

Université de Montréal

Étude de la persistance et du polymorphisme génétique des papillomavirus humains de  
type 31, 33 et 35

par

Simon Gagnon

Département de microbiologie et immunologie

Faculté de médecine

Mémoire présenté à la Faculté des études supérieures  
en vue de l'obtention du grade de  
Maître ès sciences (M.Sc.)  
en Microbiologie/ Immunologie

Août, 2003

© Simon Gagnon

Université de Montréal



W

4

U58

2004

v. 085

**Direction des bibliothèques**

**AVIS**

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

**NOTICE**

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal  
Faculté des études supérieures

Ce mémoire intitulé :

Étude de la persistance et du polymorphisme génétique des papillomavirus humain de  
type 31, 33 et 35

présenté par :  
Simon Gagnon

a été évalué par un jury composé des personnes suivantes :

Louise Labrecque  
François Coutlée  
Éric Cohen

Mémoire accepté le :

## RÉSUMÉ

Les papillomavirus humains (VPH) oncogènes sont maintenant reconnus comme les agents étiologiques du cancer du col utérin. Le rôle du polymorphisme génétique dans le développement de lésions précancéreuses a été clairement défini pour VHP-16. Dans une étude antérieure, nous avons aussi démontré que les variants non-prototypiques du VPH-52 étaient associés avec une infection persistante, un facteur de risque reconnu pour le développement de lésions malignes. Aucune recherche similaire sur le polymorphisme et la persistance des types 31, 33 et 35 n'a encore été faite. Nous disposons d'une cohorte de 1055 femmes dépistées pour le VIH, desquelles nous avons obtenu des lavages cervico-vaginaux à des intervalles de 6 mois. Un isolat de chaque femme positive pour les types 31, 33 ou 35 a été analysé par PCR-séquençage du LCR (Long Control Region), de E6 et de E7. Les variants étaient ensuite classifiés en «prototype-like» et «non-prototype». Pour le VPH-31, nous avons défini respectivement 18, 10 et 8 variants pour le LCR, E6 et E7, en comparaison avec 11, 6 et 6 variants pour le VPH-33 pour le LCR, E6 et E7. Pour VPH-35 8, 8 et 7 variants ont pu être définis pour le LCR, E6 et E7 respectivement. Pour tous les types, le polymorphisme était plus important dans le LCR que dans les gènes codant pour les oncoprotéines E6 et E7. De plus, le polymorphisme du VPH-31 était relié à l'origine ethnique. Cependant, pour ce même type, l'infection par des variants non-prototypiques n'était pas reliée à la persistance. En contrepartie, pour les types 33 et 35, une infection par un variant non-prototypique au niveau de E7 était associée avec la persistance avec un risque relatif de 2.0.

**Mots-clés : virus du papillome humain, polymorphisme, persistance, PCR.**

## ABSTRACT

High-risk human papillomaviruses are the etiologic agents of cervical cancer. It is now recognized that the genetic polymorphism of HPV-16 is associated with persistence of infection, an important predictor for precancerous and cancerous cervical lesions. Also, we have previously shown that non-prototypic HPV-52 variants were associated with persistent infection. No similar polymorphism study was done on HPV-31, 33 and 35. We disposed of cervicovaginal lavages obtained at 6-month intervals from 1055 women, screened for HPV and HIV. One HPV-31, 33 and 35 isolate from each women was tested for molecular variant analysis by PCR-sequencing of the LCR (Long Control Region), E6 and E7 genes. Isolates were classified into prototype-like and non-prototype variants. For HPV-31, we defined 18, 10 and 8 variants for the LCR, E6 and E7, respectively. For the same genomic regions of HPV 33 and 35, we defined 11, 6, 6 and 8, 8, 7 variants, respectively. In all cases, LCR was more polymorphic than E6 or E7. HPV-31 polymorphism was associated with the ethnic origin of participants. However, persistence was not associated with non-prototypic HPV-31 variants infection. Interestingly, for HPV-33 and 35, infection with an E7 non-prototypic isolate was associated with HPV persistence (relative risk (RR) of 2.0).

**Keywords : human papillomavirus, polymorphism, persistence, PCR.**

## PRÉSENTATIONS À DES CONGRÈS SCIENTIFIQUES

Cette étude a été présentée dans le cadre de trois congrès scientifiques :

1) Présentation orale

Gagnon S, Hankins C, The Canadian Women's HIV Study Group, Coutlée F.

«Polymorphisme génétique et persistance des papillomavirus humains des types 31 et 33».

Congrès du CHUM, 23 janvier 2003

2) Par affiche

Gagnon S, Hankins C, The Canadian Women's HIV Study Group, Coutlée F.

«Genetic polymorphism and persistence of human papillomavirus type 31 and 33 in HIV-infected women».

