proliferation and cancer. *BioEssays : news and reviews in molecular, cellular and developmental biology* **25**:542-553.
330. **Grm HS, Banks L.** 2004. Degradation of hDlg and MAGIs by human papillomavirus E6 is E6-AP-independent. *The Journal of general virology* **85**:2815-2819.
331. **Lee C, Laimins LA.** 2004. Role of the PDZ domain-binding motif of the oncoprotein E6 in the pathogenesis of human papillomavirus type 31. *Journal of virology* **78**:12366-12377.
332. **Klingelutz AJ, Foster SA, McDougall JK.** 1996. Telomerase activation by the E6 gene product of human papillomavirus type 16. *Nature* **380**:79-82.
333. **Kelley ML, Keiger KE, Lee CJ, Huibregtse JM.** 2005. The global transcriptional effects of the human papillomavirus E6 protein in cervical carcinoma cell lines are mediated by the E6AP ubiquitin ligase. *Journal of virology* **79**:3737-3747.
334. **Gewin L, Galloway DA.** 2001. E box-dependent activation of telomerase by human papillomavirus type 16 E6 does not require induction of c-myc. *Journal of virology* **75**:7198-7201.
335. **Gewin L, Myers H, Kiyono T, Galloway DA.** 2004. Identification of a novel telomerase repressor that interacts with the human papillomavirus type-16 E6/E6-AP complex. *Genes & development* **18**:2269-2282.
336. **Nomine Y, Masson M, Charbonnier S, Zanier K, Ristriani T, Deryckere F, Sibler AP, Desplancq D, Atkinson RA, Weiss E, Orfanoudakis G, Kieffer B, Trave G.** 2006. Structural and functional analysis of E6 oncoprotein: insights in the molecular pathways of human papillomavirus-mediated pathogenesis. *Molecular cell* **21**:665-678.
337. **Pim D, Tomaic 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. *Journal of virology* **83**:9863-9874.
338. **Clements A, Johnston K, Mazzarelli JM, Ricciardi RP, Marmorstein R.** 2000. Oligomerization properties of the viral oncoproteins adenovirus E1A and human papillomavirus E7 and their complexes with the retinoblastoma protein. *Biochemistry* **39**:16033-16045.

339. **Ohlenschlager O, Seiboth T, Zengerling H, Briese L, Marchanka A, Ramachandran R, Baum M, Korbass M, Meyer-Klaucke W, Durst M, Gorlach M.** 2006. Solution structure of the partially folded high-risk human papilloma virus 45 oncoprotein E7. *Oncogene* **25**:5953-5959.
340. **Firzlaff JM, Galloway DA, Eisenman RN, Lüscher B.** 1989. The E7 protein of human papillomavirus type 16 is phosphorylated by casein kinase II. *The New biologist* **1**:44-53.
341. **Wang J, Sampath A, Raychaudhuri P, Bagchi S.** 2001. Both Rb and E7 are regulated by the ubiquitin proteasome pathway in HPV-containing cervical tumor cells. *Oncogene* **20**:4740-4749.
342. **Phelps WC, Yee CL, Munger K, Howley PM.** 1988. The human papillomavirus type 16 E7 gene encodes transactivation and transformation functions similar to those of adenovirus E1A. *Cell* **53**:539-547.
343. **Davies R, Hicks R, Crook T, Morris J, Vousden K.** 1993. Human papillomavirus type 16 E7 associates with a histone H1 kinase and with p107 through sequences necessary for transformation. *Journal of virology* **67**:2521-2528.
344. **Cobrinik D.** 2005. Pocket proteins and cell cycle control. *Oncogene* **24**:2796-2809.
345. **Dimova DK, Dyson NJ.** 2005. The E2F transcriptional network: old acquaintances with new faces. *Oncogene* **24**:2810-2826.
346. **McLaughlin-Drubin ME, Munger K.** 2009. The human papillomavirus E7 oncoprotein. *Virology* **384**:335-344.
347. **Huh K, Zhou X, Hayakawa H, Cho JY, Libermann TA, Jin J, Harper JW, Munger K.** 2007. Human papillomavirus type 16 E7 oncoprotein associates with the cullin 2 ubiquitin ligase complex, which contributes to degradation of the retinoblastoma tumor suppressor. *Journal of virology* **81**:9737-9747.
348. **Hwang SG, Lee D, Kim J, Seo T, Choe J.** 2002. Human papillomavirus type 16 E7 binds to E2F1 and activates E2F1-driven transcription in a retinoblastoma protein-independent manner. *The Journal of biological chemistry* **277**:2923-2930.
349. **McLaughlin-Drubin ME, Huh KW, Munger K.** 2008. Human papillomavirus type 16 E7 oncoprotein associates with E2F6. *Journal of virology* **82**:8695-8705.
350. **Cheng S, Schmidt-Grimminger DC, Murrant T, Broker TR, Chow LT.** 1995. Differentiation-dependent up-regulation of the human papillomavirus E7 gene reactivates cellular DNA replication in suprabasal differentiated keratinocytes. *Genes & development* **9**:2335-2349.
351. **Longworth MS, Wilson R, Laimins LA.** 2005. HPV31 E7 facilitates replication by activating E2F2 transcription through its interaction with HDACs. *The EMBO journal* **24**:1821-1830.
352. **Funk JO, Waga S, Harry JB, Espling E, Stillman B, Galloway DA.** 1997. Inhibition of CDK activity and PCNA-dependent DNA replication by p21 is blocked by interaction with the HPV-16 E7 oncoprotein. *Genes & development* **11**:2090-2100.
353. **Zerfass-Thome K, Zwerschke W, Mannhardt B, Tindle R, Botz JW, Jansen-Durr P.** 1996. Inactivation of the cdk inhibitor p27KIP1 by the human papillomavirus type 16 E7 oncoprotein. *Oncogene* **13**:2323-2330.
354. **Abbas T, Dutta A.** 2009. p21 in cancer: intricate networks and multiple activities. *Nature reviews. Cancer* **9**:400-414.



355. **Chu IM, Hengst L, Slingerland JM.** 2008. The Cdk inhibitor p27 in human cancer: prognostic potential and relevance to anticancer therapy. *Nature reviews. Cancer* **8**:253-267.
356. **Katich SC, Zerfass-Thome K, Hoffmann I.** 2001. Regulation of the Cdc25A gene by the human papillomavirus Type 16 E7 oncogene. *Oncogene* **20**:543-550.
357. **Duensing S, Duensing A, Crum CP, Munger K.** 2001. Human papillomavirus type 16 E7 oncoprotein-induced abnormal centrosome synthesis is an early event in the evolving malignant phenotype. *Cancer research* **61**:2356-2360.
358. **Duensing S, Duensing A, Flores ER, Do A, Lambert PF, Munger K.** 2001. Centrosome abnormalities and genomic instability by episomal expression of human papillomavirus type 16 in raft cultures of human keratinocytes. *Journal of virology* **75**:7712-7716.
359. **Duensing S, Lee LY, Duensing A, Basile J, Piboonniyom S, Gonzalez S, Crum CP, Munger K.** 2000. The human papillomavirus type 16 E6 and E7 oncoproteins cooperate to induce mitotic defects and genomic instability by uncoupling centrosome duplication from the cell division cycle. *Proceedings of the National Academy of Sciences of the United States of America* **97**:10002-10007.
360. **Duensing A, Liu Y, Tseng M, Malumbres M, Barbacid M, Duensing S.** 2006. Cyclin-dependent kinase 2 is dispensable for normal centrosome duplication but required for oncogene-induced centrosome overduplication. *Oncogene* **25**:2943-2949.
361. **Nguyen CL, Eichwald C, Nibert ML, Munger K.** 2007. Human papillomavirus type 16 E7 oncoprotein associates with the centrosomal component gamma-tubulin. *Journal of virology* **81**:13533-13543.
362. **Moody CA, Laimins LA.** 2009. Human papillomaviruses activate the ATM DNA damage pathway for viral genome amplification upon differentiation. *PLoS pathogens* **5**:e1000605.
363. **Jones DL, Munger K.** 1997. Analysis of the p53-mediated G1 growth arrest pathway in cells expressing the human papillomavirus type 16 E7 oncoprotein. *Journal of virology* **71**:2905-2912.
364. **Sun Y, Han H, McCance DJ.** 1998. Active domains of human papillomavirus type 11 E1 protein for origin replication. *The Journal of general virology* **79 ( Pt 7)**:1651-1658.
365. **Ferran MC, McBride AA.** 1998. Transient viral DNA replication and repression of viral transcription are supported by the C-terminal domain of the bovine papillomavirus type 1 E1 protein. *Journal of virology* **72**:796-801.
366. **Yang L, Mohr I, Fouts E, Lim DA, Nohaile M, Botchan M.** 1993. The E1 protein of bovine papilloma virus 1 is an ATP-dependent DNA helicase. *Proceedings of the National Academy of Sciences of the United States of America* **90**:5086-5090.
367. **Raj K, Stanley MA.** 1995. The ATP-binding and ATPase activities of human papillomavirus type 16 E1 are significantly weakened by the absence of prolines in its ATP-binding domain. *The Journal of general virology* **76 ( Pt 12)**:2949-2956.
368. **Sedman J, Stenlund A.** 1998. The papillomavirus E1 protein forms a DNA-dependent hexameric complex with ATPase and DNA helicase activities. *Journal of virology* **72**:6893-6897.
369. **Mechali F, Hsu CY, Castro A, Lorca T, Bonne-Andrea C.** 2004. Bovine papillomavirus replicative helicase E1 is a target of the ubiquitin ligase APC. *Journal of virology* **78**:2615-2619.

370. **Fradet-Turcotte A, Brault K, Titolo S, Howley PM, Archambault J.** 2009. Characterization of papillomavirus E1 helicase mutants defective for interaction with the SUMO-conjugating enzyme Ubc9. *Virology* **395**:190-201.
371. **Rangasamy D, Wilson VG.** 2000. Bovine papillomavirus E1 protein is sumoylated by the host cell Ubc9 protein. *The Journal of biological chemistry* **275**:30487-30495.
372. **Fouts ET, Yu X, Egelman EH, Botchan MR.** 1999. Biochemical and electron microscopic image analysis of the hexameric E1 helicase. *The Journal of biological chemistry* **274**:4447-4458.
373. **Titolo S, Pelletier A, Pulichino AM, Brault K, Wardrop E, White PW, Cordingley MG, Archambault J.** 2000. Identification of domains of the human papillomavirus type 11 E1 helicase involved in oligomerization and binding to the viral origin. *Journal of virology* **74**:7349-7361.
374. **White PW, Pelletier A, Brault K, Titolo S, Welchner E, Thauvette L, Fazekas M, Cordingley MG, Archambault J.** 2001. Characterization of recombinant HPV6 and 11 E1 helicases: effect of ATP on the interaction of E1 with E2 and mapping of a minimal helicase domain. *The Journal of biological chemistry* **276**:22426-22438.
375. **Whelan F, Stead JA, Shkumatov AV, Svergun DI, Sanders CM, Antson AA.** 2012. A flexible brace maintains the assembly of a hexameric replicative helicase during DNA unwinding. *Nucleic acids research* **40**:2271-2283.
376. **Lusky M, Fontane E.** 1991. Formation of the complex of bovine papillomavirus E1 and E2 proteins is modulated by E2 phosphorylation and depends upon sequences within the carboxyl terminus of E1. *Proceedings of the National Academy of Sciences of the United States of America* **88**:6363-6367.
377. **Chen G, Stenlund A.** 1998. Characterization of the DNA-binding domain of the bovine papillomavirus replication initiator E1. *Journal of virology* **72**:2567-2576.
378. **Dixon EP, Pahel GL, Rocque WJ, Barnes JA, Lobe DC, Hanlon MH, Alexander KA, Chao SF, Lindley K, Phelps WC.** 2000. The E1 helicase of human papillomavirus type 11 binds to the origin of replication with low sequence specificity. *Virology* **270**:345-357.
379. **Conti E, Uy M, Leighton L, Blobel G, Kuriyan J.** 1998. Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha. *Cell* **94**:193-204.
380. **Bian XL, Rosas-Acosta G, Wu YC, Wilson VG.** 2007. Nuclear import of bovine papillomavirus type 1 E1 protein is mediated by multiple alpha importins and is negatively regulated by phosphorylation near a nuclear localization signal. *Journal of virology* **81**:2899-2908.
381. **Lentz MR, Pak D, Mohr I, Botchan MR.** 1993. The E1 replication protein of bovine papillomavirus type 1 contains an extended nuclear localization signal that includes a p34cdc2 phosphorylation site. *Journal of virology* **67**:1414-1423.
382. **Zanardi TA, Stanley CM, Saviile BM, Spacek SM, Lentz MR.** 1997. Modulation of bovine papillomavirus DNA replication by phosphorylation of the viral E1 protein. *Virology* **228**:1-10.
383. **Yu JH, Lin BY, Deng W, Broker TR, Chow LT.** 2007. Mitogen-activated protein kinases activate the nuclear localization sequence of human papillomavirus type 11 E1 DNA helicase to promote efficient nuclear import. *Journal of virology* **81**:5066-5078.

384. **Deng W, Lin BY, Jin G, Wheeler CG, Ma T, Harper JW, Broker TR, Chow LT.** 2004. Cyclin/CDK regulates the nucleocytoplasmic localization of the human papillomavirus E1 DNA helicase. *Journal of virology* **78**:13954-13965.
385. **Fradet-Turcotte A, Moody C, Laimins LA, Archambault J.** 2010. Nuclear export of human papillomavirus type 31 E1 is regulated by Cdk2 phosphorylation and required for viral genome maintenance. *Journal of virology* **84**:11747-11760.
386. **Cueille N, Nougarede R, Mechali F, Philippe M, Bonne-Andrea C.** 1998. Functional interaction between the bovine papillomavirus virus type 1 replicative helicase E1 and cyclin E-Cdk2. *Journal of virology* **72**:7255-7262.
387. **Ma T, Zou N, Lin BY, Chow LT, Harper JW.** 1999. Interaction between cyclin-dependent kinases and human papillomavirus replication-initiation protein E1 is required for efficient viral replication. *Proceedings of the National Academy of Sciences of the United States of America* **96**:382-387.
388. **Dong X, Biswas A, Suel KE, Jackson LK, Martinez R, Gu H, Chook YM.** 2009. Structural basis for leucine-rich nuclear export signal recognition by CRM1. *Nature* **458**:1136-1141.
389. **Hsu CY, Mechali F, Bonne-Andrea C.** 2007. Nucleocytoplasmic shuttling of bovine papillomavirus E1 helicase downregulates viral DNA replication in S phase. *Journal of virology* **81**:384-394.
390. **McShan GD, Wilson VG.** 1997. Casein kinase II phosphorylates bovine papillomavirus type 1 E1 in vitro at a conserved motif. *The Journal of general virology* **78 ( Pt 1)**:171-177.
391. **Lentz MR, Stevens SM, Jr., Raynes J, Elkhoury N.** 2006. A phosphorylation map of the bovine papillomavirus E1 helicase. *Virology journal* **3**:13.
392. **Fradet-Turcotte A, Bergeron-Labrecque F, Moody CA, Lehoux M, Laimins LA, Archambault J.** 2011. Nuclear accumulation of the papillomavirus E1 helicase blocks S-phase progression and triggers an ATM-dependent DNA damage response. *Journal of virology* **85**:8996-9012.
393. **Reinson T, Toots M, Kadaja M, Pipitch R, Allik M, Ustav E, Ustav M.** 2013. Engagement of the ATR-dependent DNA damage response at the human papillomavirus 18 replication centers during the initial amplification. *Journal of virology* **87**:951-964.
394. **Gillespie KA, Mehta KP, Laimins LA, Moody CA.** 2012. Human papillomaviruses recruit cellular DNA repair and homologous recombination factors to viral replication centers. *Journal of virology* **86**:9520-9526.
395. **Park P, Copeland W, Yang L, Wang T, Botchan MR, Mohr IJ.** 1994. The cellular DNA polymerase alpha-primase is required for papillomavirus DNA replication and associates with the viral E1 helicase. *Proceedings of the National Academy of Sciences of the United States of America* **91**:8700-8704.
396. **Bonne-Andrea C, Santucci S, Clertant P, Tillier F.** 1995. Bovine papillomavirus E1 protein binds specifically DNA polymerase alpha but not replication protein A. *Journal of virology* **69**:2341-2350.
397. **Loo YM, Melendy T.** 2004. Recruitment of replication protein A by the papillomavirus E1 protein and modulation by single-stranded DNA. *Journal of virology* **78**:1605-1615.
398. **Swindle CS, Engler JA.** 1998. Association of the human papillomavirus type 11 E1 protein with histone H1. *Journal of virology* **72**:1994-2001.