Canadian Association on HIV Research, XII th Annual Conference, Halifax, 12-15 avril 2003.

3) Par affiche

Gagnon S, , Hankins C, Tremblay C, Pourreaux K, Forest P, The Canadian Women's HIV Study Group, Coutlée F.

«Polymorphisms of HPV-33 and HPV-35 are associated with persistence of HPV infections».

The CACMID Annual Meeting, Montréal, 2-5 novembre 2003.

# TABLE DES MATIÈRES

<b>RÉSUMÉ</b> .....	<b>iii</b>
<b>ABSTRACT</b> .....	<b>iv</b>
Présentations à des congrès scientifiques .....	v
<b>TABLE DES MATIÈRES</b> .....	<b>vi</b>
<b>LISTE DES FIGURES</b> .....	<b>viii</b>
<u>Revue de littérature</u> .....	viii
<u>Article 1</u> .....	ix
<u>Article 2</u> .....	ix
<b>LISTE DES TABLES</b> .....	<b>xi</b>
<u>Article 1</u> .....	xi
<u>Article 2</u> .....	xi
<b>LISTE DES ABBRÉVIATIONS</b> .....	<b>xii</b>
<i>Dédicace</i> .....	<i>xv</i>
<i>Remerciements</i> .....	<i>xvi</i>
<b><u>REVUE DE LITTÉRATURE</u></b> .....	<b>1</b>
<i>Introduction</i> .....	<i>2</i>
<b>1. <i>Le virus du papillome humain</i></b> .....	<b>4</b>
1.1 Historique.....	4
1.2 Classification .....	5
1.3 Structure du virus et organisation génomique .....	6
1.3.1 Structure du virion.....	6
1.3.2 Organisation génomique .....	7
1.4 Transcription et réplication chez les papillomavirus .....	8
1.4.1 Fonction des protéines précoces .....	8
1.4.2 Relation virus-kératinocytes.....	10
1.4.3 Réplication.....	12
1.5 Caractéristiques biologiques des papillomavirus .....	13
1.5.1 Spécificités tissulaire et d'espèce .....	13
1.5.2 Prolifération et transformation cellulaire .....	14
<b>2. <i>Papillomavirus humains génitaux et cancer du col utérin</i></b> .....	<b>15</b>
2.1 Association épidémiologique.....	15
2.2 Épidémiologie du cancer du col de l'utérus.....	16
2.3 Pathogenèse du cancer du col utérin .....	17
2.4 Mécanismes d'induction du cancer.....	18



2.5 La persistance.....	24
<b>3. VPH et VIH .....</b>	<b>26</b>
3.1 Interactions indirectes.....	26
3.2 Interactions directes.....	28
<b>4. Les variants chez les papillomavirus.....</b>	<b>31</b>
4.1 Phylogénie des VPHs.....	31
4.2 Variants et facteur de risque.....	32
4.3 Les variants du LCR.....	34
4.4 Les variants de E6 et E7.....	35
4.5 Autres variants.....	38
<i>Objectifs et hypothèses.....</i>	<i>40</i>
<b><u>ARTICLES</u> .....</b>	<b>42</b>
<b><u>Article 1</u></b> .....	<b>44</b>
<b><u>Article 2</u></b> .....	<b>87</b>
<b><u>DISCUSSION</u>.....</b>	<b>132</b>
<i>Conclusion.....</i>	<i>143</i>
<b>RÉFÉRENCES .....</b>	<b>144</b>

## LISTE DES FIGURES

### Revue de littérature

#### Figure 1 :

*Organisation structurale de la capsid de VPH.*

Source : [http://bilbo.bio.purdue.edu/~baker/projects/papova/papilloma/hpv\\_10810.html](http://bilbo.bio.purdue.edu/~baker/projects/papova/papilloma/hpv_10810.html)

#### Figure 2 :

*Organisation génomique du VPH-31.*

Source : [http://bugs.mimnet.northwestern.edu/laimins\\_lab/labsite/Research.html](http://bugs.mimnet.northwestern.edu/laimins_lab/labsite/Research.html)

#### Figure 3 :

*Représentation schématique d'un papillome.*

Source : <http://gsbs.utmb.edu/microbook/ch066.htm>

#### Figure 4 :

*Zone de transformation.*

Source : Aho, 2002

#### Figure 5 :

*Représentation schématique des lésions précurseurs du cancer du col utérin.*

Source : Palefsky, 1999

#### Figure 6 :

*Représentation schématique de l'inhibition du suppresseur de tumeur p53 par la protéine transformante E6.*

Source : <http://www.baclesse.fr/cours/fondamentale/7-carcino-virale/Virale-6.htm>

Figure 7 :

*Représentation schématique de l'inhibition de pRb par l'oncoprotéine E7.*

Source : <http://www.baclesse.fr/cours/fondamentale/7-carcino-virale/Virale-6.htm>

Figure 8 :

*Intégration du génôme des papillomavirus.*

## **Article 1**

Figure 1 :

Sequence variations in the 3'end of LCR of 79 HPV-31 isolates.

Figure 2 :

2a) Sequence variations in E6 of 79 HPV-31 isolates. 2b) Sequence variations in the E7 of 79 HPV-31 isolates.

## **Article 2**

Figure 1 :

Sequence variations in the 3'end of the LCR of HPV-33.

Figure 2 :

2a) Sequence variations in the E6 open reading frame of HPV-33. 2b) Sequence variations in the E7 open reading frames of HPV-33.

Figure 3 :

Sequence variations in the 3'end of the LCR of HPV-35.

Figure 4 :

4a) Sequence variations in the E6 open reading frame of HPV-35. 4b) Sequence variations in the E7 open reading frames of HPV-35.

# **LISTE DES TABLES**

*tableaux*

## **Article 1**

### **Table 1 :**

Association between HPV-31 variants, ethnicity and HIV infection.

### **Table 2 :**

HPV-31 variant classification in function of genomic region analyzed for 41 isolates.

## **Article 2**

### **Table 1 :**

Sequence variations in the 3'end of the LCR of HPV-33.

### **Table 2 :**

2a) Sequence variations in the E6 open reading frame of HPV-33. 2b) Sequence variations in the E7 open reading frames of HPV-33.

### **Table 3 :**

Sequence variations in the 3'end of the LCR of HPV-35.

### **Table 4 :**

4a) Sequence variations in the E6 open reading frame of HPV-35. 4b) Sequence variations in the E7 open reading frames of HPV-35.

## LISTE DES ABBRÉVIATIONS

ADN : Acide désoxyribonucléique

AP-1 : «Activator protein-1»

ARN : Acide ribonucléique

ARNm : Acide ribonucléique messenger

BPV : «Bovine papillomavirus»

CBP : «CRE binding factor»

C/EBP $\beta$  : «CCAAT/enhancer binding protein beta»

CRE : «cAMP response element»

DPV : «Deer papillomavirus»

E : Européen

EGF: «Epidermal growth factor»

EGFR : «Epidermal growth factor receptor»

HSIL : «High-grade intraepithelial lesion»

hERT : «Human telomerase reverse transcriptase»

KRF-1 : «Keratinocyte-specific transcription factor-1»

LCR : «Long control region»

LSIL : «Low-grade intraepithelial lesion»

NE : Non-européen

NF-1 : «Nuclear factor-1»

Nm : Nanomètres

NPL : «non-prototype-like»

Nt : Nucléotide

Oct-1 : «Octamer binding factor-1»

ORF : «Open reading frame»

Pb : Paire de bases

PL : «prototype-like»

RRE : «Rev-responsive element»

SIDA : Syndrome de l'immunodéficience acquise

SIL : «Squamous intraepithelial lesion»

Sp1 : «Promoter-specific transcription factor-1»

TEF-1 : «Transcriptional enhancer factor-1»

TRAIL : «Tumor necrosis factor-related apoptosis-inducing ligand»

URR : Upstream regulatory region

VIH : Virus de l'immunodéficience humaine

VLP : «Virus-like particles»

VPH : Virus du papillome humain

YY1 : «Yin Yang-1»



## DÉDICACE

*Je dédie ce mémoire à Val, mon amour, pour toute la joie qu'elle m'apporte, à mon frère, pour tout l'intérêt qu'il a manifesté face à mes travaux et enfin, à mes parents, pour leur incroyable support et pour la fierté qu'ils ont toujours démontrée à mon égard.*

## REMERCIEMENTS

Mes remerciements vont d'abord à mon directeur de recherche, François Coutlée, pour sa disponibilité tout au court de ma maîtrise, et ce malgré un horaire chargé, pour son appui pédagogique hors pair et pour ses conseils éclairés. Mes remerciements vont aussi aux assistants de recherche, Pierre Forest et Hélène Voyer, qui ont grandement contribué à ma formation en laboratoire ainsi qu'à France Dion et Julie Lacaille pour leur support technique.

Enfin, je me dois de mentionner mes collègues étudiants : Joséphine Aho, Jonas Lefevre, Julie Fontaine, Igal Ifergan et Julie Lajoie pour les discussions toujours hautement pertinentes et sensées que nous avons générées ensemble.