399. **Lee D, Sohn H, Kalpana GV, Choe J.** 1999. Interaction of E1 and hSNF5 proteins stimulates replication of human papillomavirus DNA. *Nature* **399**:487-491.
400. **Liu JS, Kuo SR, Makhov AM, Cyr DM, Griffith JD, Broker TR, Chow LT.** 1998. Human Hsp70 and Hsp40 chaperone proteins facilitate human papillomavirus-11 E1 protein binding to the origin and stimulate cell-free DNA replication. *The Journal of biological chemistry* **273**:30704-30712.
401. **Moody CA, Fradet-Turcotte A, Archambault J, Laimins LA.** 2007. Human papillomaviruses activate caspases upon epithelial differentiation to induce viral genome amplification. *Proceedings of the National Academy of Sciences of the United States of America* **104**:19541-19546.
402. **Johansson C, Graham SV, Dornan ES, Morgan IM.** 2009. The human papillomavirus 16 E2 protein is stabilised in S phase. *Virology* **394**:194-199.
403. **Sekhar V, McBride AA.** 2012. Phosphorylation regulates binding of the human papillomavirus type 8 E2 protein to host chromosomes. *Journal of virology* **86**:10047-10058.
404. **Quinlan EJ, Culleton SP, Wu SY, Chiang CM, Androphy EJ.** 2013. Acetylation of conserved lysines in bovine papillomavirus E2 by p300. *Journal of virology* **87**:1497-1507.
405. **Wu YC, Roark AA, Bian XL, Wilson VG.** 2008. Modification of papillomavirus E2 proteins by the small ubiquitin-like modifier family members (SUMOs). *Virology* **378**:329-338.
406. **Gagnon D, Joubert S, Senechal H, Fradet-Turcotte A, Torre S, Archambault J.** 2009. Proteasomal degradation of the papillomavirus E2 protein is inhibited by overexpression of bromodomain-containing protein 4. *Journal of virology* **83**:4127-4139.
407. **Lee AY, Chiang CM.** 2009. Chromatin adaptor Brd4 modulates E2 transcription activity and protein stability. *The Journal of biological chemistry* **284**:2778-2786.
408. **King LE, Dornan ES, Donaldson MM, Morgan IM.** 2011. Human papillomavirus 16 E2 stability and transcriptional activation is enhanced by E1 via a direct protein-protein interaction. *Virology* **414**:26-33.
409. **O'Connor MJ, Chan SY, Bernard HU.** 1995. Transcription factor binding sites in the long control regions of genital HPVs, part III-A, p. pp. 21-40. *In* Myers G BHU, Delius H, Baker C, Icenogle J, Halpern A, Wheeler C (ed.), *Human papillomaviruses 1995 compendium*. Los Alamos National Laboratory, Los Alamos, N.M. USA.
410. **Romanczuk H, Thierry F, Howley PM.** 1990. Mutational analysis of cis elements involved in E2 modulation of human papillomavirus type 16 P97 and type 18 P105 promoters. *Journal of virology* **64**:2849-2859.
411. **Steger G, Corbach S.** 1997. Dose-dependent regulation of the early promoter of human papillomavirus type 18 by the viral E2 protein. *Journal of virology* **71**:50-58.
412. **Lim DA, Gossen M, Lehman CW, Botchan MR.** 1998. Competition for DNA binding sites between the short and long forms of E2 dimers underlies repression in bovine papillomavirus type 1 DNA replication control. *Journal of virology* **72**:1931-1940.
413. **Grm HS, Massimi P, Gammoh N, Banks L.** 2005. Crosstalk between the human papillomavirus E2 transcriptional activator and the E6 oncoprotein. *Oncogene* **24**:5149-5164.

414. **Gammoh N, Isaacson E, Tomaic V, Jackson DJ, Doorbar J, Banks L.** 2009. Inhibition of HPV-16 E7 oncogenic activity by HPV-16 E2. *Oncogene* **28**:2299-2304.
415. **Baxter MK, McPhillips MG, Ozato K, McBride AA.** 2005. The mitotic chromosome binding activity of the papillomavirus E2 protein correlates with interaction with the cellular chromosomal protein, Brd4. *Journal of virology* **79**:4806-4818.
416. **Senechal H, Poirier GG, Coulombe B, Laimins LA, Archambault J.** 2007. Amino acid substitutions that specifically impair the transcriptional activity of papillomavirus E2 affect binding to the long isoform of Brd4. *Virology* **358**:10-17.
417. **Wang X, Helfer CM, Pancholi N, Bradner JE, You J.** 2013. Recruitment of Brd4 to the human papillomavirus type 16 DNA replication complex is essential for replication of viral DNA. *Journal of virology* **87**:3871-3884.
418. **Poddar A, Reed SC, McPhillips MG, Spindler JE, McBride AA.** 2009. The human papillomavirus type 8 E2 tethering protein targets the ribosomal DNA loci of host mitotic chromosomes. *Journal of virology* **83**:640-650.
419. **Oliveira JG, Colf LA, McBride AA.** 2006. Variations in the association of papillomavirus E2 proteins with mitotic chromosomes. *Proceedings of the National Academy of Sciences of the United States of America* **103**:1047-1052.
420. **Zhao KN, Hengst K, Liu WJ, Liu YH, Liu XS, McMillan NA, Frazer IH.** 2000. BPV1 E2 protein enhances packaging of full-length plasmid DNA in BPV1 pseudovirions. *Virology* **272**:382-393.
421. **Webster K, Parish J, Pandya M, Stern PL, Clarke AR, Gaston K.** 2000. The human papillomavirus (HPV) 16 E2 protein induces apoptosis in the absence of other HPV proteins and via a p53-dependent pathway. *The Journal of biological chemistry* **275**:87-94.
422. **Thierry F, Demeret C.** 2008. Direct activation of caspase 8 by the proapoptotic E2 protein of HPV18 independent of adaptor proteins. *Cell death and differentiation* **15**:1356-1363.
423. **Chiang CM, Broker TR, Chow LT.** 1991. An E1M--E2C fusion protein encoded by human papillomavirus type 11 is a sequence-specific transcription repressor. *Journal of virology* **65**:3317-3329.
424. **Alloul N, Sherman L.** 1999. Transcription-modulatory activities of differentially spliced cDNAs encoding the E2 protein of human papillomavirus type 16. *The Journal of general virology* **80 ( Pt 9)**:2461-2470.
425. **Hubbert NL, Schiller JT, Lowy DR, Androphy EJ.** 1988. Bovine papilloma virus-transformed cells contain multiple E2 proteins. *Proceedings of the National Academy of Sciences of the United States of America* **85**:5864-5868.
426. **Barsoum J, Prakash SS, Han P, Androphy EJ.** 1992. Mechanism of action of the papillomavirus E2 repressor: repression in the absence of DNA binding. *Journal of virology* **66**:3941-3945.
427. **Doorbar J, Campbell D, Grand RJ, Gallimore PH.** 1986. Identification of the human papilloma virus-1a E4 gene products. *The EMBO journal* **5**:355-362.
428. **Milligan SG, Veerapraditsin T, Ahamet B, Mole S, Graham SV.** 2007. Analysis of novel human papillomavirus type 16 late mRNAs in differentiated W12 cervical epithelial cells. *Virology* **360**:172-181.
429. **Khan J, Davy CE, McIntosh PB, Jackson DJ, Hinz S, Wang Q, Doorbar J.** 2011. Role of calpain in the formation of human papillomavirus type 16 E1<sup>E4</sup>

- amyloid fibers and reorganization of the keratin network. *Journal of virology* **85**:9984-9997.
430. **Tan CL, Gunaratne J, Lai D, Carthage L, Wang Q, Xue YZ, Quek LS, Doorbar J, Bachelier F, Thierry F, Bellanger S.** 2012. HPV-18 E2<sup>E4</sup> chimera: 2 new spliced transcripts and proteins induced by keratinocyte differentiation. *Virology* **429**:47-56.
431. **Nakahara T, Peh WL, Doorbar J, Lee D, Lambert PF.** 2005. Human papillomavirus type 16 E1<sup>circumflex</sup>E4 contributes to multiple facets of the papillomavirus life cycle. *Journal of virology* **79**:13150-13165.
432. **Roberts S, Ashmole I, Gibson LJ, Rookes SM, Barton GJ, Gallimore PH.** 1994. Mutational analysis of human papillomavirus E4 proteins: identification of structural features important in the formation of cytoplasmic E4/cytokeratin networks in epithelial cells. *Journal of virology* **68**:6432-6445.
433. **Knight GL, Grainger JR, Gallimore PH, Roberts S.** 2004. Cooperation between different forms of the human papillomavirus type 1 E4 protein to block cell cycle progression and cellular DNA synthesis. *Journal of virology* **78**:13920-13933.
434. **Peh WL, Brandsma JL, Christensen ND, Cladel NM, Wu X, Doorbar J.** 2004. The viral E4 protein is required for the completion of the cottontail rabbit papillomavirus productive cycle in vivo. *Journal of virology* **78**:2142-2151.
435. **Wilson R, Fehrmann F, Laimins LA.** 2005. Role of the E1--E4 protein in the differentiation-dependent life cycle of human papillomavirus type 31. *Journal of virology* **79**:6732-6740.
436. **Doorbar J, Ely S, Sterling J, McLean C, Crawford L.** 1991. Specific interaction between HPV-16 E1-E4 and cytokeratins results in collapse of the epithelial cell intermediate filament network. *Nature* **352**:824-827.
437. **Roberts S, Ashmole I, Johnson GD, Kreider JW, Gallimore PH.** 1993. Cutaneous and mucosal human papillomavirus E4 proteins form intermediate filament-like structures in epithelial cells. *Virology* **197**:176-187.
438. **McIntosh PB, Martin SR, Jackson DJ, Khan J, Isaacson ER, Calder L, Raj K, Griffin HM, Wang Q, Laskey P, Eccleston JF, Doorbar J.** 2008. Structural analysis reveals an amyloid form of the human papillomavirus type 16 E1--E4 protein and provides a molecular basis for its accumulation. *Journal of virology* **82**:8196-8203.
439. **McIntosh PB, Laskey P, Sullivan K, Davy C, Wang Q, Jackson DJ, Griffin HM, Doorbar J.** 2010. E1--E4-mediated keratin phosphorylation and ubiquitylation: a mechanism for keratin depletion in HPV16-infected epithelium. *Journal of cell science* **123**:2810-2822.
440. **Wang Q, Kennedy A, Das P, McIntosh PB, Howell SA, Isaacson ER, Hinz SA, Davy C, Doorbar J.** 2009. Phosphorylation of the human papillomavirus type 16 E1--E4 protein at T57 by ERK triggers a structural change that enhances keratin binding and protein stability. *Journal of virology* **83**:3668-3683.
441. **Davy CE, Jackson DJ, Wang Q, Raj K, Masterson PJ, Fenner NF, Southern S, Cuthill S, Millar JB, Doorbar J.** 2002. Identification of a G(2) arrest domain in the E1 wedge E4 protein of human papillomavirus type 16. *Journal of virology* **76**:9806-9818.
442. **Davy CE, Jackson DJ, Raj K, Peh WL, Southern SA, Das P, Sorathia R, Laskey P, Middleton K, Nakahara T, Wang Q, Masterson PJ, Lambert PF, Cuthill S, Millar JB, Doorbar J.** 2005. Human papillomavirus type 16 E1 E4-

- induced G2 arrest is associated with cytoplasmic retention of active Cdk1/cyclin B1 complexes. *Journal of virology* **79**:3998-4011.
443. **Knight GL, Pugh AG, Yates E, Bell I, Wilson R, Moody CA, Laimins LA, Roberts S.** 2011. A cyclin-binding motif in human papillomavirus type 18 (HPV18) E1<sup>E4</sup> is necessary for association with CDK-cyclin complexes and G2/M cell cycle arrest of keratinocytes, but is not required for differentiation-dependent viral genome amplification or L1 capsid protein expression. *Virology* **412**:196-210.
444. **Raj K, Berguerand S, Southern S, Doorbar J, Beard P.** 2004. E1 empty set E4 protein of human papillomavirus type 16 associates with mitochondria. *Journal of virology* **78**:7199-7207.
445. **DiMaio D, Petti LM.** 2013. The E5 proteins. *Virology* **445**:99-114.
446. **Burnett S, Jareborg N, DiMaio D.** 1992. Localization of bovine papillomavirus type 1 E5 protein to transformed basal keratinocytes and permissive differentiated cells in fibropapilloma tissue. *Proceedings of the National Academy of Sciences of the United States of America* **89**:5665-5669.
447. **Bergman P, Ustav M, Sedman J, Moreno-Lopez J, Vennstrom B, Pettersson U.** 1988. The E5 gene of bovine papillomavirus type 1 is sufficient for complete oncogenic transformation of mouse fibroblasts. *Oncogene* **2**:453-459.
448. **DiMaio D, Guralski D, Schiller JT.** 1986. Translation of open reading frame E5 of bovine papillomavirus is required for its transforming activity. *Proceedings of the National Academy of Sciences of the United States of America* **83**:1797-1801.
449. **Petti LM, Ray FA.** 2000. Transformation of mortal human fibroblasts and activation of a growth inhibitory pathway by the bovine papillomavirus E5 oncoprotein. *Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research* **11**:395-408.
450. **Schiller JT, Vass WC, Vousden KH, Lowy DR.** 1986. E5 open reading frame of bovine papillomavirus type 1 encodes a transforming gene. *Journal of virology* **57**:1-6.
451. **Drummond-Barbosa DA, Vaillancourt RR, Kazlauskas A, DiMaio D.** 1995. Ligand-independent activation of the platelet-derived growth factor beta receptor: requirements for bovine papillomavirus E5-induced mitogenic signaling. *Molecular and cellular biology* **15**:2570-2581.
452. **Lai CC, Edwards AP, DiMaio D.** 2005. Productive interaction between transmembrane mutants of the bovine papillomavirus E5 protein and the platelet-derived growth factor beta receptor. *Journal of virology* **79**:1924-1929.
453. **McBride AA, Dlugosz A, Baker CC.** 2000. Production of infectious bovine papillomavirus from cloned viral DNA by using an organotypic raft/xenograft technique. *Proceedings of the National Academy of Sciences of the United States of America* **97**:5534-5539.
454. **Straight SW, Hinkle PM, Jewers RJ, McCance DJ.** 1993. The E5 oncoprotein of human papillomavirus type 16 transforms fibroblasts and effects the downregulation of the epidermal growth factor receptor in keratinocytes. *Journal of virology* **67**:4521-4532.
455. **Leechanachai P, Banks L, Moreau F, Matlashewski G.** 1992. The E5 gene from human papillomavirus type 16 is an oncogene which enhances growth factor-mediated signal transduction to the nucleus. *Oncogene* **7**:19-25.

456. **Pim D, Collins M, Banks L.** 1992. Human papillomavirus type 16 E5 gene stimulates the transforming activity of the epidermal growth factor receptor. *Oncogene* **7**:27-32.
457. **Bouvard V, Matlashewski G, Gu ZM, Storey A, Banks L.** 1994. The human papillomavirus type 16 E5 gene cooperates with the E7 gene to stimulate proliferation of primary cells and increases viral gene expression. *Virology* **203**:73-80.
458. **Stoppler MC, Straight SW, Tsao G, Schlegel R, McCance DJ.** 1996. The E5 gene of HPV-16 enhances keratinocyte immortalization by full-length DNA. *Virology* **223**:251-254.
459. **Maufort JP, Shai A, Pitot HC, Lambert PF.** 2010. A role for HPV16 E5 in cervical carcinogenesis. *Cancer research* **70**:2924-2931.
460. **Cortese MS, Ashrafi GH, Campo MS.** 2010. All 4 di-leucine motifs in the first hydrophobic domain of the E5 oncoprotein of human papillomavirus type 16 are essential for surface MHC class I downregulation activity and E5 endomembrane localization. *International journal of cancer. Journal international du cancer* **126**:1675-1682.
461. **Miura S, Kawana K, Schust DJ, Fujii T, Yokoyama T, Iwasawa Y, Nagamatsu T, Adachi K, Tomio A, Tomio K, Kojima S, Yasugi T, Kozuma S, Taketani Y.** 2010. CD1d, a sentinel molecule bridging innate and adaptive immunity, is downregulated by the human papillomavirus (HPV) E5 protein: a possible mechanism for immune evasion by HPV. *Journal of virology* **84**:11614-11623.
462. **Wang X, Meyers C, Wang HK, Chow LT, Zheng ZM.** 2011. Construction of a full transcription map of human papillomavirus type 18 during productive viral infection. *Journal of virology* **85**:8080-8092.
463. **Fertey J, Hurst J, Straub E, Schenker A, Iftner T, Stubenrauch F.** 2011. Growth inhibition of HeLa cells is a conserved feature of high-risk human papillomavirus E8<sup>E2C</sup> proteins and can also be achieved by an artificial repressor protein. *Journal of virology* **85**:2918-2926.
464. **Stubenrauch F, Zobel T, Iftner T.** 2001. The E8 domain confers a novel long-distance transcriptional repression activity on the E8E2C protein of high-risk human papillomavirus type 31. *Journal of virology* **75**:4139-4149.
465. **Ammermann I, Bruckner M, Matthes F, Iftner T, Stubenrauch F.** 2008. Inhibition of transcription and DNA replication by the papillomavirus E8-E2C protein is mediated by interaction with corepressor molecules. *Journal of virology* **82**:5127-5136.
466. **Powell ML, Smith JA, Sowa ME, Harper JW, Iftner T, Stubenrauch F, Howley PM.** 2010. NCoR1 mediates papillomavirus E8;E2C transcriptional repression. *Journal of virology* **84**:4451-4460.
467. **Kurg R, Uusen P, Vosa L, Ustav M.** 2010. Human papillomavirus E2 protein with single activation domain initiates HPV18 genome replication, but is not sufficient for long-term maintenance of virus genome. *Virology* **408**:159-166.
468. **Zobel T, Iftner T, Stubenrauch F.** 2003. The papillomavirus E8-E2C protein represses DNA replication from extrachromosomal origins. *Molecular and cellular biology* **23**:8352-8362.
469. **Straub E, Dreer M, Fertey J, Iftner T, Stubenrauch F.** 2014. The viral E8<sup>E2C</sup> repressor limits productive replication of human papillomavirus 16. *Journal of virology* **88**:937-947.