**REVUE DE LITTÉRATURE**

## INTRODUCTION

Les papillomavirus humains regroupent plus d'une centaine de génotypes différents, dont 40 infectent les muqueuses génitales. Ceux-ci ont une grande importance du fait qu'ils causent les condylomes, une des maladies transmises sexuellement les plus fréquentes (Santé Canada, 2003), mais surtout de par le fait que certains de ces génotypes constituent les agents étiologiques du cancer du col de l'utérus [10]. Ce dernier est le troisième cancer en importance chez la femme mondialement et prend une importance plus grande encore dans les pays en voie de développement, où les tests de dépistage cytologique sont l'exception plutôt que la règle. Il est à noter qu'une infection par un VPH oncogène n'entraînera pas obligatoirement un cancer. Dans les faits, seule une proportion minime de ces infections résulteront en un cancer, tout dépendant de la présence d'une conjonction de facteurs.

Les facteurs auxquels nous nous intéresserons particulièrement sont le génotype du VPH et les variants intratypiques. Le génotype est un facteur primordial dans le développement du cancer, étant donné que seuls les types dit «à haut risque» causent cette pathologie et que pour un type à haut risque donné, par exemple VPH-16, certains variants semblent plus ou moins agressifs que d'autres. Autre facteur d'importance à aborder: la co-infection avec le VIH. En effet, l'infection au VIH prédispose à l'infection par les VPHs et au développement de lésions précancéreuses, ce phénomène étant directement relié au niveau d'immuno-suppression de l'individu.

La principale particularité de cette étude repose sur le fait que le polymorphisme des types à l'étude, soit les types 31, 33 et 35, n'a pas encore été caractérisé jusqu'à présent. En effet, il n'y a que peu de publications sur ces génotypes, et ce malgré le fait qu'ils sont des types à haut risque.

Mais avant de s'attaquer à la méthodologie et aux résultats concrets apportés par cette étude, penchons-nous d'abord sur l'état actuel des connaissances générales sur le VPH, c'est-à-dire sa classification, sa structure, son organisation génomique, sa

réplication et ses caractéristiques biologiques. Ensuite, nous aborderons la pathogenèse de l'infection par les VPHs oncogènes en ciblant particulièrement le développement des lésions malignes et la persistance virale, facteur jouant un rôle central dans l'apparition de lésions précancéreuses. Puis nous mettrons en lumière la co-infection VPH-VIH de façon à discuter de leurs interactions multiples. Enfin, nous concluerons avec le thème principal de notre étude, en l'occurrence les variants phylogénétiques, que nous décrirons en nous basant notamment sur un type déjà très bien caractérisé, le VPH-16.

# 1. LE VIRUS DU PAPILLOME HUMAIN

## *1.1 Historique*

Il y a de cela plusieurs siècles, les médecins suspectaient les verrues d'être de nature infectieuse. Ce n'est qu'en 1907 que Ciuffo démontra que les verrues pouvaient être transmises de personne à personne par l'application d'extraits verruqueux exempts de cellules, indiquant ainsi une implication virale dans cette pathogénie (Fields, p. 2045). Le «cotton-tail rabbit papillomavirus» fut le premier de la grande famille des papillomavirus à être décrit, mais fut aussi la première démonstration d'un potentiel oncogénique d'origine virale chez les mammifères (Fields, p. 2077), le virus induisant la conversion rapide des papillomes en carcinome chez le lapin domestique [189]. Dans les années soixante, l'opinion générale voulait qu'il n'y ait qu'un type de papillomavirus et que la morphologie et le comportement caractéristique d'une verrue à un site corporel particulier était dû uniquement à la nature spécifique de l'épithélium à ce site [136]. Dans les années soixante-dix, la recherche sur les papillomavirus évolua rapidement, principalement en conséquence de nouveaux développements technologiques tels que l'utilisation des endonucléases de restriction et des techniques de clonage, mais aussi et surtout en raison de l'émergence de nouvelles idées par rapport au rôle des papillomavirus dans le développement de tumeurs malignes chez l'homme [189]. Cependant, ce n'est qu'en 1976-77 que la pluralité des papillomavirus humains fut reconnue [48]. Ainsi, au début des années 80, le champs de recherche explosa virtuellement avec la détection de nouveaux types provenant de biopsies de patientes avec tumeur au niveau du col de l'utérus [31]. La présence régulière de ces types (notamment les VPH-16 et VPH-18) dans les cancers du col utérin se solda par un intérêt croissant pour ce groupe de virus et par l'isolation et la découverte de nombreux autres types. À ce jour, plus de 70 génotypes du VPH ont été entièrement caractérisés [146].

## 1.2 Classification

Initialement, les papillomavirus furent regroupés avec les polyomavirus pour former la famille des *Papovavirus*. Les propriétés communes propres à ces derniers sont notamment une petite taille, l'absence d'enveloppe, une capsidie icosaédrique, une molécule d'ADN double-brin circulaire et une multiplication nucléaire [60]. Cependant, des études portant sur la biologie et l'organisation génomique de ces deux groupes de virus mirent en évidence des différences fondamentales entre eux et il fut décidé d'en faire deux familles distinctes. Les papillomavirus se distinguent des polyomavirus par un génôme plus gros (de l'ordre de 8000 pb versus 5000 pb) et par la grosseur du virion, les papillomavirus étant légèrement plus volumineux (55 nm versus 40 nm) [60].

Les papillomavirus se retrouvent partout dans la nature et ont été caractérisés chez nombre d'espèces, l'agent infectieux étant sous-classifié et nommé selon l'espèce infectée, le bovin (BPV) et le renne (DPV) en sont d'ailleurs des exemples. Fait important, aucune infection productive inter-espèce n'a jamais été rapportée, révélant ainsi une haute spécificité d'espèce chez cette famille virale [145].

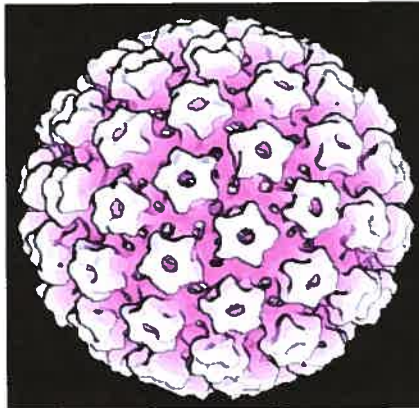
Les papillomavirus d'une espèce donnée sont sous-classifiés en types selon l'homologie de leur séquence génomique. Ainsi, il y a maintenant plus de 70 types reconnus chez l'homme, que l'on distingue par comparaison des séquences nucléotidiques de régions spécifiques du génôme, soit les régions E6, E7 et L1. En fait, selon le «Papillomavirus Nomenclature Committee», deux types différents auront une homologie de moins de 90% au niveau de ces cadres de lecture. Une homologie se situant entre 90% et 98% définirait des sous-types et une homologie au-dessus de 98% définirait des variants.

Il est à noter que les virologues ont élaboré des méthodes de classification non-taxonomique alternatives se basant sur certains facteurs, notamment le site d'infection des différents types de VPH. Ainsi, on fait une distinction entre les VPHs cutanés et les VPHs muqueux. On peut aussi faire une distinction entre les VPHs

génétaux et les VPHs infectant d'autres régions corporelles, les VPHs génitaux représentant environ 40% des VPHs totaux. Enfin, il existe la classification selon le risque de développement de lésions malignes. Ainsi, on pourra classer les VPHs en types à risque faible (ex. VPH-6, VPH-11, etc.) ou à haut risque (ex. VPH-16, VPH-18, 31, 33, 35, 45, etc.) selon que l'infection soit bénigne ou génère des lésions de plus haut grade.

### ***1.3 Structure du virus et organisation génomique***

#### **1.3.1 Structure du virion**



*Figure 1 : Organisation structurale de la capsidie icosahédrique de VPH. Elle est constituée de 72 capsomères de la protéine structurale majeure (L1).*

Les papillomavirus sont non-enveloppés, de forme icosahédrique et d'un diamètre oscillant entre 52 et 55 nm [49]. La capsidie est constituée des protéines virales L1 et L2, la protéine majeure étant L1. En fait, la création de VLPs par l'expression de la protéine L1 seule à l'aide d'un système d'expression utilisant un Baculovirus a démontré que L2 n'était point requise pour l'assemblage de la particule virale [55]. La



morphologie des capsides ne contenant que L1 semble identique que L2 soit présente ou non [54]. Cependant, il fut observé que l'efficacité de l'assemblage du virion (en terme quantitatif) était plus élevée lorsque L2 était co-exprimée avec L1 [80]. D'ailleurs, L2 serait possiblement impliquée dans l'encapsidation et dans l'interaction de la capside avec l'ADN viral [187].