470. **Buck CB, Day PM, Trus BL.** 2013. The papillomavirus major capsid protein L1. *Virology* **445**:169-174.
471. **Hagensee ME, Yaegashi N, Galloway DA.** 1993. Self-assembly of human papillomavirus type 1 capsids by expression of the L1 protein alone or by coexpression of the L1 and L2 capsid proteins. *Journal of virology* **67**:315-322.
472. **Kirnbauer R, Booy F, Cheng N, Lowy DR, Schiller JT.** 1992. Papillomavirus L1 major capsid protein self-assembles into virus-like particles that are highly immunogenic. *Proceedings of the National Academy of Sciences of the United States of America* **89**:12180-12184.
473. **Kirnbauer R, Taub J, Greenstone H, Roden R, Durst M, Gissmann L, Lowy DR, Schiller JT.** 1993. Efficient self-assembly of human papillomavirus type 16 L1 and L1-L2 into virus-like particles. *Journal of virology* **67**:6929-6936.
474. **Christensen ND, Cladel NM, Reed CA, Budgeon LR, Embers ME, Skulsky DM, McClements WL, Ludmerer SW, Jansen KU.** 2001. Hybrid papillomavirus L1 molecules assemble into virus-like particles that reconstitute conformational epitopes and induce neutralizing antibodies to distinct HPV types. *Virology* **291**:324-334.
475. **Carter JJ, Wipf GC, Benki SF, Christensen ND, Galloway DA.** 2003. Identification of a human papillomavirus type 16-specific epitope on the C-terminal arm of the major capsid protein L1. *Journal of virology* **77**:11625-11632.
476. **Wang Z, Christensen N, Schiller JT, Dillner J.** 1997. A monoclonal antibody against intact human papillomavirus type 16 capsids blocks the serological reactivity of most human sera. *The Journal of general virology* **78 ( Pt 9)**:2209-2215.
477. **Stanley M, Lowy DR, Frazer I.** 2006. Chapter 12: Prophylactic HPV vaccines: underlying mechanisms. *Vaccine* **24 Suppl 3**:S3/106-113.
478. **Buck CB, Cheng N, Thompson CD, Lowy DR, Steven AC, Schiller JT, Trus BL.** 2008. Arrangement of L2 within the papillomavirus capsid. *Journal of virology* **82**:5190-5197.
479. **Heino P, Skyldberg B, Lehtinen M, Rantala I, Hagmar B, Kreider JW, Kirnbauer R, Dillner J.** 1995. Human papillomavirus type 16 capsids expose multiple type-restricted and type-common antigenic epitopes. *The Journal of general virology* **76 ( Pt 5)**:1141-1153.
480. **Day PM, Gambhira R, Roden RB, Lowy DR, Schiller JT.** 2008. Mechanisms of human papillomavirus type 16 neutralization by L2 cross-neutralizing and L1 type-specific antibodies. *Journal of virology* **82**:4638-4646.
481. **Rubio I, Seitz H, Canali E, Sehr P, Bolchi A, Tommasino M, Ottonello S, Muller M.** 2011. The N-terminal region of the human papillomavirus L2 protein contains overlapping binding sites for neutralizing, cross-neutralizing and non-neutralizing antibodies. *Virology* **409**:348-359.
482. **Gambhira R, Karanam B, Jagu S, Roberts JN, Buck CB, Bossis I, Alphs H, Culp T, Christensen ND, Roden RB.** 2007. A protective and broadly cross-neutralizing epitope of human papillomavirus L2. *Journal of virology* **81**:13927-13931.
483. **Campos SK, Ozbun MA.** 2009. Two highly conserved cysteine residues in HPV16 L2 form an intramolecular disulfide bond and are critical for infectivity in human keratinocytes. *PloS one* **4**:e4463.

484. **Heino P, Zhou J, Lambert PF.** 2000. Interaction of the papillomavirus transcription/replication factor, E2, and the viral capsid protein, L2. *Virology* **276**:304-314.
485. **Lin Z, Yemelyanova AV, Gambhira R, Jagu S, Meyers C, Kirnbauer R, Ronnett BM, Gravitt PE, Roden RB.** 2009. Expression pattern and subcellular localization of human papillomavirus minor capsid protein L2. *The American journal of pathology* **174**:136-143.
486. **Lee D, Kim H, Lee Y, Choe J.** 1997. Identification of sequence requirement for the origin of DNA replication in human papillomavirus type 18. *Virus research* **52**:97-108.
487. **Sverdrup F, Khan SA.** 1995. Two E2 binding sites alone are sufficient to function as the minimal origin of replication of human papillomavirus type 18 DNA. *Journal of virology* **69**:1319-1323.
488. **Holt SE, Wilson VG.** 1995. Mutational analysis of the 18-base-pair inverted repeat element at the bovine papillomavirus origin of replication: identification of critical sequences for E1 binding and in vivo replication. *Journal of virology* **69**:6525-6532.
489. **Chen G, Stenlund A.** 2001. The E1 initiator recognizes multiple overlapping sites in the papillomavirus origin of DNA replication. *Journal of virology* **75**:292-302.
490. **Titolo S, Brault K, Majewski J, White PW, Archambault J.** 2003. Characterization of the minimal DNA binding domain of the human papillomavirus e1 helicase: fluorescence anisotropy studies and characterization of a dimerization-defective mutant protein. *Journal of virology* **77**:5178-5191.
491. **Frattini MG, Laimins LA.** 1994. Binding of the human papillomavirus E1 origin-recognition protein is regulated through complex formation with the E2 enhancer-binding protein. *Proceedings of the National Academy of Sciences of the United States of America* **91**:12398-12402.
492. **Lusky M, Hurwitz J, Seo YS.** 1993. Cooperative assembly of the bovine papilloma virus E1 and E2 proteins on the replication origin requires an intact E2 binding site. *The Journal of biological chemistry* **268**:15795-15803.
493. **Lusky M, Hurwitz J, Seo YS.** 1994. The bovine papillomavirus E2 protein modulates the assembly of but is not stably maintained in a replication-competent multimeric E1-replication origin complex. *Proceedings of the National Academy of Sciences of the United States of America* **91**:8895-8899.
494. **Sanders CM, Stenlund A.** 2001. Mechanism and requirements for bovine papillomavirus, type 1, E1 initiator complex assembly promoted by the E2 transcription factor bound to distal sites. *The Journal of biological chemistry* **276**:23689-23699.
495. **Bonne-Andrea C, Tillier F, McShan GD, Wilson VG, Clertant P.** 1997. Bovine papillomavirus type 1 DNA replication: the transcriptional activator E2 acts in vitro as a specificity factor. *Journal of virology* **71**:6805-6815.
496. **Enemark EJ, Chen G, Vaughn DE, Stenlund A, Joshua-Tor L.** 2000. Crystal structure of the DNA binding domain of the replication initiation protein E1 from papillomavirus. *Molecular cell* **6**:149-158.
497. **Sedman J, Stenlund A.** 1996. The initiator protein E1 binds to the bovine papillomavirus origin of replication as a trimeric ring-like structure. *The EMBO journal* **15**:5085-5092.
498. **Schuck S, Stenlund A.** 2005. Assembly of a double hexameric helicase. *Molecular cell* **20**:377-389.

499. **Schuck S, Stenlund A.** 2011. Mechanistic analysis of local ori melting and helicase assembly by the papillomavirus E1 protein. *Molecular cell* **43**:776-787.
500. **Enemark EJ, Joshua-Tor L.** 2006. Mechanism of DNA translocation in a replicative hexameric helicase. *Nature* **442**:270-275.
501. **Sanders CM, Kovalevskiy OV, Sizov D, Lebedev AA, Isupov MN, Antson AA.** 2007. Papillomavirus E1 helicase assembly maintains an asymmetric state in the absence of DNA and nucleotide cofactors. *Nucleic acids research* **35**:6451-6457.
502. **Nallaseth FS, DePamphilis ML.** 1994. Papillomavirus contains cis-acting sequences that can suppress but not regulate origins of DNA replication. *Journal of virology* **68**:3051-3064.
503. **Yates JL, Guan N.** 1991. Epstein-Barr virus-derived plasmids replicate only once per cell cycle and are not amplified after entry into cells. *Journal of virology* **65**:483-488.
504. **Gilbert DM, Cohen SN.** 1987. Bovine papilloma virus plasmids replicate randomly in mouse fibroblasts throughout S phase of the cell cycle. *Cell* **50**:59-68.
505. **Lee TI, Young RA.** 2000. Transcription of eukaryotic protein-coding genes. *Annual review of genetics* **34**:77-137.
506. **Howley P, Lowy D.** 2007. Papillomaviruses, p. 2299–2354. *In* Knipe D, Howley P (ed.), *Fields Virology*, vol. Vol. 5. Lippincott Williams & Wilkins, Philadelphia, PA.
507. **Tan SH, Baker CC, Stunkel W, Bernard HU.** 2003. A transcriptional initiator overlaps with a conserved YY1 binding site in the long control region of human papillomavirus type 16. *Virology* **305**:486-501.
508. **Johansson C, Schwartz S.** 2013. Regulation of human papillomavirus gene expression by splicing and polyadenylation. *Nature reviews. Microbiology* **11**:239-251.
509. **Graham SV.** 2010. Human papillomavirus: gene expression, regulation and prospects for novel diagnostic methods and antiviral therapies. *Future microbiology* **5**:1493-1506.
510. **Jiang C, Pugh BF.** 2009. Nucleosome positioning and gene regulation: advances through genomics. *Nature reviews. Genetics* **10**:161-172.
511. **Stunkel W, Bernard HU.** 1999. The chromatin structure of the long control region of human papillomavirus type 16 represses viral oncoprotein expression. *Journal of virology* **73**:1918-1930.
512. **Thierry F.** 2009. Transcriptional regulation of the papillomavirus oncogenes by cellular and viral transcription factors in cervical carcinoma. *Virology* **384**:375-379.
513. **O'Connor MJ, Stunkel W, Zimmermann H, Koh CH, Bernard HU.** 1998. A novel YY1-independent silencer represses the activity of the human papillomavirus type 16 enhancer. *Journal of virology* **72**:10083-10092.
514. **O'Connor MJ, Stunkel W, Koh CH, Zimmermann H, Bernard HU.** 2000. The differentiation-specific factor CDP/Cut represses transcription and replication of human papillomaviruses through a conserved silencing element. *Journal of virology* **74**:401-410.
515. **Ai W, Toussaint E, Roman A.** 1999. CCAAT displacement protein binds to and negatively regulates human papillomavirus type 6 E6, E7, and E1 promoters. *Journal of virology* **73**:4220-4229.

516. **Demeret C, Yaniv M, Thierry F.** 1994. The E2 transcriptional repressor can compensate for Sp1 activation of the human papillomavirus type 18 early promoter. *Journal of virology* **68**:7075-7082.
517. **Tan SH, Gloss B, Bernard HU.** 1992. During negative regulation of the human papillomavirus-16 E6 promoter, the viral E2 protein can displace Sp1 from a proximal promoter element. *Nucleic acids research* **20**:251-256.
518. **Gloss B, Bernard HU.** 1990. The E6/E7 promoter of human papillomavirus type 16 is activated in the absence of E2 proteins by a sequence-aberrant Sp1 distal element. *Journal of virology* **64**:5577-5584.
519. **Tan SH, Leong LE, Walker PA, Bernard HU.** 1994. 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. *Journal of virology* **68**:6411-6420.
520. **Dostatni N, Lambert PF, Sousa R, Ham J, Howley PM, Yaniv M.** 1991. The functional BPV-1 E2 trans-activating protein can act as a repressor by preventing formation of the initiation complex. *Genes & development* **5**:1657-1671.
521. **Dong G, Broker TR, Chow LT.** 1994. Human papillomavirus type 11 E2 proteins repress the homologous E6 promoter by interfering with the binding of host transcription factors to adjacent elements. *Journal of virology* **68**:1115-1127.
522. **Chin MT, Broker TR, Chow LT.** 1989. Identification of a novel constitutive enhancer element and an associated binding protein: implications for human papillomavirus type 11 enhancer regulation. *Journal of virology* **63**:2967-2976.
523. **Cripe TP, Haugen TH, Turk JP, Tabatabai F, Schmid PG, 3rd, Durst M, Gissmann L, Roman A, Turek LP.** 1987. Transcriptional regulation of the human papillomavirus-16 E6-E7 promoter by a keratinocyte-dependent enhancer, and by viral E2 trans-activator and repressor gene products: implications for cervical carcinogenesis. *The EMBO journal* **6**:3745-3753.
524. **Gloss B, Bernard HU, Seedorf K, Klock G.** 1987. The upstream regulatory region of the human papilloma virus-16 contains an E2 protein-independent enhancer which is specific for cervical carcinoma cells and regulated by glucocorticoid hormones. *The EMBO journal* **6**:3735-3743.
525. **Lung MS, Mak WM, Murray V.** 2011. The use of a human papillomavirus 18 promoter for tissue-specific expression in cervical carcinoma cells. *Cellular & molecular biology letters* **16**:477-492.
526. **Swift FV, Bhat K, Younghusband HB, Hamada H.** 1987. Characterization of a cell type-specific enhancer found in the human papilloma virus type 18 genome. *The EMBO journal* **6**:1339-1344.
527. **Zhao W, Chow LT, Broker TR.** 1997. Transcription activities of human papillomavirus type 11 E6 promoter-proximal elements in raft and submerged cultures of foreskin keratinocytes. *Journal of virology* **71**:8832-8840.
528. **Cuthill S, Sibbet GJ, Campo MS.** 1993. Characterization of a nuclear factor, papilloma enhancer binding factor-1, that binds the long control region of human papillomavirus type 16 and contributes to enhancer activity. *Molecular carcinogenesis* **8**:96-104.
529. **Sibbet GJ, Cuthill S, Campo MS.** 1995. The enhancer in the long control region of human papillomavirus type 16 is up-regulated by PEF-1 and down-regulated by Oct-1. *Journal of virology* **69**:4006-4011.

530. **Mack DH, Laimins LA.** 1991. A keratinocyte-specific transcription factor, KRF-1, interacts with AP-1 to activate expression of human papillomavirus type 18 in squamous epithelial cells. *Proceedings of the National Academy of Sciences of the United States of America* **88**:9102-9106.
531. **Chan WK, Klock G, Bernard HU.** 1989. Progesterone and glucocorticoid response elements occur in the long control regions of several human papillomaviruses involved in anogenital neoplasia. *Journal of virology* **63**:3261-3269.
532. **Mittal R, Pater A, Pater MM.** 1993. Multiple human papillomavirus type 16 glucocorticoid response elements functional for transformation, transient expression, and DNA-protein interactions. *Journal of virology* **67**:5656-5659.
533. **Mittal R, Tsutsumi K, Pater A, Pater MM.** 1993. Human papillomavirus type 16 expression in cervical keratinocytes: role of progesterone and glucocorticoid hormones. *Obstetrics and gynecology* **81**:5-12.
534. **Cancer ICoESoC.** 2006. Cervical carcinoma and reproductive factors: collaborative reanalysis of individual data on 16,563 women with cervical carcinoma and 33,542 women without cervical carcinoma from 25 epidemiological studies. *International journal of cancer. Journal international du cancer* **119**:1108-1124.
535. **Roman A, Brown D.** 1995. Sequence variation in the extreme 5' end of the human papillomavirus type 6a long control region. *The Journal of infectious diseases* **171**:697-700.
536. **Kanaya T, Kyo S, Laimins LA.** 1997. The 5' region of the human papillomavirus type 31 upstream regulatory region acts as an enhancer which augments viral early expression through the action of YY1. *Virology* **237**:159-169.
537. **Hubert WG, Kanaya T, Laimins LA.** 1999. DNA replication of human papillomavirus type 31 is modulated by elements of the upstream regulatory region that lie 5' of the minimal origin. *Journal of virology* **73**:1835-1845.
538. **Tan SH, Bartsch D, Schwarz E, Bernard HU.** 1998. Nuclear matrix attachment regions of human papillomavirus type 16 point toward conservation of these genomic elements in all genital papillomaviruses. *Journal of virology* **72**:3610-3622.
539. **Stunkel W, Huang Z, Tan SH, O'Connor MJ, Bernard HU.** 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. *Journal of virology* **74**:2489-2501.
540. **Pederson T.** 1998. Thinking about a nuclear matrix. *Journal of molecular biology* **277**:147-159.
541. **Bouallaga I, Massicard S, Yaniv M, Thierry F.** 2000. An enhanceosome containing the Jun B/Fra-2 heterodimer and the HMG-I(Y) architectural protein controls HPV 18 transcription. *EMBO reports* **1**:422-427.
542. **Carey M.** 1998. The enhanceosome and transcriptional synergy. *Cell* **92**:5-8.
543. **Bouallaga I, Teissier S, Yaniv M, Thierry F.** 2003. HMG-I(Y) and the CBP/p300 coactivator are essential for human papillomavirus type 18 enhanceosome transcriptional activity. *Molecular and cellular biology* **23**:2329-2340.
544. **Alvarez J, Coutlee F, Archambault J.** 2016. A Rare Polymorphism Enhances the Transcriptional Activity of the HPV33 Long Control Region by Disrupting a Binding Site for the Cellular Transcription Factor C/EBP $\beta$  (unpublished).

545. **Ottinger M, Smith JA, Schweiger MR, Robbins D, Powell ML, You J, Howley PM.** 2009. Cell-type specific transcriptional activities among different papillomavirus long control regions and their regulation by E2. *Virology* **395**:161-171.
546. **Syrjanen S.** 2003. Human papillomavirus infections and oral tumors. *Medical microbiology and immunology* **192**:123-128.
547. **Dickens P, Srivastava G, Loke SL, Larkin S.** 1991. Human papillomavirus 6, 11, and 16 in laryngeal papillomas. *The Journal of pathology* **165**:243-246.
548. **Ball SL, Winder DM, Vaughan K, Hanna N, Levy J, Sterling JC, Stanley MA, Goon PK.** 2011. Analyses of human papillomavirus genotypes and viral loads in anogenital warts. *Journal of medical virology* **83**:1345-1350.
549. **Miyauchi T, Moriuchi R, Hamade Y, Suzuki S, Nomura T, Shimizu S.** 2015. Warts in toe webs associated with human papillomavirus type 7: A specific cutaneous manifestation of this type? *The British journal of dermatology*.
550. **Mistry N, Simonsson M, Evander M.** 2007. Transcriptional activation of the human papillomavirus type 5 and 16 long control region in cells from cutaneous and mucosal origin. *Virology journal* **4**:27.
551. **Mistry N, Wibom C, Evander M.** 2008. Cutaneous and mucosal human papillomaviruses differ in net surface charge, potential impact on tropism. *Virology journal* **5**:118.
552. **Egawa N, Egawa K, Griffin H, Doorbar J.** 2015. Human Papillomaviruses; Epithelial Tropisms, and the Development of Neoplasia. *Viruses* **7**:3863-3890.
553. **Cid A, Auewarakul P, Garcia-Carranca A, Ovseiovich R, Gaissert H, Gissmann L.** 1993. Cell-type-specific activity of the human papillomavirus type 18 upstream regulatory region in transgenic mice and its modulation by tetradecanoyl phorbol acetate and glucocorticoids. *Journal of virology* **67**:6742-6752.
554. **Donalisio M, Poli A, Civra A, Landolfo S, Lembo D.** 2010. Effects of cytokines on long control region transcriptional activity in high-risk cutaneous human papillomavirus types 5 and 8. *Archives of virology* **155**:583-587.
555. **Vasiljevic N, Nielsen L, Doherty G, Dillner J, Forslund O, Norrild B.** 2008. Differences in transcriptional activity of cutaneous human papillomaviruses. *Virus research* **137**:213-219.
556. **Chong T, Apt D, Gloss B, Isa M, Bernard HU.** 1991. The enhancer of human papillomavirus type 16: binding sites for the ubiquitous transcription factors oct-1, NFA, TEF-2, NF1, and AP-1 participate in epithelial cell-specific transcription. *Journal of virology* **65**:5933-5943.
557. **Pabo CO, Sauer RT.** 1992. Transcription factors: structural families and principles of DNA recognition. *Annual review of biochemistry* **61**:1053-1095.
558. **Spitz F, Furlong EE.** 2012. Transcription factors: from enhancer binding to developmental control. *Nature reviews. Genetics* **13**:613-626.
559. **Schenker A, Straub E, Iftner T, Stubenrauch F.** 2013. Cell-type-dependent activities of regulatory regions and E2 proteins derived from carcinogenic and non-carcinogenic human alphapapillomaviruses. *The Journal of general virology* **94**:1343-1350.
560. **Sailaja G, Watts RM, Bernard HU.** 1999. Many different papillomaviruses have low transcriptional activity in spite of strong epithelial specific enhancers. *The Journal of general virology* **80 ( Pt 7)**:1715-1724.