### 1.3.2 Organisation génomique

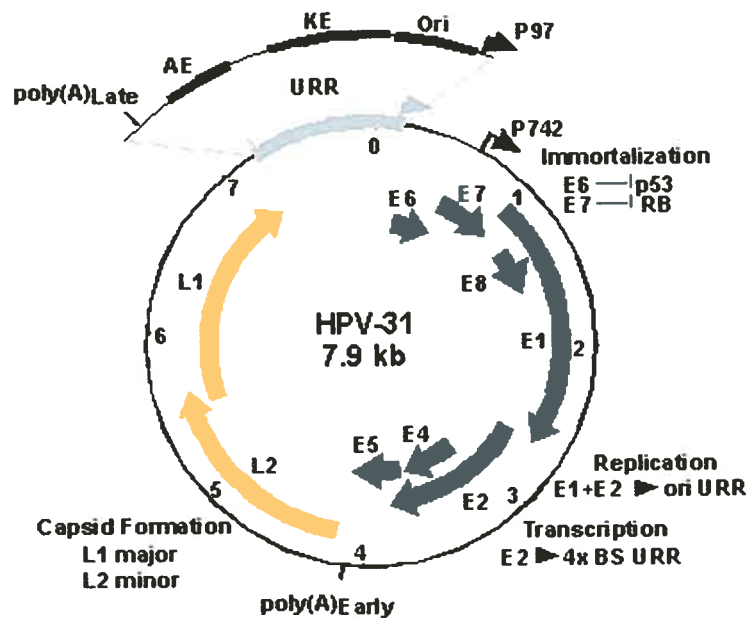


Figure 2 : Génôme du VPH-31. Les gènes tardifs de la capside L1 et L2 (en orange) ont leur propre promoteur p742 et leur propre site de poly-adénylation poly(A). Les gènes précoces E1, E2, E4, E5, E6, E7 et E8 (en gris foncé) utilisent le promoteur p97. Les gènes tardifs utilisent le promoteur p742. La région régulatrice non-codante URR (Upstream Regulatory Region), contient l'origine de réplication et les activateurs transcriptionnels en cis : le domaine AE (Auxiliary Enhancer) et le domaine KE (Keratinocyte Enhancer) auxquels se fixent une panoplie de facteurs de transcription cellulaires régulant la transcription virale (ex. YY1, C/EBP, Sp1). La région URR contient aussi 4 sites de liaison à la protéine régulatrice virale E2.

Quel que soit le type du VPH ou l'espèce infectée, les papillomavirus ont une organisation génomique remarquablement similaire entre eux. Leur génôme consiste en une molécule circulaire d'ADN double-brin d'environ 8 kb. Autre point commun à tous les membres de cette famille virale : les cadres de lecture sont transcrits à partir du même brin, ce qui indique qu'un seul brin de la molécule d'ADN sert de matrice pour la transcription [17]. Peu importe son type, le génôme d'un papillomavirus peut être divisé en trois grandes régions, soit la région des gènes «Late» (tardifs), la région «Early» (précoce) et enfin une région non-codante connue sous la dénomination de LCR (Long Control Region) ou URR (Upstream Regulatory Region). Les gènes tardifs codent pour les protéines de la capsid et ne sont actifs que lors des infections productives. En revanche, les gènes précoces sont impliqués dans la réplication et la transcription virale ainsi que dans la transformation cellulaire. Le LCR, quant à lui, contient les activateurs transcriptionnels, les promoteurs ainsi que l'origine de réplication virale. Le LCR constitue ainsi le modulateur de la réplication et de la transcription virale.

## ***1.4 Transcription et réplication chez les papillomavirus***

### **1.4.1 Fonction des protéines précoces**

La transcription des papillomavirus, s'effectue en deux phases. Dans un premier temps, il y aura transcription des gènes précoces, comprenant les cadres de lecture E1 à E8, sous la direction d'un promoteur spécifique (p97 dans le cas du VPH-31). Les fonctions attribuées aux protéines tardives ayant été décrit suffisamment dans la section 1.3.1, nous allons maintenant faire un survol des protéines précoces.

La protéine E1 a la propriété de lier l'ADN et possède des fonctions ATPase et hélicase [143]. En fait, E1 lie spécifiquement l'origine de réplication du virus [161], bien que cette affinité soit moindre en l'absence de la protéine E2 [163]. E1 aurait aussi la propriété de recruter la machinerie cellulaire pour l'initiation de la réplication au site de l'origine de réplication virale [118]. En somme, E1 apparaît comme la seule protéine

directement impliquée dans la réplication du VPH, bien qu'il semble que E2 ait un rôle auxiliaire dans cette réplication.

En plus de son rôle comme facteur auxiliaire pour la réplication virale, E2 est un important régulateur de la transcription qui peut jouer le rôle d'activateur ou de répresseur selon la position de son site de liaison relativement aux différents promoteurs et aussi selon sa forme. En effet, E2, hormis sa forme complète, existe sous deux formes tronquées (E2TR et E8/E2) qui compétitionnent avec la protéine complète pour les sites de liaison sans en avoir la propriété transactivatrice [92]. La protéine complète, quant à elle, peut jouer le rôle de répresseur ou d'activateur entre autres selon la proximité du promoteur impliqué dans la transcription. Par exemple, le promoteur p97, situé à proximité de sites de liaison à E2, sera réprimé par la fixation de E2, cette dernière provoquant un encombrement stérique empêchant l'assemblage du complexe de pré-initiation de la transcription [160] ainsi que la liaison d'autres facteurs de transcription trans-activateurs tels que Sp1 («promoter-specific transcription factor») et TFIID («transcription factor-IID») [158]. En contrepartie, E2 peut être un transactivateur pour les promoteurs suffisamment distants [100]. Ce mécanisme de trans-activation n'est pas encore bien compris, mais impliquerait la liaison de E2 avec des facteurs cellulaires tels que CBP («CRE binding protein») [93], Sp1 [25] ou p300 [127].

Parmi les gènes précoces, on retrouve aussi le gène E4, dont le patron d'expression est d'ailleurs beaucoup plus caractéristique des gènes tardifs que des gènes précoces. En effet, l'expression de E4 est d'abord détectée au niveau de la couche parabasale de l'épithélium, où la réplication végétative est déjà initiée [18]. La principale caractéristique biologique de E4 semble être son association avec le réseau de cytokératine cellulaire, association qui résulte en la déstabilisation du dit réseau, ce qui faciliterait le relâchement et la transmission du virus [30]. Plus récemment, il fut suggéré que E4 pouvait jouer un rôle dans la régulation du cycle viral en modulant le cycle de division cellulaire [110].

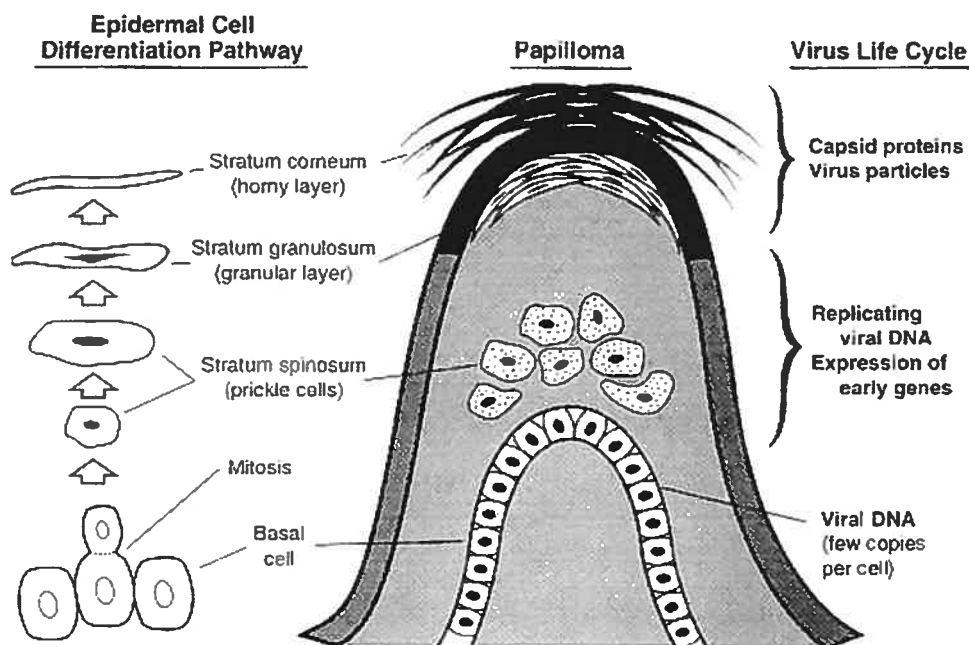
Les autres protéines précoces, soit E5, E6, E7 et E8 ne seront pas abordés dans cette section, E8 parce que n'ayant pas de fonction connue à ce jour et les trois

autres parce qu'elles ont leur rôle à jouer dans la transformation cellulaire et qu'elles seront donc largement décrites dans le chapitre portant sur le cancer du col de l'utérus.

### 1.4.2 Relation virus-kératinocytes

Les papillomavirus infectent les cellules basales de l'épithélium. Cependant, les nouvelles particules virales produites par cette infection ne seront relâchées que par des cellules épithéliales totalement différenciées. C'est donc dire que le cycle végétatif des papillomavirus est intimement lié à l'état de différenciation des cellules de l'hôte infecté, d'où l'impossibilité de produire ce virus en culture monocellulaire. Cette coordination entre le virus et la cellule est en grande partie possible de par la nature du LCR. En vérité, ce dernier constitue le mode de communication principal du virus avec la cellule. Il contient de nombreux sites de liaison pour des activateurs ou répresseurs transcriptionnels agissant en *trans* de façon à moduler la réplication et la transcription virale. Ces facteurs peuvent être d'origine virale (E1 et E2) aussi bien que d'origine cellulaire, comme par exemple les facteurs de transcription AP-1 («activator protein-1»), NFI («nuclear factor-1»), KRF-1 («keratinocyte-specific transcription factor»), Oct-1 («octamer binding factor»), C/EBP $\beta$  («CCAAT/enhancer binding protein beta»), TEF-1 («transcriptional enhancer factor») et YY1 («Yin Yang-1»), et jouent un rôle direct dans l'expression génomique du VPH. Ainsi, AP-1, NFI, TEF-1, KRF-1 (Myers, p. III 21-III 40) contribueraient à la régulation positive du LCR alors que E2, C/EBP $\beta$  [88], Oct-1 [70] et YY1 [98] agiraient comme répresseurs. En fait, le spectre de facteurs cellulaires se liant au LCR serait un déterminant majeur du tropisme épithélial des VPHs ainsi que de leur dépendance sur l'état de différenciation cellulaire pour la complétion de leur cycle réplcatif [78]. Par exemple, le facteur Sp1 est un activateur transcriptionnel du promoteur p97 viral, qui régule l'expression des protéines de transformation cellulaire E6 et E7 [156]. Or, le facteur YY1 compétitionne avec Sp1 pour son site de liaison et agit ainsi comme un répresseur transcriptionnel [28]. L'exemple du facteur C/EBP $\beta$  (aussi nommé NF-IL6) est aussi intéressant, dans la mesure où ce facteur pourrait en partie être responsable du faible niveau de copies génomiques virales dans les couches