561. **Bromberg-White JL, Meyers C.** 2003. Comparison of the basal and glucocorticoid-inducible activities of the upstream regulatory regions of HPV18 and HPV31 in multiple epithelial cell lines. *Virology* **306**:197-202.
562. **Ralph WM, Jr., Liu K, Auborn KJ.** 2006. CCAAT/enhancer-binding protein beta represses human papillomavirus 11 upstream regulatory region expression through a promoter-proximal YY1-binding site. *The Journal of general virology* **87**:51-59.
563. **Struyk L, van der Meijden E, Minnaar R, Fontaine V, Meijer I, ter Schegget J.** 2000. Transcriptional regulation of human papillomavirus type 16 LCR by different C/EBPbeta isoforms. *Molecular carcinogenesis* **28**:42-50.
564. **Hegde RS.** 2002. The papillomavirus E2 proteins: structure, function, and biology. *Annual review of biophysics and biomolecular structure* **31**:343-360.
565. **Li R, Knight J, Bream G, Stenlund A, Botchan M.** 1989. Specific recognition nucleotides and their DNA context determine the affinity of E2 protein for 17 binding sites in the BPV-1 genome. *Genes & development* **3**:510-526.
566. **Stenlund A, Botchan MR.** 1990. The E2 trans-activator can act as a repressor by interfering with a cellular transcription factor. *Genes & development* **4**:476.
567. **Russell J, Botchan MR.** 1995. cis-Acting components of human papillomavirus (HPV) DNA replication: linker substitution analysis of the HPV type 11 origin. *Journal of virology* **69**:651-660.
568. **Stubenrauch F, Lim HB, Laimins LA.** 1998. Differential requirements for conserved E2 binding sites in the life cycle of oncogenic human papillomavirus type 31. *Journal of virology* **72**:1071-1077.
569. **Hoppe-Seyler F, Butz K.** 1992. Activation of human papillomavirus type 18 E6-E7 oncogene expression by transcription factor Sp1. *Nucleic acids research* **20**:6701-6706.
570. **Lung MS, Zhang N, Murray V.** 2012. Site-directed mutagenesis of human papillomavirus 18 promoter elements and tissue-specific expression in cervical carcinoma cells. *Virus genes* **44**:395-402.
571. **Dong XP, Pfister H.** 1999. Overlapping YY1- and aberrant SP1-binding sites proximal to the early promoter of human papillomavirus type 16. *The Journal of general virology* **80 ( Pt 8)**:2097-2101.
572. **Apt D, Watts RM, Suske G, Bernard HU.** 1996. High Sp1/Sp3 ratios in epithelial cells during epithelial differentiation and cellular transformation correlate with the activation of the HPV-16 promoter. *Virology* **224**:281-291.
573. **Shaulian E, Karin M.** 2002. AP-1 as a regulator of cell life and death. *Nature cell biology* **4**:E131-136.
574. **Zenz R, Wagner EF.** 2006. Jun signalling in the epidermis: From developmental defects to psoriasis and skin tumors. *The international journal of biochemistry & cell biology* **38**:1043-1049.
575. **Zhou H, Zarubin T, Ji Z, Min Z, Zhu W, Downey JS, Lin S, Han J.** 2005. Frequency and distribution of AP-1 sites in the human genome. *DNA research : an international journal for rapid publication of reports on genes and genomes* **12**:139-150.
576. **Hess J, Angel P, Schorpp-Kistner M.** 2004. AP-1 subunits: quarrel and harmony among siblings. *Journal of cell science* **117**:5965-5973.
577. **Mehic D, Bakiri L, Ghannadan M, Wagner EF, Tschachler E.** 2005. Fos and jun proteins are specifically expressed during differentiation of human keratinocytes. *The Journal of investigative dermatology* **124**:212-220.

578. **Nakshatri H, Pater MM, Pater A.** 1990. Ubiquitous and cell-type-specific protein interactions with human papillomavirus type 16 and type 18 enhancers. *Virology* **178**:92-103.
579. **Offord EA, Chappuis PO, Beard P.** 1993. Different stability of AP1 proteins in human keratinocyte and fibroblast cells: possible role in the cell-type specific expression of human papillomavirus type 18 genes. *Carcinogenesis* **14**:2447-2455.
580. **Kyo S, Tam A, Laimins LA.** 1995. Transcriptional activity of human papillomavirus type 31b enhancer is regulated through synergistic interaction of AP1 with two novel cellular factors. *Virology* **211**:184-197.
581. **Kyo S, Klumpp DJ, Inoue M, Kanaya T, Laimins LA.** 1997. Expression of AP1 during cellular differentiation determines human papillomavirus E6/E7 expression in stratified epithelial cells. *The Journal of general virology* **78 ( Pt 2)**:401-411.
582. **Garcia-Carranca A, Thierry F, Yaniv M.** 1988. Interplay of viral and cellular proteins along the long control region of human papillomavirus type 18. *Journal of virology* **62**:4321-4330.
583. **Chan WK, Chong T, Bernard HU, Klock G.** 1990. Transcription of the transforming genes of the oncogenic human papillomavirus-16 is stimulated by tumor promoters through AP1 binding sites. *Nucleic acids research* **18**:763-769.
584. **Sibbet GJ, Campo MS.** 1990. Multiple interactions between cellular factors and the non-coding region of human papillomavirus type 16. *The Journal of general virology* **71 ( Pt 11)**:2699-2707.
585. **Offord EA, Beard P.** 1990. A member of the activator protein 1 family found in keratinocytes but not in fibroblasts required for transcription from a human papillomavirus type 18 promoter. *Journal of virology* **64**:4792-4798.
586. **Thierry F, Spyrou G, Yaniv M, Howley P.** 1992. Two AP1 sites binding JunB are essential for human papillomavirus type 18 transcription in keratinocytes. *Journal of virology* **66**:3740-3748.
587. **Butz K, Hoppe-Seyler F.** 1993. Transcriptional control of human papillomavirus (HPV) oncogene expression: composition of the HPV type 18 upstream regulatory region. *Journal of virology* **67**:6476-6486.
588. **Dollard SC, Broker TR, Chow LT.** 1993. Regulation of the human papillomavirus type 11 E6 promoter by viral and host transcription factors in primary human keratinocytes. *Journal of virology* **67**:1721-1726.
589. **Gronostajski RM.** 2000. Roles of the NFI/CTF gene family in transcription and development. *Gene* **249**:31-45.
590. **Gloss B, Yeo-Gloss M, Meisterenst M, Rogge L, Winnacker EL, Bernard HU.** 1989. Clusters of nuclear factor I binding sites identify enhancers of several papillomaviruses but alone are not sufficient for enhancer function. *Nucleic acids research* **17**:3519-3533.
591. **Chong T, Chan WK, Bernard HU.** 1990. Transcriptional activation of human papillomavirus 16 by nuclear factor I, AP1, steroid receptors and a possibly novel transcription factor, PVF: a model for the composition of genital papillomavirus enhancers. *Nucleic acids research* **18**:465-470.
592. **Rho J, Lee S, de Villiers EM, Choe J.** 1997. Identification of cis-regulatory elements in the upstream regulatory region of human papillomavirus type 59. *Virus research* **47**:155-166.



593. **Hoppe-Seyler F, Butz K, zur Hausen H.** 1991. Repression of the human papillomavirus type 18 enhancer by the cellular transcription factor Oct-1. *Journal of virology* **65**:5613-5618.
594. **Royer HD, Freyaldenhoven MP, Napierski I, Spitkovsky DD, Bauknecht T, Dathan N.** 1991. Delineation of human papillomavirus type 18 enhancer binding proteins: the intracellular distribution of a novel octamer binding protein p92 is cell cycle regulated. *Nucleic acids research* **19**:2363-2371.
595. **Dent CL, McIndoe GA, Latchman DS.** 1991. The constitutively expressed octamer binding protein OTF-1 and a novel octamer binding protein expressed specifically in cervical cells bind to an octamer-related sequence in the human papillomavirus 16 enhancer. *Nucleic acids research* **19**:4531-4535.
596. **Morris PJ, Ring CJ, Lillycrop KA, Latchman DS.** 1993. Transactivation of the human papilloma virus 16 octamer motif by the octamer binding protein Oct-2 requires both the N and C terminal activation domains. *Nucleic acids research* **21**:4506-4510.
597. **O'Connor M, Bernard HU.** 1995. Oct-1 activates the epithelial-specific enhancer of human papillomavirus type 16 via a synergistic interaction with NFI at a conserved composite regulatory element. *Virology* **207**:77-88.
598. **Gordon S, Akopyan G, Garban H, Bonavida B.** 2006. Transcription factor YY1: structure, function, and therapeutic implications in cancer biology. *Oncogene* **25**:1125-1142.
599. **Margolis DM, Somasundaran M, Green MR.** 1994. Human transcription factor YY1 represses human immunodeficiency virus type 1 transcription and virion production. *Journal of virology* **68**:905-910.
600. **Mills LK, Shi Y, Millette RL.** 1994. YY1 is the cellular factor shown previously to bind to regulatory regions of several leaky-late (beta gamma, gamma 1) genes of herpes simplex virus type 1. *Journal of virology* **68**:1234-1238.
601. **O'Connor MJ, Tan SH, Tan CH, Bernard HU.** 1996. YY1 represses human papillomavirus type 16 transcription by quenching AP-1 activity. *Journal of virology* **70**:6529-6539.
602. **Bauknecht T, Jundt F, Herr I, Oehler T, Delius H, Shi Y, Angel P, Zur Hausen H.** 1995. A switch region determines the cell type-specific positive or negative action of YY1 on the activity of the human papillomavirus type 18 promoter. *Journal of virology* **69**:1-12.
603. **Dong XP, Stubenrauch F, Beyer-Finkler E, Pfister H.** 1994. Prevalence of deletions of YY1-binding sites in episomal HPV 16 DNA from cervical cancers. *International journal of cancer. Journal international du cancer* **58**:803-808.
604. **Veress G, Murvai M, Szarka K, Juhasz A, Konya J, Gergely L.** 2001. Transcriptional activity of human papillomavirus type 16 variants having deletions in the long control region. *European journal of cancer (Oxford, England : 1990)* **37**:1946-1952.
605. **Lace MJ, Yamakawa Y, Ushikai M, Anson JR, Haugen TH, Turek LP.** 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. *The Journal of general virology* **90**:2402-2412.
606. **Bauknecht T, Shi Y.** 1998. Overexpression of C/EBPbeta represses human papillomavirus type 18 upstream regulatory region activity in HeLa cells by

- interfering with the binding of TATA-binding protein. *Journal of virology* **72**:2113-2124.
607. **Bauknecht T, Angel P, Royer HD, zur Hausen H.** 1992. Identification of a negative regulatory domain in the human papillomavirus type 18 promoter: interaction with the transcriptional repressor YY1. *The EMBO journal* **11**:4607-4617.
608. **Jundt F, Herr I, Angel P, Zur Hausen H, Bauknecht T.** 1995. Transcriptional control of human papillomavirus type 18 oncogene expression in different cell lines: role of transcription factor YY1. *Virus genes* **11**:53-58.
609. **Bauknecht T, See RH, Shi Y.** 1996. A novel C/EBP beta-YY1 complex controls the cell-type-specific activity of the human papillomavirus type 18 upstream regulatory region. *Journal of virology* **70**:7695-7705.
610. **Seligson D, Horvath S, Huerta-Yepez S, Hanna S, Garban H, Roberts A, Shi T, Liu X, Chia D, Goodglick L, Bonavida B.** 2005. Expression of transcription factor Yin Yang 1 in prostate cancer. *International journal of oncology* **27**:131-141.
611. **Wan M, Huang W, Kute TE, Miller LD, Zhang Q, Hatcher H, Wang J, Stovall DB, Russell GB, Cao PD, Deng Z, Wang W, Zhang Q, Lei M, Torti SV, Akman SA, Sui G.** 2012. Yin Yang 1 plays an essential role in breast cancer and negatively regulates p27. *The American journal of pathology* **180**:2120-2133.
612. **Baritaki S, Sifakis S, Huerta-Yepez S, Neonakis IK, Soufla G, Bonavida B, Spandidos DA.** 2007. Overexpression of VEGF and TGF-beta1 mRNA in Pap smears correlates with progression of cervical intraepithelial neoplasia to cancer: implication of YY1 in cervical tumorigenesis and HPV infection. *International journal of oncology* **31**:69-79.
613. **He G, Wang Q, Zhou Y, Wu X, Wang L, Duru N, Kong X, Zhang P, Wan B, Sui L, Guo Q, Li JJ, Yu L.** 2011. YY1 is a novel potential therapeutic target for the treatment of HPV infection-induced cervical cancer by arsenic trioxide. *International journal of gynecological cancer : official journal of the International Gynecological Cancer Society* **21**:1097-1104.
614. **May M, Dong XP, Beyer-Finkler E, Stubenrauch F, Fuchs PG, Pfister H.** 1994. The E6/E7 promoter of extrachromosomal HPV16 DNA in cervical cancers escapes from cellular repression by mutation of target sequences for YY1. *The EMBO journal* **13**:1460-1466.
615. **Lace MJ, Isacson C, Anson JR, Lorincz AT, Wilczynski SP, Haugen TH, Turek LP.** 2009. Upstream regulatory region alterations found in human papillomavirus type 16 (HPV-16) isolates from cervical carcinomas increase transcription, ori function, and HPV immortalization capacity in culture. *Journal of virology* **83**:7457-7466.
616. **Lee KY, Broker TR, Chow LT.** 1998. Transcription factor YY1 represses cell-free replication from human papillomavirus origins. *Journal of virology* **72**:4911-4917.
617. **Nerlov C.** 2007. The C/EBP family of transcription factors: a paradigm for interaction between gene expression and proliferation control. *Trends in cell biology* **17**:318-324.
618. **Tsukada J, Yoshida Y, Kominato Y, Auron PE.** 2011. The CCAAT/enhancer (C/EBP) family of basic-leucine zipper (bZIP) transcription factors is a multifaceted highly-regulated system for gene regulation. *Cytokine* **54**:6-19.
619. **Niehof M, Manns MP, Trautwein C.** 1997. CREB controls LAP/C/EBP beta transcription. *Molecular and cellular biology* **17**:3600-3613.

620. **Berrier A, Siu G, Calame K.** 1998. Transcription of a minimal promoter from the NF-IL6 gene is regulated by CREB/ATF and SP1 proteins in U937 promonocytic cells. *Journal of immunology (Baltimore, Md. : 1950)* **161**:2267-2275.
621. **Lin WC, Shen BJ, Tsay YG, Yen HC, Lee SC, Chang CJ.** 2002. Transcriptional activation of C/EBPbeta gene by c-Jun and ATF2. *DNA and cell biology* **21**:551-560.
622. **Zhang JW, Klemm DJ, Vinson C, Lane MD.** 2004. Role of CREB in transcriptional regulation of CCAAT/enhancer-binding protein beta gene during adipogenesis. *The Journal of biological chemistry* **279**:4471-4478.
623. **Niehof M, Kubicka S, Zender L, Manns MP, Trautwein C.** 2001. Autoregulation enables different pathways to control CCAAT/enhancer binding protein beta (C/EBP beta) transcription. *Journal of molecular biology* **309**:855-868.
624. **Jimenez-Preitner M, Berney X, Uldry M, Vitali A, Cinti S, Ledford JG, Thorens B.** 2011. Plac8 is an inducer of C/EBPbeta required for brown fat differentiation, thermoregulation, and control of body weight. *Cell metabolism* **14**:658-670.
625. **Osada S, Yamamoto H, Nishihara T, Imagawa M.** 1996. DNA binding specificity of the CCAAT/enhancer-binding protein transcription factor family. *The Journal of biological chemistry* **271**:3891-3896.
626. **Pulido-Salgado M, Vidal-Taboada JM, Saura J.** 2015. C/EBPbeta and C/EBPdelta transcription factors: Basic biology and roles in the CNS. *Progress in neurobiology* **132**:1-33.
627. **Descombes P, Schibler U.** 1991. A liver-enriched transcriptional activator protein, LAP, and a transcriptional inhibitory protein, LIP, are translated from the same mRNA. *Cell* **67**:569-579.
628. **Gunasekharan V, Hache G, Laimins L.** 2012. Differentiation-dependent changes in levels of C/EBPbeta repressors and activators regulate human papillomavirus type 31 late gene expression. *Journal of virology* **86**:5393-5398.
629. **Begay V, Smink JJ, Loddenkemper C, Zimmermann K, Rudolph C, Scheller M, Steinemann D, Leser U, Schlegelberger B, Stein H, Leutz A.** 2015. Deregulation of the endogenous C/EBPbeta LIP isoform predisposes to tumorigenesis. *Journal of molecular medicine (Berlin, Germany)* **93**:39-49.
630. **Dearth LR, Hutt J, Sattler A, Gigliotti A, DeWille J.** 2001. Expression and function of CCAAT/enhancer binding proteinbeta (C/EBPbeta) LAP and LIP isoforms in mouse mammary gland, tumors and cultured mammary epithelial cells. *Journal of cellular biochemistry* **82**:357-370.
631. **Milde-Langosch K, Loning T, Bamberger AM.** 2003. Expression of the CCAAT/enhancer-binding proteins C/EBPalpha, C/EBPbeta and C/EBPdelta in breast cancer: correlations with clinicopathologic parameters and cell-cycle regulatory proteins. *Breast cancer research and treatment* **79**:175-185.
632. **Arnett B, Soisson P, Ducatman BS, Zhang P.** 2003. Expression of CAAT enhancer binding protein beta (C/EBP beta) in cervix and endometrium. *Molecular cancer* **2**:21.
633. **Rask K, Thorn M, Ponten F, Kraaz W, Sundfeldt K, Hedin L, Enerback S.** 2000. Increased expression of the transcription factors CCAAT-enhancer binding protein-beta (C/EBPbeta) and C/EBPzeta (CHOP) correlate with invasiveness of human colorectal cancer. *International journal of cancer. Journal international du cancer* **86**:337-343.

634. **Gomis RR, Alarcon C, Nadal C, Van Poznak C, Massague J.** 2006. C/EBPbeta at the core of the TGFbeta cytosstatic response and its evasion in metastatic breast cancer cells. *Cancer cell* **10**:203-214.
635. **Auborn KJ, Galli RL, Dilorenzo TP, Steinberg BM.** 1989. Identification of DNA-protein interactions and enhancer activity at the 5' end of the upstream regulatory region in human papillomavirus type 11. *Virology* **170**:123-130.
636. **Auborn KJ, Steinberg BM.** 1991. A key DNA-protein interaction determines the function of the 5'URR enhancer in human papillomavirus type 11. *Virology* **181**:132-138.
637. **Kyo S, Inoue M, Nishio Y, Nakanishi K, Akira S, Inoue H, Yutsudo M, Tanizawa O, Hakura A.** 1993. NF-IL6 represses early gene expression of human papillomavirus type 16 through binding to the noncoding region. *Journal of virology* **67**:1058-1066.
638. **Wang H, Liu K, Yuan F, Berdichevsky L, Taichman LB, Auborn K.** 1996. C/EBPbeta is a negative regulator of human papillomavirus type 11 in keratinocytes. *Journal of virology* **70**:4839-4844.
639. **Zhao W, Chow LT, Broker TR.** 1999. A distal element in the HPV-11 upstream regulatory region contributes to promoter repression in basal keratinocytes in squamous epithelium. *Virology* **253**:219-229.
640. **Gyongyosi E, Szalmas A, Ferenczi A, Poliska S, Konya J, Veress G.** 2015. Transcriptional regulation of genes involved in keratinocyte differentiation by human papillomavirus 16 oncoproteins. *Archives of virology* **160**:389-398.
641. **Sperling T, Oldak M, Walch-Ruckheim B, Wickenhauser C, Doorbar J, Pfister H, Malejczyk M, Majewski S, Keates AC, Smola S.** 2012. Human papillomavirus type 8 interferes with a novel C/EBPbeta-mediated mechanism of keratinocyte CCL20 chemokine expression and Langerhans cell migration. *PLoS pathogens* **8**:e1002833.
642. **Medina-Martinez O, Morales-Peza N, Yaniv M, Garcia-Carranca A, Thierry F.** 1996. A single element mediates glucocorticoid hormone response of HPV18 with no functional interactions with AP1 or hbrm. *Virology* **217**:392-396.
643. **Kyo S, Inoue M, Hayasaka N, Inoue T, Yutsudo M, Tanizawa O, Hakura A.** 1994. Regulation of early gene expression of human papillomavirus type 16 by inflammatory cytokines. *Virology* **200**:130-139.
644. **Baldwin A, Pirisi L, Creek KE.** 2004. NFI-Ski interactions mediate transforming growth factor beta modulation of human papillomavirus type 16 early gene expression. *Journal of virology* **78**:3953-3964.
645. **Lembo D, Donalisio M, De Andrea M, Cornaglia M, Scutera S, Musso T, Landolfo S.** 2006. A cell-based high-throughput assay for screening inhibitors of human papillomavirus-16 long control region activity. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **20**:148-150.
646. **Kim K, Garner-Hamrick PA, Fisher C, Lee D, Lambert PF.** 2003. Methylation patterns of papillomavirus DNA, its influence on E2 function, and implications in viral infection. *Journal of virology* **77**:12450-12459.
647. **Bhattacharjee B, Sengupta S.** 2006. CpG methylation of HPV 16 LCR at E2 binding site proximal to P97 is associated with cervical cancer in presence of intact E2. *Virology* **354**:280-285.

648. **Kalantari M, Calleja-Macias IE, Tewari D, Hagmar B, Lie K, Barrera-Saldana HA, Wiley DJ, Bernard HU.** 2004. Conserved methylation patterns of human papillomavirus type 16 DNA in asymptomatic infection and cervical neoplasia. *Journal of virology* **78**:12762-12772.
649. **Hublarova P, Hrstka R, Rotterova P, Rotter L, Coupkova M, Badal V, Nenutil R, Vojtesek B.** 2009. Prediction of human papillomavirus 16 e6 gene expression and cervical intraepithelial neoplasia progression by methylation status. *International journal of gynecological cancer : official journal of the International Gynecological Cancer Society* **19**:321-325.
650. **Jacquin E, Baraquin A, Ramanah R, Carcopino X, Morel A, Valmary-Degano S, Bravo IG, de Sanjose S, Riethmuller D, Mouglin C, Pretet JL.** 2013. Methylation of human papillomavirus Type 16 CpG sites at E2-binding site 1 (E2BS1), E2BS2, and the Sp1-binding site in cervical cancer samples as determined by high-resolution melting analysis-PCR. *Journal of clinical microbiology* **51**:3207-3215.
651. **Park IS, Chang X, Loyo M, Wu G, Chuang A, Kim MS, Chae YK, Lyford-Pike S, Westra WH, Saunders JR, Sidransky D, Pai SI.** 2011. Characterization of the methylation patterns in human papillomavirus type 16 viral DNA in head and neck cancers. *Cancer prevention research (Philadelphia, Pa.)* **4**:207-217.
652. **Hummel M, Hudson JB, Laimins LA.** 1992. Differentiation-induced and constitutive transcription of human papillomavirus type 31b in cell lines containing viral episomes. *Journal of virology* **66**:6070-6080.
653. **Ruesch MN, Stubenrauch F, Laimins LA.** 1998. Activation of papillomavirus late gene transcription and genome amplification upon differentiation in semisolid medium is coincident with expression of involucrin and transglutaminase but not keratin-10. *Journal of virology* **72**:5016-5024.
654. **Bodily JM, Meyers C.** 2005. Genetic analysis of the human papillomavirus type 31 differentiation-dependent late promoter. *Journal of virology* **79**:3309-3321.
655. **Bodily JM, Mehta KP, Cruz L, Meyers C, Laimins LA.** 2011. The E7 open reading frame acts in cis and in trans to mediate differentiation-dependent activities in the human papillomavirus type 16 life cycle. *Journal of virology* **85**:8852-8862.
656. **Ozbun MA, Meyers C.** 1997. Characterization of late gene transcripts expressed during vegetative replication of human papillomavirus type 31b. *Journal of virology* **71**:5161-5172.
657. **Ozbun MA, Meyers C.** 1998. Temporal usage of multiple promoters during the life cycle of human papillomavirus type 31b. *Journal of virology* **72**:2715-2722.
658. **Grassmann K, Rapp B, Maschek H, Petry KU, Iftner T.** 1996. Identification of a differentiation-inducible promoter in the E7 open reading frame of human papillomavirus type 16 (HPV-16) in raft cultures of a new cell line containing high copy numbers of episomal HPV-16 DNA. *Journal of virology* **70**:2339-2349.
659. **Kukimoto I, Takeuchi T, Kanda T.** 2006. CCAAT/enhancer binding protein beta binds to and activates the P670 promoter of human papillomavirus type 16. *Virology* **346**:98-107.
660. **Bernard HU.** 2002. Gene expression of genital human papillomaviruses and considerations on potential antiviral approaches. *Antiviral therapy* **7**:219-237.
661. **Wilson R, Ryan GB, Knight GL, Laimins LA, Roberts S.** 2007. The full-length E1E4 protein of human papillomavirus type 18 modulates differentiation-dependent viral DNA amplification and late gene expression. *Virology* **362**:453-460.

662. **Muller-Schiffmann A, Beckmann J, Steger G.** 2006. The E6 protein of the cutaneous human papillomavirus type 8 can stimulate the viral early and late promoters by distinct mechanisms. *Journal of virology* **80**:8718-8728.
663. **Sen E, Alam S, Meyers C.** 2004. Genetic and biochemical analysis of cis regulatory elements within the keratinocyte enhancer region of the human papillomavirus type 31 upstream regulatory region during different stages of the viral life cycle. *Journal of virology* **78**:612-629.
664. **Carson A, Khan SA.** 2006. Characterization of transcription factor binding to human papillomavirus type 16 DNA during cellular differentiation. *Journal of virology* **80**:4356-4362.
665. **Spink KM, Laimins LA.** 2005. Induction of the human papillomavirus type 31 late promoter requires differentiation but not DNA amplification. *Journal of virology* **79**:4918-4926.
666. **Kajitani N, Schwartz S.** 2015. RNA Binding Proteins that Control Human Papillomavirus Gene Expression. *Biomolecules* **5**:758-774.
667. **Rush M, Zhao X, Schwartz S.** 2005. A splicing enhancer in the E4 coding region of human papillomavirus type 16 is required for early mRNA splicing and polyadenylation as well as inhibition of premature late gene expression. *Journal of virology* **79**:12002-12015.
668. **Hindmarsh PL, Laimins LA.** 2007. Mechanisms regulating expression of the HPV 31 L1 and L2 capsid proteins and pseudovirion entry. *Virology journal* **4**:19.
669. **Jia R, Liu X, Tao M, Kruhlak M, Guo M, Meyers C, Baker CC, Zheng ZM.** 2009. Control of the papillomavirus early-to-late switch by differentially expressed SRp20. *Journal of virology* **83**:167-180.
670. **Li X, Johansson C, Glahder J, Mossberg AK, Schwartz S.** 2013. Suppression of HPV-16 late L1 5'-splice site SD3632 by binding of hnRNP D proteins and hnRNP A2/B1 to upstream AUAGUA RNA motifs. *Nucleic acids research* **41**:10488-10508.
671. **Pajunk HS, May C, Pfister H, Fuchs PG.** 1997. Regulatory interactions of transcription factor YY1 with control sequences of the E6 promoter of human papillomavirus type 8. *The Journal of general virology* **78 ( Pt 12)**:3287-3295.
672. **Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, Barre-Dirrie A, Reuter I, Chekmenev D, Krull M, Hornischer K, Voss N, Stegmaier P, Lewicki-Potapov B, Saxel H, Kel AE, Wingender E.** 2006. TRANSFAC and its module TRANSCOMP: transcriptional gene regulation in eukaryotes. *Nucleic acids research* **34**:D108-110.
673. **Mathelier A, Zhao X, Zhang AW, Parcy F, Worsley-Hunt R, Arenillas DJ, Buchman S, Chen CY, Chou A, Ienasescu H, Lim J, Shyr C, Tan G, Zhou M, Lenhard B, Sandelin A, Wasserman WW.** 2014. JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. *Nucleic acids research* **42**:D142-147.
674. **Veprintsev DB, Fersht AR.** 2008. Algorithm for prediction of tumour suppressor p53 affinity for binding sites in DNA. *Nucleic acids research* **36**:1589-1598.
675. **Zhou Q, Liu JS.** 2004. Modeling within-motif dependence for transcription factor binding site predictions. *Bioinformatics (Oxford, England)* **20**:909-916.
676. **Ishiji T, Lacey MJ, Parkkinen S, Anderson RD, Haugen TH, Cripe TP, Xiao JH, Davidson I, Chambon P, Turek LP.** 1992. Transcriptional enhancer factor (TEF)-1 and its cell-specific co-activator activate human papillomavirus-16 E6 and

- E7 oncogene transcription in keratinocytes and cervical carcinoma cells. The EMBO journal **11**:2271-2281.
677. **Geisen C, Kahn T.** 1996. Promoter activity of sequences located upstream of the human papillomavirus types of 16 and 18 late regions. The Journal of general virology **77 ( Pt 9)**:2193-2200.
678. **Weyn C, Vanderwinden JM, Rasschaert J, Englert Y, Fontaine V.** 2011. Regulation of human papillomavirus type 16 early gene expression in trophoblastic and cervical cells. Virology **412**:146-155.
679. **Sichero L, Sobrinho JS, Villa LL.** 2012. Identification of novel cellular transcription factors that regulate early promoters of human papillomavirus types 18 and 16. The Journal of infectious diseases **206**:867-874.
680. **Chen YH, Huang LH, Chen TM.** 1996. Differential effects of progestins and estrogens on long control regions of human papillomavirus types 16 and 18. Biochemical and biophysical research communications **224**:651-659.
681. **Chow LT.** 2015. Model systems to study the life cycle of human papillomaviruses and HPV-associated cancers. Virologica Sinica **30**:92-100.
682. **Anacker D, Moody C.** 2012. Generation of organotypic raft cultures from primary human keratinocytes. Journal of visualized experiments : JoVE.
683. **Ozbun MA.** 2002. Infectious human papillomavirus type 31b: purification and infection of an immortalized human keratinocyte cell line. The Journal of general virology **83**:2753-2763.
684. **Delvenne P, Hubert P, Jacobs N, Giannini SL, Havard L, Renard I, Saboulard D, Boniver J.** 2001. The organotypic culture of HPV-transformed keratinocytes: an effective in vitro model for the development of new immunotherapeutic approaches for mucosal (pre)neoplastic lesions. Vaccine **19**:2557-2564.
685. **Flores ER, Allen-Hoffmann BL, Lee D, Sattler CA, Lambert PF.** 1999. Establishment of the human papillomavirus type 16 (HPV-16) life cycle in an immortalized human foreskin keratinocyte cell line. Virology **262**:344-354.
686. **Wang HK, Duffy AA, Broker TR, Chow LT.** 2009. Robust production and passaging of infectious HPV in squamous epithelium of primary human keratinocytes. Genes & development **23**:181-194.
687. **Chen RW, Aalto Y, Teesalu T, Durst M, Knuutila S, Aaltonen LM, Vaheri A.** 2003. Establishment and characterisation of human papillomavirus type 16 DNA immortalised human tonsillar epithelial cell lines. European journal of cancer (Oxford, England : 1990) **39**:698-707.
688. **Westphal K, Akgul B, Storey A, Nindl I.** 2009. Cutaneous human papillomavirus E7 type-specific effects on differentiation and proliferation of organotypic skin cultures. Cellular oncology : the official journal of the International Society for Cellular Oncology **31**:213-226.
689. **Akgul B, Garcia-Escudero R, Ghali L, Pfister HJ, Fuchs PG, Navsaria H, Storey A.** 2005. The E7 protein of cutaneous human papillomavirus type 8 causes invasion of human keratinocytes into the dermis in organotypic cultures of skin. Cancer research **65**:2216-2223.
690. **Garner-Hamrick PA, Fostel JM, Chien WM, Banerjee NS, Chow LT, Broker TR, Fisher C.** 2004. Global effects of human papillomavirus type 18 E6/E7 in an organotypic keratinocyte culture system. Journal of virology **78**:9041-9050.

691. **Longworth MS, Laimins LA.** 2004. The binding of histone deacetylases and the integrity of zinc finger-like motifs of the E7 protein are essential for the life cycle of human papillomavirus type 31. *Journal of virology* **78**:3533-3541.
692. **McCloskey R, Menges C, Friedman A, Patel D, McCance DJ.** 2010. Human papillomavirus type 16 E6/E7 upregulation of nucleophosmin is important for proliferation and inhibition of differentiation. *Journal of virology* **84**:5131-5139.
693. **Ho L, Chan SY, Burk RD, Das BC, Fujinaga K, Icenogle JP, Kahn T, Kiviat N, Lancaster W, Mavromara-Nazos P, et al.** 1993. The genetic drift of human papillomavirus type 16 is a means of reconstructing prehistoric viral spread and the movement of ancient human populations. *Journal of virology* **67**:6413-6423.
694. **Woolford L, Rector A, Van Ranst M, Ducki A, Bennett MD, Nicholls PK, Warren KS, Swan RA, Wilcox GE, O'Hara AJ.** 2007. A novel virus detected in papillomas and carcinomas of the endangered western barred bandicoot (*Perameles bougainville*) exhibits genomic features of both the Papillomaviridae and Polyomaviridae. *Journal of virology* **81**:13280-13290.
695. **Marongiu L, Godi A, Parry JV, Beddows S.** 2014. Human papillomavirus type 16 long control region and E6 variants stratified by cervical disease stage. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* **26**:8-13.
696. **Pientong C, Wongwarissara P, Ekalaksananan T, Swangphon P, Kleebkaow P, Kongyingyoes B, Siriaunkgul S, Tungsinmunkong K, Suthipintawong C.** 2013. Association of human papillomavirus type 16 long control region mutation and cervical cancer. *Virology journal* **10**:30.
697. **Kammer C, Tommasino M, Syrjanen S, Delius H, Hebling U, Warthorst U, Pfister H, Zehbe I.** 2002. Variants of the long control region and the E6 oncogene in European human papillomavirus type 16 isolates: implications for cervical disease. *British journal of cancer* **86**:269-273.
698. **Xi LF, Koutsky LA, Hildesheim A, Galloway DA, Wheeler CM, Winer RL, Ho J, Kiviat NB.** 2007. Risk for high-grade cervical intraepithelial neoplasia associated with variants of human papillomavirus types 16 and 18. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **16**:4-10.
699. **Chen AA, Gheit T, Franceschi S, Tommasino M, Clifford GM.** 2015. Human Papillomavirus 18 Genetic Variation and Cervical Cancer Risk Worldwide. *Journal of virology* **89**:10680-10687.
700. **Lizano M, De la Cruz-Hernandez E, Carrillo-Garcia A, Garcia-Carranca A, Ponce de Leon-Rosales S, Duenas-Gonzalez A, Hernandez-Hernandez DM, Mohar A.** 2006. Distribution of HPV16 and 18 intratypic variants in normal cytology, intraepithelial lesions, and cervical cancer in a Mexican population. *Gynecologic oncology* **102**:230-235.
701. **Gagnon S, Hankins C, Tremblay C, Forest P, Pourreaux K, Coutlee F.** 2004. Viral polymorphism in human papillomavirus types 33 and 35 and persistent and transient infection in the genital tract of women. *The Journal of infectious diseases* **190**:1575-1585.
702. **Chen Z, Terai M, Fu L, Herrero R, DeSalle R, Burk RD.** 2005. Diversifying selection in human papillomavirus type 16 lineages based on complete genome analyses. *Journal of virology* **79**:7014-7023.



703. **Chen Z, DeSalle R, Schiffman M, Herrero R, Burk RD.** 2009. Evolutionary dynamics of variant genomes of human papillomavirus types 18, 45, and 97. *Journal of virology* **83**:1443-1455.
704. **Prado JC, Calleja-Macias IE, Bernard HU, Kalantari M, Macay SA, Allan B, Williamson AL, Chung LP, Collins RJ, Zuna RE, Dunn ST, Ortiz-Lopez R, Barrera-Saldana HA, Cubie HA, Cuschieri K, von Knebel-Doeberitz M, Sanchez GI, Bosch FX, Villa LL.** 2005. Worldwide genomic diversity of the human papillomaviruses-53, 56, and 66, a group of high-risk HPVs unrelated to HPV-16 and HPV-18. *Virology* **340**:95-104.
705. **Alvarez J, Coutlee F, Archambault J.** 2016. The Combination of Specific Nucleotide Variations Modulates the Transcriptional Activity of the Human Papillomavirus Type 33 Long Control Region (unpublished).
706. **Kukimoto I, Muramatsu M.** 2015. Genetic variations of human papillomavirus type 16: implications for cervical carcinogenesis. *Japanese journal of infectious diseases* **68**:169-175.
707. **Xi LF, Schiffman M, Koutsky LA, He Z, Winer RL, Hulbert A, Lee SK, Ke Y, Kiviat NB.** 2013. Persistence of newly detected human papillomavirus type 31 infection, stratified by variant lineage. *International journal of cancer. Journal international du cancer* **132**:549-555.
708. **Xi LF, Schiffman M, Koutsky LA, Hulbert A, Lee SK, Defilippis V, Shen Z, Kiviat NB.** 2012. Association of human papillomavirus type 31 variants with risk of cervical intraepithelial neoplasia grades 2-3. *International journal of cancer. Journal international du cancer* **131**:2300-2307.
709. **Formentin A, Archambault J, Koushik A, Richardson H, Brassard P, Franco EL, Coutlee F.** 2013. Human papillomavirus type 52 polymorphism and high-grade lesions of the uterine cervix. *International journal of cancer. Journal international du cancer* **132**:1821-1830.
710. **Chan PK, Lam CW, Cheung TH, Li WW, Lo KW, Chan MY, Cheung JL, Cheng AF.** 2002. Association of human papillomavirus type 58 variant with the risk of cervical cancer. *Journal of the National Cancer Institute* **94**:1249-1253.
711. **Kammer C, Warthorst U, Torrez-Martinez N, Wheeler CM, Pfister H.** 2000. Sequence analysis of the long control region of human papillomavirus type 16 variants and functional consequences for P97 promoter activity. *The Journal of general virology* **81**:1975-1981.
712. **Veress G, Szarka K, Dong XP, Gergely L, Pfister H.** 1999. Functional significance of sequence variation in the E2 gene and the long control region of human papillomavirus type 16. *The Journal of general virology* **80 ( Pt 4)**:1035-1043.
713. **Sichero L, Franco EL, Villa LL.** 2005. Different P105 promoter activities among natural variants of human papillomavirus type 18. *The Journal of infectious diseases* **191**:739-742.
714. **Hubert WG.** 2005. Variant upstream regulatory region sequences differentially regulate human papillomavirus type 16 DNA replication throughout the viral life cycle. *Journal of virology* **79**:5914-5922.
715. **Amador-Molina A, Gonzalez-Montoya JL, Garcia-Carranca A, Mohar A, Lizano M.** 2013. Intratypic changes of the E1 gene and the long control region affect ori function of human papillomavirus type 18 variants. *The Journal of general virology* **94**:393-402.

716. **Cornet I, Gheit T, Iannacone MR, Vignat J, Sylla BS, Del Mistro A, Franceschi S, Tommasino M, Clifford GM.** 2013. HPV16 genetic variation and the development of cervical cancer worldwide. *British journal of cancer* **108**:240-244.
717. **Richard C, Lanner C, Naryzhny SN, Sherman L, Lee H, Lambert PF, Zehbe I.** 2010. The immortalizing and transforming ability of two common human papillomavirus 16 E6 variants with different prevalences in cervical cancer. *Oncogene* **29**:3435-3445.
718. **Mesplede T, Gagnon D, Bergeron-Labrecque F, Azar I, Senechal H, Coutlee F, Archambault J.** 2012. p53 degradation activity, expression, and subcellular localization of E6 proteins from 29 human papillomavirus genotypes. *Journal of virology* **86**:94-107.
719. **Gurgel AP, Chagas BS, do Amaral CM, Nascimento KC, Leal LR, Silva Neto Jda C, Cartaxo Muniz MT, de Freitas AC.** 2015. Prevalence of Human Papillomavirus Variants and Genetic Diversity in the L1 Gene and Long Control Region of HPV16, HPV31, and HPV58 Found in North-East Brazil. *BioMed research international* **2015**:130828.
720. **Mandal P, Bhattacharjee B, Das Ghosh D, Mondal NR, Roy Chowdhury R, Roy S, Sengupta S.** 2013. Differential expression of HPV16 L2 gene in cervical cancers harboring episomal HPV16 genomes: influence of synonymous and non-coding region variations. *PloS one* **8**:e65647.
721. **Qmichou Z, Khyatti M, Berraho M, Ennaji MM, Benbacer L, Nejari C, Benjaafar N, Benider A, Attaleb M, El Mzibri M.** 2013. Analysis of mutations in the E6 oncogene of human papillomavirus 16 in cervical cancer isolates from Moroccan women. *BMC infectious diseases* **13**:378.
722. **Niccoli S, Abraham S, Richard C, Zehbe I.** 2012. The Asian-American E6 variant protein of human papillomavirus 16 alone is sufficient to promote immortalization, transformation, and migration of primary human foreskin keratinocytes. *Journal of virology* **86**:12384-12396.
723. **Smith B, Chen Z, Reimers L, van Doorslaer K, Schiffman M, Desalle R, Herrero R, Yu K, Wacholder S, Wang T, Burk RD.** 2011. Sequence imputation of HPV16 genomes for genetic association studies. *PloS one* **6**:e21375.
724. **Lei YJ, Makhaola K, Pittayakhajonwut D, Wood C, Angeletti PC.** 2011. Human papillomavirus 16 variants from Zambian women with normal pap smears. *Journal of medical virology* **83**:1230-1237.
725. **Sanchez GI, Kleter B, Gheit T, van Doorn LJ, de Koning MN, de Sanjose S, Alemany L, Bosch XF, Tommasino M, Munoz N, Quint WG.** 2011. Clinical evaluation of polymerase chain reaction reverse hybridization assay for detection and identification of human papillomavirus type 16 variants. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology* **51**:165-169.
726. **Quint KD, de Koning MN, van Doorn LJ, Quint WG, Pirog EC.** 2010. HPV genotyping and HPV16 variant analysis in glandular and squamous neoplastic lesions of the uterine cervix. *Gynecologic oncology* **117**:297-301.
727. **Zuna RE, Moore WE, Shanesmith RP, Dunn ST, Wang SS, Schiffman M, Blakey GL, Teel T.** 2009. Association of HPV16 E6 variants with diagnostic severity in cervical cytology samples of 354 women in a US population. *International journal of cancer. Journal international du cancer* **125**:2609-2613.

728. **Chopjitt P, Ekalaksananan T, Pientong C, Kongyingyoes B, Kleebkaow P, Charoensri N.** 2009. Prevalence of human papillomavirus type 16 and its variants in abnormal squamous cervical cells in Northeast Thailand. *International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases* **13**:212-219.
729. **Grodzki M, Besson G, Clavel C, Arslan A, Franceschi S, Birembaut P, Tommasino M, Zehbe I.** 2006. Increased risk for cervical disease progression of French women infected with the human papillomavirus type 16 E6-350G variant. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **15**:820-822.
730. **del Refugio Gonzalez-Losa M, Laviada Mier y Teran MA, Puerto-Solis M, Garcia-Carranca A.** 2004. Molecular variants of HPV type 16 E6 among Mexican women with LSIL and invasive cancer. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology* **29**:95-98.
731. **Berumen J, Ordonez RM, Lazcano E, Salmeron J, Galvan SC, Estrada RA, Yunes E, Garcia-Carranca A, Gonzalez-Lira G, Madrigal-de la Campa A.** 2001. Asian-American variants of human papillomavirus 16 and risk for cervical cancer: a case-control study. *Journal of the National Cancer Institute* **93**:1325-1330.
732. **Andersson S, Alemi M, Rylander E, Strand A, Larsson B, Sallstrom J, Wilander E.** 2000. Uneven distribution of HPV 16 E6 prototype and variant (L83V) oncoprotein in cervical neoplastic lesions. *British journal of cancer* **83**:307-310.
733. **Xi LF, Koutsky LA, Galloway DA, Kuypers J, Hughes JP, Wheeler CM, Holmes KK, Kiviat NB.** 1997. Genomic variation of human papillomavirus type 16 and risk for high grade cervical intraepithelial neoplasia. *Journal of the National Cancer Institute* **89**:796-802.
734. **Sun Z, Lu Z, Liu J, Wang G, Zhou W, Yang L, Liu C, Wang B, Ruan Q.** 2013. Genetic variations of E6 and long control region of human papillomavirus type 16 from patients with cervical lesion in Liaoning, China. *BMC cancer* **13**:459.
735. **Junes-Gill K, Sichero L, Maciag PC, Mello W, Noronha V, Villa LL.** 2008. Human papillomavirus type 16 variants in cervical cancer from an admixed population in Brazil. *Journal of medical virology* **80**:1639-1645.
736. **Chang YJ, Chen HC, Pan MH, Lee BH, You SL, Lin CY, Chou YC, Hsieh CY, Cheng YJ, Liaw KL, Hsing AW, Schiffman M, Chen CJ.** 2013. Intratypic variants of human papillomavirus type 16 and risk of cervical neoplasia in Taiwan. *Journal of medical virology* **85**:1567-1576.
737. **Pista A, Oliveira A, Barateiro A, Costa H, Verdasca N, Paixao MT.** 2007. Molecular variants of human papillomavirus type 16 and 18 and risk for cervical neoplasia in Portugal. *Journal of medical virology* **79**:1889-1897.
738. **Perez S, Cid A, Inarrea A, Pato M, Lamas MJ, Couso B, Gil M, Alvarez MJ, Rey S, Lopez-Miragaya I, Melon S, Ona M.** 2014. Prevalence of HPV 16 and HPV 18 lineages in Galicia, Spain. *PloS one* **9**:e104678.
739. **Lopez-Saavedra A, Gonzalez-Maya L, Ponce-de-Leon S, Garcia-Carranca A, Mohar A, Lizano M.** 2009. Functional implication of sequence variation in the long control region and E2 gene among human papillomavirus type 18 variants. *Archives of virology* **154**:747-754.

740. **Ferenczi A, Gyongyosi E, Szalmas A, Hernadi Z, Toth Z, Konya J, Veress G.** 2013. Sequence variation of human papillomavirus type 31 long control region: phylogenetic and functional implications. *Journal of medical virology* **85**:852-859.
741. **Gagnon S, Hankins C, Tremblay C, Pourreaux K, Forest P, Rouah F, Coutlee F.** 2005. Polymorphism of human papillomavirus type 31 isolates infecting the genital tract of HIV-seropositive and HIV-seronegative women at risk for HIV infection. *Journal of medical virology* **75**:213-221.
742. **Chagas BS, Batista MV, Crovella S, Gurgel AP, Silva Neto Jda C, Serra IG, Amaral CM, Balbino VQ, Muniz MT, Freitas AC.** 2013. Novel E6 and E7 oncogenes variants of human papillomavirus type 31 in Brazilian women with abnormal cervical cytology. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* **16**:13-18.
743. **Chagas BS, Batista MV, Guimaraes V, Balbino VQ, Crovella S, Freitas AC.** 2011. New variants of E6 and E7 oncogenes of human papillomavirus type 31 identified in Northeastern Brazil. *Gynecologic oncology* **123**:284-288.
744. **Xi LF, Schiffman M, Koutsky LA, Hughes JP, Winer RL, Mao C, Hulbert A, Lee SK, Shen Z, Kiviat NB.** 2014. Lineages of oncogenic human papillomavirus types other than type 16 and 18 and risk for cervical intraepithelial neoplasia. *Journal of the National Cancer Institute* **106**.
745. **Chen AA, Heideman DA, Boon D, Chen Z, Burk RD, De Vuyst H, Gheit T, Snijders PJ, Tommasino M, Franceschi S, Clifford GM.** 2014. Human papillomavirus 33 worldwide genetic variation and associated risk of cervical cancer. *Virology* **448**:356-362.
746. **Zhang C, Park JS, Grce M, Hibbitts S, Palefsky JM, Konno R, Smith-McCune KK, Giovannelli L, Chu TY, Picconi MA, Pina-Sanchez P, Settheetham-Ishida W, Coutlee F, De Marco F, Woo YL, Ho WC, Wong MC, Chirenje MZ, Magure T, Moscicki AB, Sabol I, Fiander AN, Chen Z, Chan MC, Cheung TH, Burk RD, Chan PK.** 2014. Geographical distribution and risk association of human papillomavirus genotype 52-variant lineages. *The Journal of infectious diseases* **210**:1600-1604.
747. **Sun Z, Lu Z, Liu J, Wang G, Zhou W, Yang L, Liu C, Ruan Q.** 2012. Genomic polymorphism of human papillomavirus type 52 in women from Northeast China. *International journal of molecular sciences* **13**:14962-14972.
748. **Aho J, Hankins C, Tremblay C, Forest P, Pourreaux K, Rouah F, Coutlee F.** 2004. Genomic polymorphism of human papillomavirus type 52 predisposes toward persistent infection in sexually active women. *The Journal of infectious diseases* **190**:46-52.
749. **Chang YJ, Chen HC, Lee BH, You SL, Lin CY, Pan MH, Chou YC, Hsieh CY, Chen YM, Cheng YJ, Chen CJ.** 2011. Unique variants of human papillomavirus genotypes 52 and 58 and risk of cervical neoplasia. *International journal of cancer. Journal international du cancer* **129**:965-973.
750. **Rodrigues-Coutlee C, Archambault J, Money D, Ramanakumar AV, Raboud J, Hankins C, Koushik A, Richardson H, Brassard P, Franco EL, Coutlee F.** 2013. Human papillomavirus type 56 polymorphism in Canadian women with and without cervical lesions. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology* **58**:660-665.
751. **Chan PK, Luk AC, Park JS, Smith-McCune KK, Palefsky JM, Konno R, Giovannelli L, Coutlee F, Hibbitts S, Chu TY, Settheetham-Ishida W, Picconi**

- MA, Ferrera A, De Marco F, Woo YL, Raiol T, Pina-Sanchez P, Cheung JL, Bae JH, Chirenje MZ, Magure T, Moscicki AB, Fiander AN, Di Stefano R, Cheung TH, Yu MM, Tsui SK, Pim D, Banks L.** 2011. Identification of human papillomavirus type 58 lineages and the distribution worldwide. *The Journal of infectious diseases* **203**:1565-1573.
752. **Wyant PS, Cerqueira DM, Moraes DS, Leite JP, Martins CR, de Macedo Brigido M, Raiol T.** 2011. Phylogeny and polymorphism in the long control region, E6, and L1 of human papillomavirus types 53, 56, and 66 in central Brazil. *International journal of gynecological cancer : official journal of the International Gynecological Cancer Society* **21**:222-229.
753. **Godinez JM, Heideman DA, Gheit T, Alemany L, Snijders PJ, Tommasino M, Meijer CJ, de Sanjose S, Bosch FX, Bravo IG.** 2013. Differential presence of Papillomavirus variants in cervical cancer: an analysis for HPV33, HPV45 and HPV58. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* **13**:96-104.
754. **Vrtacnik Bokal E, Kocjan BJ, Poljak M, Bogovac Z, Jancar N.** 2010. Genomic variants of human papillomavirus genotypes 16, 18, and 33 in women with cervical cancer in Slovenia. *The journal of obstetrics and gynaecology research* **36**:1204-1213.
755. **Raiol T, Wyant PS, de Amorim RM, Cerqueira DM, Milanezi N, Brigido Mde M, Sichero L, Martins CR.** 2009. Genetic variability and phylogeny of the high-risk HPV-31, -33, -35, -52, and -58 in central Brazil. *Journal of medical virology* **81**:685-692.
756. **Ceribelli M, Dolfini D, Merico D, Gatta R, Vigano AM, Pavesi G, Mantovani R.** 2008. The histone-like NF-Y is a bifunctional transcription factor. *Molecular and cellular biology* **28**:2047-2058.
757. **Dolfini D, Gatta R, Mantovani R.** 2012. NF-Y and the transcriptional activation of CCAAT promoters. *Critical reviews in biochemistry and molecular biology* **47**:29-49.
758. **Qyang Y, Luo X, Lu T, Ismail PM, Krylov D, Vinson C, Sawadogo M.** 1999. Cell-type-dependent activity of the ubiquitous transcription factor USF in cellular proliferation and transcriptional activation. *Molecular and cellular biology* **19**:1508-1517.
759. **Corre S, Galibert MD.** 2005. Upstream stimulating factors: highly versatile stress-responsive transcription factors. *Pigment cell research / sponsored by the European Society for Pigment Cell Research and the International Pigment Cell Society* **18**:337-348.
760. **Fontaine V, van der Meijden E, ter Schegget J.** 2001. Inhibition of human papillomavirus-16 long control region activity by interferon-gamma overcome by p300 overexpression. *Molecular carcinogenesis* **31**:27-36.
761. **Vo N, Goodman RH.** 2001. CREB-binding protein and p300 in transcriptional regulation. *The Journal of biological chemistry* **276**:13505-13508.
762. **Farley EK, Olson KM, Zhang W, Rokhsar DS, Levine MS.** 2016. Syntax compensates for poor binding sites to encode tissue specificity of developmental enhancers. *Proceedings of the National Academy of Sciences of the United States of America* **113**:6508-6513.
763. **Aguiar R, Grandy R, Meza D, Sepulveda H, Pihan P, van Wijnen AJ, Lian JB, Stein GS, Stein JL, Montecino M.** 2014. A functional N-terminal domain in

- C/EBPbeta-LAP\* is required for interacting with SWI/SNF and to repress Ric-8B gene transcription in osteoblasts. *Journal of cellular physiology* **229**:1521-1528.
764. **Kovacs KA, Steinmann M, Magistretti PJ, Halfon O, Cardinaux JR.** 2003. CCAAT/enhancer-binding protein family members recruit the coactivator CREB-binding protein and trigger its phosphorylation. *The Journal of biological chemistry* **278**:36959-36965.
765. **Agalioti T, Lomvardas S, Parekh B, Yie J, Maniatis T, Thanos D.** 2000. Ordered recruitment of chromatin modifying and general transcription factors to the IFN-beta promoter. *Cell* **103**:667-678.
766. **Edelstein LC, Lagos L, Simmons M, Tirumalai H, Gelinas C.** 2003. NF-kappa B-dependent assembly of an enhanceosome-like complex on the promoter region of apoptosis inhibitor Bfl-1/A1. *Molecular and cellular biology* **23**:2749-2761.
767. **Yie J, Senger K, Thanos D.** 1999. Mechanism by which the IFN-beta enhanceosome activates transcription. *Proceedings of the National Academy of Sciences of the United States of America* **96**:13108-13113.
768. **Ellwood KB, Yen YM, Johnson RC, Carey M.** 2000. Mechanism for specificity by HMG-1 in enhanceosome assembly. *Molecular and cellular biology* **20**:4359-4370.
769. **Wong AK, Shanahan F, Chen Y, Lian L, Ha P, Hendricks K, Ghaffari S, Iliev D, Penn B, Woodland AM, Smith R, Salada G, Carillo A, Laity K, Gupte J, Swedlund B, Tavtigian SV, Teng DH, Lees E.** 2000. BRG1, a component of the SWI-SNF complex, is mutated in multiple human tumor cell lines. *Cancer research* **60**:6171-6177.
770. **Alvarez J, Gagnon D, Coutlee F, Archambault J.** 2016. Construction of an Energy-Based Position Weight Matrix for C/EBPβ Dimers and its Use to Predict High-Affinity Sites in the Long Control Region of 175 HPV Types (unpublished).
771. **Isok-Paas H, Mannik 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. *Virology journal* **12**:59.
772. **Geimanen J, Isok-Paas H, Pipitch R, Salk K, Laos T, Orav M, Reinson T, Ustav M, Jr., Ustav M, Ustav E.** 2011. Development of a cellular assay system to study the genome replication of high- and low-risk mucosal and cutaneous human papillomaviruses. *Journal of virology* **85**:3315-3329.
773. **Gloss B, Chong T, Bernard HU.** 1989. Numerous nuclear proteins bind the long control region of human papillomavirus type 16: a subset of 6 of 23 DNase I-protected segments coincides with the location of the cell-type-specific enhancer. *Journal of virology* **63**:1142-1152.
774. **Fawcett TW, Martindale JL, Guyton KZ, Hai T, Holbrook NJ.** 1999. Complexes containing activating transcription factor (ATF)/cAMP-responsive-element-binding protein (CREB) interact with the CCAAT/enhancer-binding protein (C/EBP)-ATF composite site to regulate Gadd153 expression during the stress response. *The Biochemical journal* **339** ( Pt 1):135-141.
775. **Hsu W, Kerppola TK, Chen PL, Curran T, Chen-Kiang S.** 1994. Fos and Jun repress transcription activation by NF-IL6 through association at the basic zipper region. *Molecular and cellular biology* **14**:268-276.
776. **Catron KM, Brickwood JR, Shang C, Li Y, Shannon MF, Parks TP.** 1998. Cooperative binding and synergistic activation by RelA and C/EBPbeta on the intercellular adhesion molecule-1 promoter. *Cell growth & differentiation : the*

- molecular biology journal of the American Association for Cancer Research **9**:949-959.
777. **Shaywitz AJ, Greenberg ME.** 1999. CREB: a stimulus-induced transcription factor activated by a diverse array of extracellular signals. Annual review of biochemistry **68**:821-861.
778. **Brenner S, Prosch S, Schenke-Layland K, Riese U, Gausmann U, Platzer C.** 2003. cAMP-induced Interleukin-10 promoter activation depends on CCAAT/enhancer-binding protein expression and monocytic differentiation. The Journal of biological chemistry **278**:5597-5604.
779. **Mayr B, Montminy M.** 2001. Transcriptional regulation by the phosphorylation-dependent factor CREB. Nature reviews. Molecular cell biology **2**:599-609.
780. **Trautwein C, van der Geer P, Karin M, Hunter T, Chojkier M.** 1994. Protein kinase A and C site-specific phosphorylations of LAP (NF-IL6) modulate its binding affinity to DNA recognition elements. The Journal of clinical investigation **93**:2554-2561.
781. **Vallejo M, Gosse ME, Beckman W, Habener JF.** 1995. Impaired cyclic AMP-dependent phosphorylation renders CREB a repressor of C/EBP-induced transcription of the somatostatin gene in an insulinoma cell line. Molecular and cellular biology **15**:415-424.
782. **Rozenberg JM, Bhattacharya P, Chatterjee R, Glass K, Vinson C.** 2013. Combinatorial recruitment of CREB, C/EBPbeta and c-Jun determines activation of promoters upon keratinocyte differentiation. PloS one **8**:e78179.
783. **Crusius K, Rodriguez I, Alonso A.** 2000. The human papillomavirus type 16 E5 protein modulates ERK1/2 and p38 MAP kinase activation by an EGFR-independent process in stressed human keratinocytes. Virus genes **20**:65-69.
784. **Gu Z, Matlashewski G.** 1995. Effect of human papillomavirus type 16 oncogenes on MAP kinase activity. Journal of virology **69**:8051-8056.
785. **Branca M, Ciotti M, Santini D, Bonito LD, Benedetto A, Giorgi C, Paba P, Favalli C, Costa S, Agarossi A, Alderisio M, Syrjanen K.** 2004. Activation of the ERK/MAP kinase pathway in cervical intraepithelial neoplasia is related to grade of the lesion but not to high-risk human papillomavirus, virus clearance, or prognosis in cervical cancer. American journal of clinical pathology **122**:902-911.
786. **Zhang L, Wu J, Ling MT, Zhao L, Zhao KN.** 2015. The role of the PI3K/Akt/mTOR signalling pathway in human cancers induced by infection with human papillomaviruses. Molecular cancer **14**:87.
787. **Lui VW, Hedberg ML, Li H, Vangara BS, Pendleton K, Zeng Y, Lu Y, Zhang Q, Du Y, Gilbert BR, Freilino M, Sauerwein S, Peyser ND, Xiao D, Diergaarde B, Wang L, Chiosea S, Seethala R, Johnson JT, Kim S, Duvvuri U, Ferris RL, Romkes M, Nukui T, Kwok-Shing Ng P, Garraway LA, Hammerman PS, Mills GB, Grandis JR.** 2013. Frequent mutation of the PI3K pathway in head and neck cancer defines predictive biomarkers. Cancer discovery **3**:761-769.





## Annexe

Notre laboratoire s'est toujours intéressé à la protéine E1 du VPH en particulier à la régulation de son transport nucléo-cytoplasmique (voir section 1.2.4.3 de cette thèse). Nous avons montré que l'export de la protéine E1 est essentiel pour atténuer la réponse aux dommages à l'ADN induite par ATM qui est causée par l'accumulation nucléaire de l'hélicase virale (385, 392). Le transport de E1 dépend de la présence de motifs NLS et NES. Le recrutement d'un complexe Cdk2-cycline E/A se fait via un motif CBM pour la phosphorylation de résidus clés dans la région N-terminale de E1. Une inspection des séquences du N-terminal de E1 montre la conservation de ces motifs chez la vaste majorité des VPHs avec quelques exceptions. Ainsi, la protéine E1 du VPH33 ne contient pas de motifs NES et CBM apparents, alors que le NLS semble présent. L'absence des motifs NES et CBM impliqués dans l'export nucléaire indique que la protéine E1 du VPH33 n'est pas relocalisée au cytoplasme.

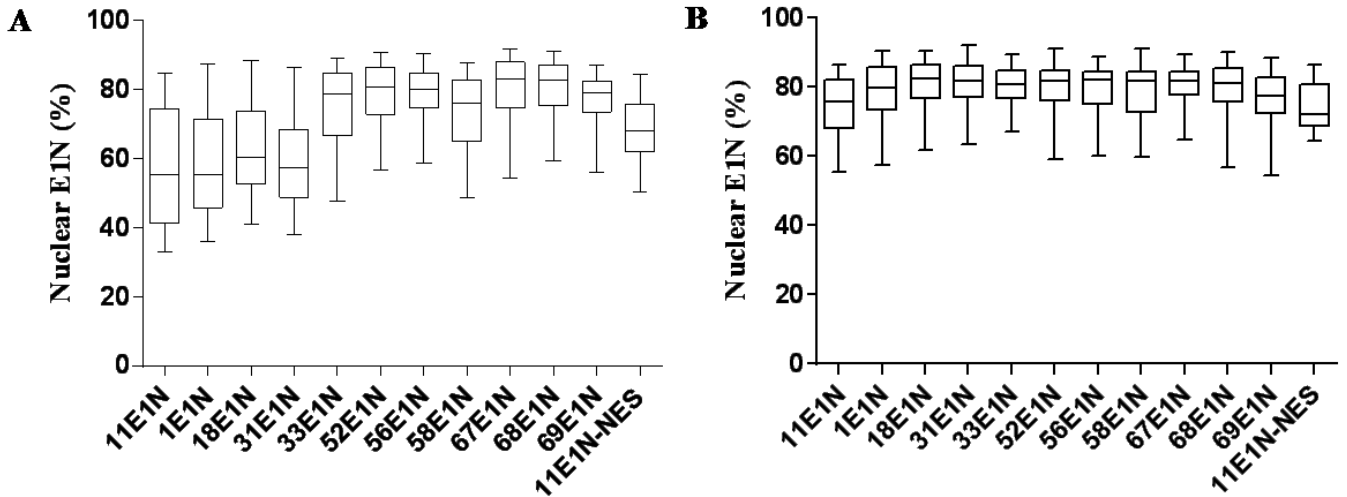
Les expériences présentées ici ont été réalisées dans le cadre d'un projet secondaire. Ces résultats, quoiqu'incomplets, indiquent que la protéine E1 du VPH33 n'est effectivement pas exportée hors du noyau et qu'elle est donc strictement nucléaire. Ceci pourrait donc signifier que la protéine E1, dite atypique, du VPH33 ne cause pas de dommages à l'ADN où que le virus utilise une autre stratégie pour éviter les effets délétères de l'accumulation au noyau.

J'ai produit toutes les figures à l'exception de la figure A3 (D) qui a été faite par mon collègue de travail, Dr David Gagnon. Dr Jacques Archambault a supervisé la réalisation des expériences.

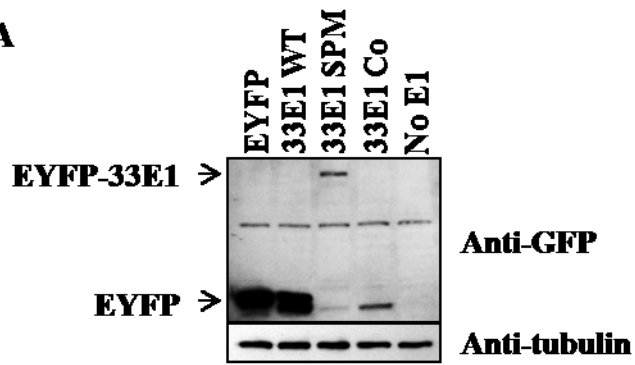
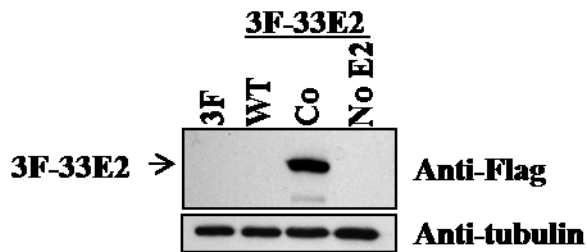
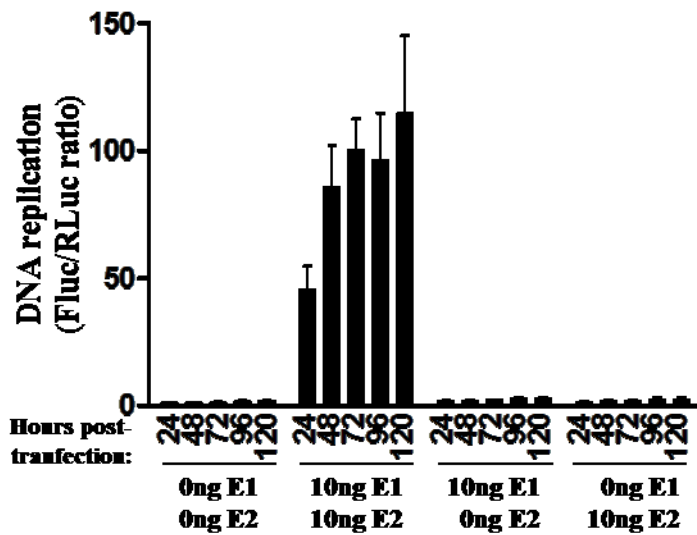
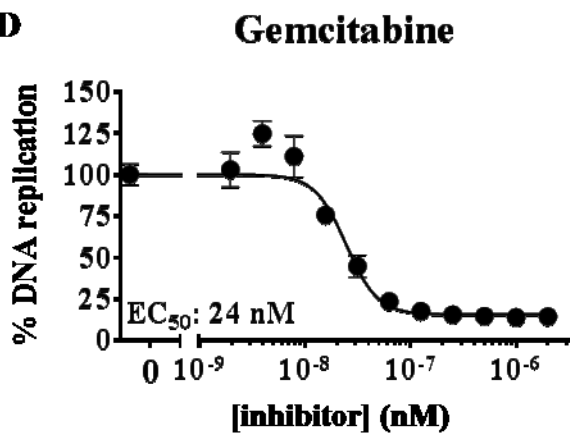
**Bi-partite NLS**

		<b>NES</b>										<b>CBM</b>			
1E1 (73)	L	KRK	LLY	SP	QAR	--	SADET--DIAS-	I	-P	RLETISI	TKQD	KKRY	R	RQL	-FS
56E1 (85)	L	KRK	YIA	--	--	SP	LRD-----	I	SN	-QQTVC-	REGV	---K	-	RRL	-IL
97E1 (85)	L	KRK	YGF	ES	KEN	SP	LGEQPG--VDYD	L	SP	RLQEISL	NSGQ	KKAK	-	RRL	-FQ
18E1 (85)	L	KRK	FAG	GS	TEN	SP	LGERLE--VDYE	L	SP	RLQEISL	NSGQ	KKAK	-	RRL	-FT
45E1 (85)	L	KRK	FAG	GS	KEN	SP	LGEQLS--VDYD	L	SP	RLQEISL	NSGH	KKAK	-	RRL	-FT
6E1 (82)	L	KRK	YLG	SP	YV-	SP	INPIAEA-VESE	I	SP	RIDAIKL	TRQP	KKVK	-	RRL	-FQ
11E1 (82)	L	KRK	YLG	SP	YV-	SP	ISNVANA-VESE	I	SP	RIDAIKL	TTQP	KKVK	-	RRL	-FE
35E1 (85)	L	KRK	YAS	--	--	SP	LSSVSL-CVMNN	I	SP	RLKAICI	ENKN	TAAK	-	RRL	-FE
16E1 (86)	L	KRK	YLG	--	--	SP	LSDISG-CVDRN	I	SP	RLKAICI	EKQS	RAAK	-	RRL	-FE
31E1 (85)	L	KRK	YVG	--	--	SP	LSDISS-CVCYN	I	SP	RLKAICI	ENNS	KTAK	-	RRL	-FE
68E1 (85)	L	KRK	YTD	SI	ES-	SP	LA-----	K	SP	-LQELSI	NVSS	TQAR	-	QPA	-YT
67E1 (85)	L	KRK	FTG	IQ	ACG	--	-GNSNG-	-	-	-IE-	NQAC	TAAK	-	RRA	Y--
69E1 (83)	L	KRK	LLG	SQ	N--	SP	LQDITN---QS	N	SQ	Q---STD	EVNN	LQAK	-	RRA	VDS
26E1 (83)	L	KRK	LLG	SQ	N--	SP	LQDITNQHQQS	D	SQ	Q---NTH	QVNN	SQAK	-	RRA	VDS
33E1 (85)	L	KRK	FAA	C-	--	SQ	SAAEDVVDRAAN	-	-P	CRTSINK	NKEC	TYRK	-	RKI	-DE
58E1 (85)	L	KRK	FAA	CS	E--	--	SAVEDCVDRAN	-	-V	CVSWKYK	NKEC	THRK	-	RKI	-IE
52E1 (85)	V	KRK	FTS	--	--	SP	ESAGQDGVKHHG	-	SP	RAKHICV	NTEC	VLPK	-	RKP	CH-

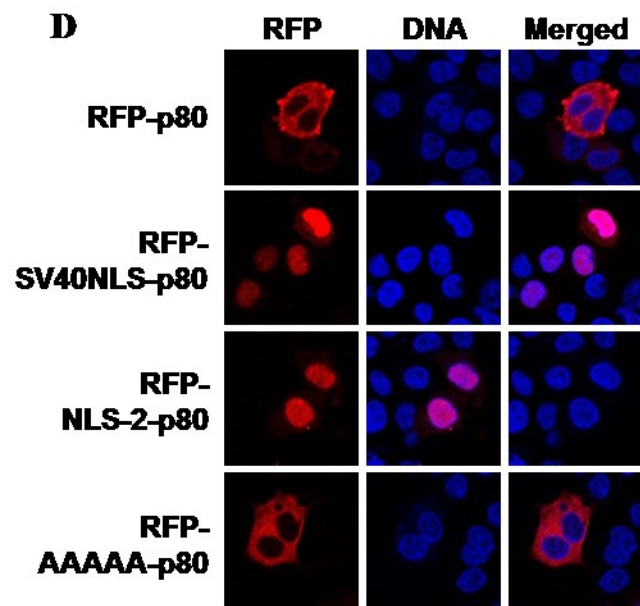
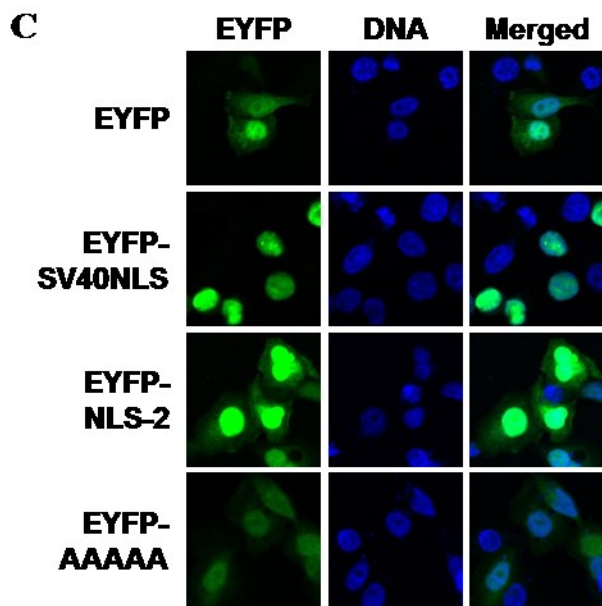
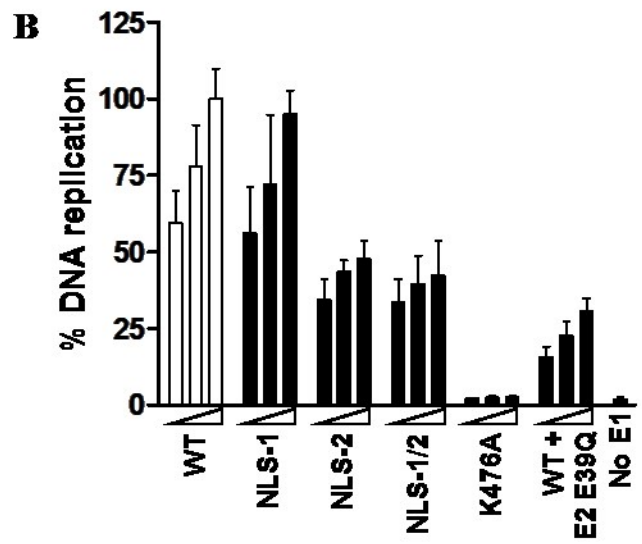
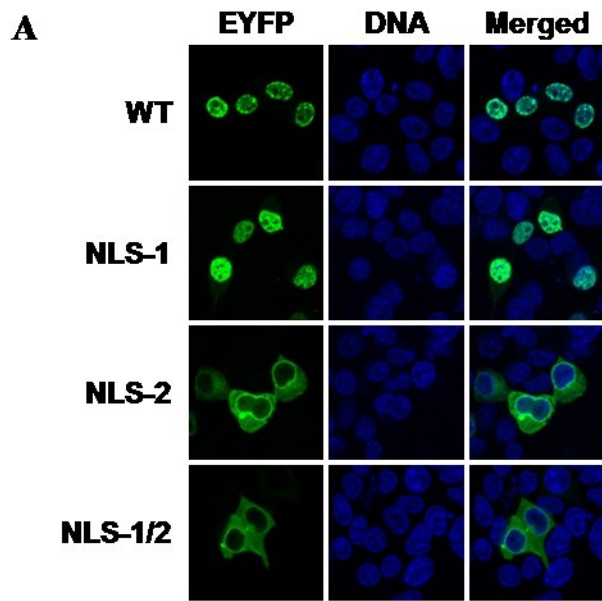
**Figure A1 Motifs de localisation cellulaire dans la région N-terminale de la protéine E1.** Alignement de la séquence en acides aminés du module de transport de différents types de VPHs. La position du premier acide aminé de la séquence est indiquée entre parenthèses. La position du signal de localisation nucléaire (NLS), du signal d'export nucléaire (NES) et du motif de liaison aux cyclines (CBM) est indiquée par les boîtes vertes, bleues et roses, respectivement. Les sites potentiels ou connus de phosphorylation de Cdk2 sont colorés en jaune. Notez l'absence de ces motifs chez certains types de VPH (en rouge).



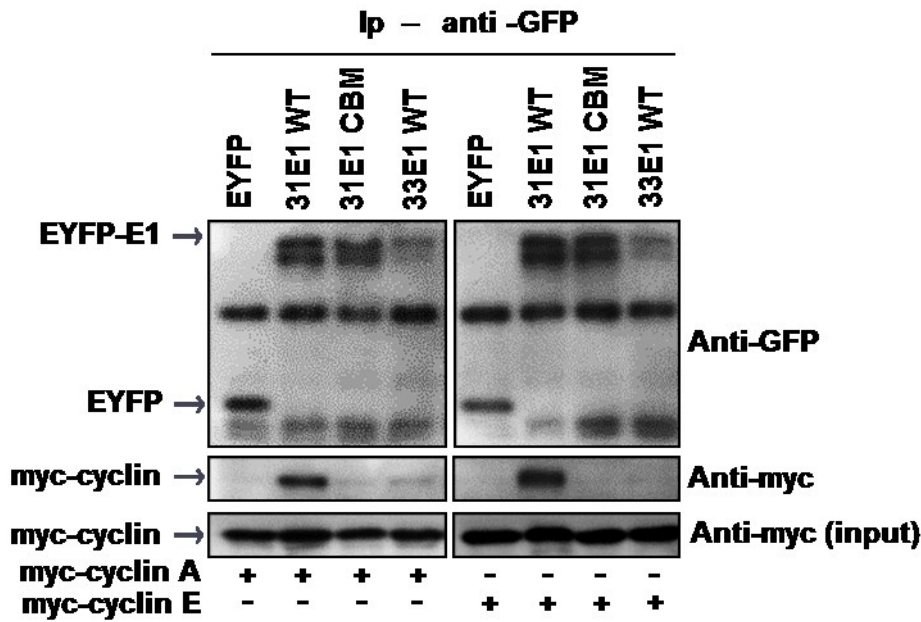
**Figure A2** La protéine E1 des VPHs atypiques, dont celle du VPH33, est **majoritairement nucléaire**. Localisation nucléaire de la région N-terminale de E1 chez différents types de VPH. Les plasmides codant une fusion GFP des régions en N-terminal de E1 des types de VPH indiqués ont été produits tels que décrits dans (229). Ces plasmides (66 ng) ont été transfectés dans des cellules HeLa (5000 cellules) préalablement cultivés dans des plaques noires de 96 puits à fond transparent. Les cellules non traitées (A) et traitées à la leptomycine B (20 nM) (B) ont été fixées et colorées au dapi suite à leur incubation avec un anticorps Alexa Fluor 633 couplé à la GFP (Life technologies). L'intensité de la fluorescence émise par la GFP a été mesurée dans un essai de criblage d'imagerie à haut débit de type HCS (*high content screening*). L'intensité nucléaire de la GFP est rapportée en pourcentage (%) de la fluorescence totale mesurée dans le noyau et le cytoplasme des cellules transfectées. Le mutant 11E1N-NES du VPH de type 11 est utilisé comme un contrôle nucléaire.

**A****B****C****D**

**Figure A3. Développement et validation d'un essai de réplication de l'ADN par les protéines E1 et E2 du VPH33.** (A) Analyse d'immunobuvardage de type Western de cellules C33A transfectées avec des vecteurs d'expression codant la protéine E1 pleine longueur de type sauvage (WT), d'un mutant d'épissage (SPM) et d'une séquence nucléotidique codante optimisée (Co). Le vecteur d'expression pour EYFP a été ajouté comme contrôle. Les protéines E1 et la tubuline ont été détectées en utilisant un anticorps contre la GFP et contre la tubuline, respectivement. (B) Analyse d'immunobuvardage de type Western de cellules C33A transfectées avec des vecteurs d'expression codant la protéine E2 pleine longueur de type sauvage (WT) et d'une séquence nucléotidique codante optimisée (Co). Le vecteur d'expression pour 3F (triple flag) a été ajouté comme contrôle. Les protéines E2 et la tubuline ont été détectées en utilisant un anticorps contre Flag et contre la tubuline, respectivement. (C) Activité répliquative de l'ADN par E1. Les protéines E1SPM et E2 Co, selon les quantités indiquées dans la figure, ont été transfectées dans des cellules C33A avec 2.5 ng d'un plasmide contenant l'ori minimale du VPH33 en cis du gène de la Fluc (pFLORI33) et de 0.5 de pRL dont la Rluc est sous le contrôle du promoteur CMV. La réplication de l'ADN a été mesurée 24, 48, 72, 96 et 120h post-transfection et est rapportée comme le ratio Fluc/Rluc. Le protocole a été adapté de (231). (D) Activité de réplication telle que mesurée dans des cellules transfectées avec 10 ng de 33E1 et 33E2 et traitées avec des concentrations croissantes en multiple de deux de gemcitabine (1.95-2000 nM) ou traitée avec du DMSO comme contrôle négatif. Les activités de réplication sont rapportées en pourcentage (%) du ratio Fluc/Rluc obtenu en présence de DMSO seulement.

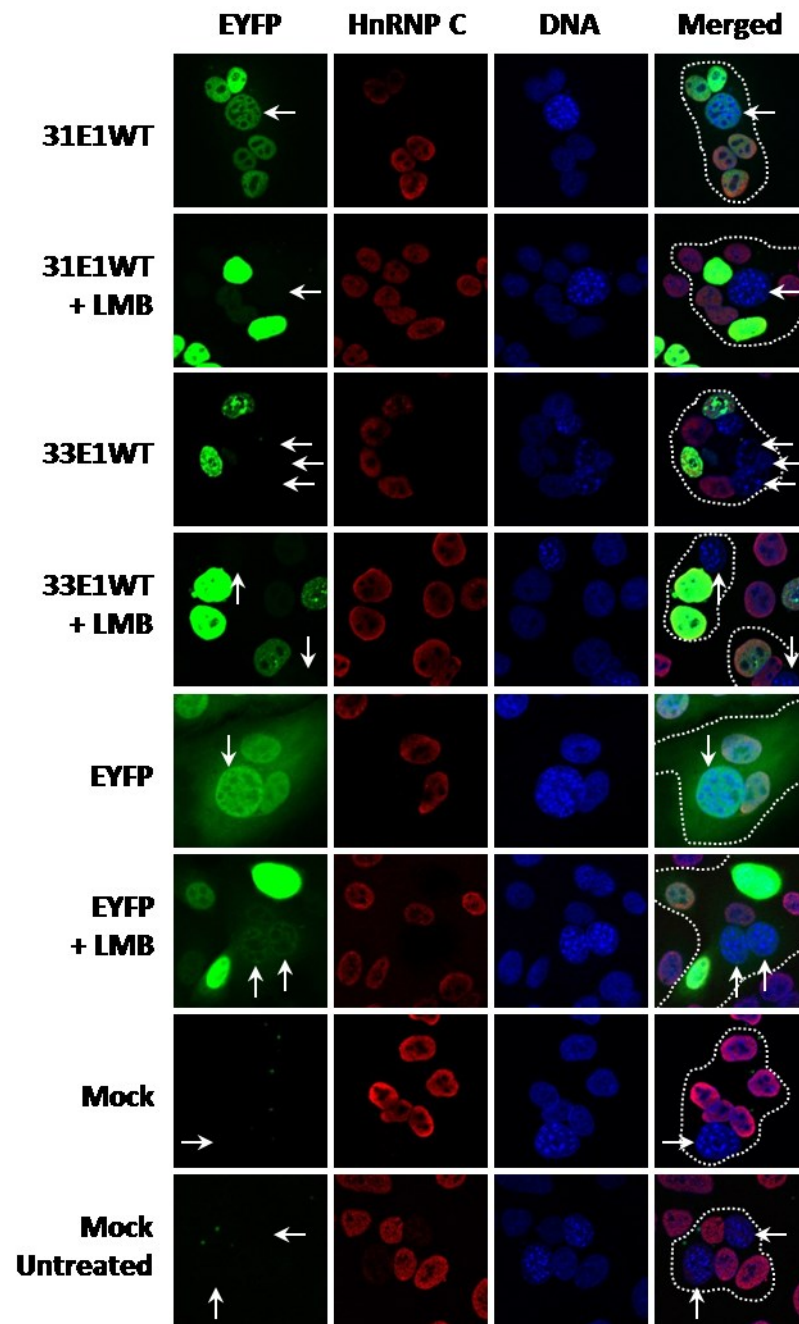


**Figure A4 Le transport nucléaire de la protéine E1 du VPH33 implique un motif NLS de type monopartite.** (A) Localisation intracellulaire des protéines E1 WT et de mutants NLS. Des alanines dans la première (NLS-1 : KRK), dans la deuxième partie du NLS (NLS-2 : RKRKI) ont été introduits individuellement ou en combinaison (NLS-1/2). Les cellules C33A préalablement transfectées avec ces protéines, et dont les noyaux ont été colorés au dapi, ont été visualisées par microscopie confocale à fluorescence. (B) Activité de réplication des mutants NLS dans des cellules C33A. Chaque mutant a été testé en trois quantités différentes (2.5, 5 et 10 ng). Les cellules non-transfectées (no E1) et transfectées avec un mutant déféctueux pour l'activité ATPase (K476A) ont été utilisées comme contrôle négatif. L'activité de réplication de la protéine E1 WT dans le contexte d'une protéine E2 mutante pour son interaction avec E1 (E39Q) a également été utilisée comme contrôle. Les activités de réplication sont rapportées en pourcentage (%) du ratio Fluc/Rluc obtenu pour 10 ng de E1 WT. Relocalisation nucléaire de la protéine EYFP (C) et de p80/UAF1 (D) par l'introduction de différents motifs NLS. Le NLS de SV40 (PKKKRKV) et le NLS de type monopartite du VPF33 (RKRKI) ou son mutant (AAAAA) ont été introduits dans un vecteur d'expression pour EYFP et dans une fusion RFP-p80 cytoplasmique. La localisation de ces protéines a été visualisée telle que décrit plus haut.



**Figure A5.** La protéine E1 du VPH33 n'interagit pas avec le complexe Cdk2/cyclines E/A *in vivo*. La protéine E1 WT ou un mutant CBM (RKI à AAA) a été co-transfectée avec la cycline E (cyc E) ou la cycline A (cyc A) fusionnée à un épitope 6xmyc dans des cellules C33A. La protéine 31E1 WT et son mutant CBM (RRL à AAA) ont été transfectées comme contrôle positif de l'interaction. Les protéines co-immunoprécipitées ont été séparées par SDS-PAGE et analysées par immunobuvardage de type Western avec un anticorps anti-GFP et un anticorps anti-myc telles qu'indiquées dans la figure. Le panneau du bas présente les niveaux d'expression de la cycline avant immunoprécipitation.





**Figure A6. La protéine E1 du VPH33 n'est pas exportée hors du noyau.** Des cellules HeLa ont été préalablement transfectées avec la protéine E1 WT du VPH33 ou VPH31 (contrôle positif de la relocalisation cellulaire) ou EYFP seulement telles qu'indiquées dans la figure. Ces cellules prétraitées avec de la cycloheximide ont été fusionnées avec des cellules de souris 3T3. Les cellules ont par la suite été incubées 2h, pour permettre le transport nucléo-cytoplasmique des protéines transfectées, et ont été traitées avec de la leptomycine B quand indiquée. La protéine HnRNP C humaine (*heterogeneous nuclear ribonucleoprotein C*) a été colorée en utilisant un anticorps contre HnRNP C et d'un anticorps Alexa Fluor 633. Cette protéine a été utilisée comme un contrôle que la synthèse de protéines *de novo* a été effectivement inhibée. Les cellules de souris sont facilement identifiées par l'allure ponctuée de leurs noyaux lorsque colorés au dapi. Les lignes blanches pointillées montrent la membrane plasmique des hétérokaryons. Les flèches blanches indiquent les noyaux des cellules de souris. La détection de la protéine E1 (EYFP) dans les cellules de souris, qui n'ont pas été transfectées au préalable, est un indicatif que E1 a été transportée et donc exportée du noyau des cellules HeLa. Le protocole a été adapté de (389).