inférieures de l'épithélium. En effet, C/EBP $\beta$  est un répresseur de la transcription et de la réplication des papillomavirus [88, 172]. Or, une diminution de sa disponibilité dans les kératinocytes lors de leur différenciation pourrait permettre une recrudescence du nombre de copies d'ADN viral dans la cellule [172].



*Figure 3 : Représentation schématique d'un papillome. Le cycle vital des papillomavirus est intimement lié à la différenciation cellulaire. Au niveau des cellules basales, l'ADN viral est maintenu à un faible nombre de copies. Les événements tardifs de la réplication virale (synthèse de la capside et assemblage des virions) ne se produisent que dans les cellules totalement différenciées.*

Ainsi, le cycle de réplication du virus se trouve rigoureusement lié au stade de différenciation cellulaire, et ce du fait que la nature et la quantité des facteurs de transcription varient selon le niveau de différenciation cellulaire. La conséquence principale de cette modulation est que les VPHs ont un patron d'expression relativement

complexe lors de l'infection productive. Suivant l'infection des cellules basales de l'épithélium, le génôme du VPH y est maintenu à un faible niveau de copies épisomales. Lorsque les cellules infectées migrent de la couche basale et se différencient, la transcription des gènes précoces est accélérée et l'ADN viral amplifié [91]. Dans les faits, cela se traduit par la présence des transcrits précoces au niveau des cellules basales de l'épithélium, notamment de transcrits E6 et E7 tel qu'établi par l'utilisation de sondes spécifiques pour ces gènes [150]. Quant aux gènes viraux tardifs encodant les protéines de la capside, leur expression n'est détectée que dans les cellules épithéliales totalement différenciées de la verrue, ainsi que l'usage d'un antisérum spécifique aux antigènes capsidiens l'a déjà démontré [73].

*tel que démontré par l'utilisation* ✓

### 1.4.3 Réplication

La forte relation entre le cycle de réplication des papillomavirus et la différenciation cellulaire pose évidemment un gros problème pour l'étude du cycle vital de ces virus. L'absence de système de culture efficace a fortement miné l'évolution de nos connaissances des événements impliqués aux différents moments de ce cycle. Ainsi, nous savons depuis peu que l'attachement du virus à la cellule épithéliale basale est médié par une interaction entre les capsomères de la protéine L1 virale et l' $\alpha_6$ -intégrine cellulaire, qui jouerait donc le rôle de récepteur pour le virus [36]. Les mécanismes d'endocytose, de transport au noyau cellulaire et de décapsidation sont encore nébuleux. Il est possible que l'entrée du virion au noyau se fasse par l'entremise d'une interaction de la protéine L1 avec le système d'import nucléaire Kap beta2 et Kap beta3 [112]. Dans le cas d'une infection productive, le virus se répliquera à un haut niveau et il y aura synthèse des protéines tardives de structure dans les kératinocytes différenciés. Le relâchement des particules virales se fera au niveau de la strate cornée de l'épithélium sans qu'il y ait lyse cellulaire causée par le virus. Il a été établi que chez le VPH de type 59, l'action du virus provoquait une baisse de transcription de la loritrine, protéine majeure de la couche cornée de l'épithélium [94]. Ce mécanisme serait impliqué dans le

relâchement du virus. À noter que de toute façon, ces cellules entrent de manière naturelle en apoptose.

Enfin, il est à noter que dans certains cas, le virus établira une infection latente. Dans ce cas, le génôme viral sera intégré au génôme cellulaire. D'ailleurs, l'intégration au génôme de l'hôte semble un pré-requis pour le phénomène de transformation observé chez les types oncogènes. La lignée de cellules HeLa en est un bon exemple. En effet, cette lignée cellulaire provient d'un cancer cervical et porte des séquences de 20 à 40 copies du virus VPH-18 intégrées dans son génôme [97]. Le phénomène d'intégration sera traité plus largement ultérieurement.

## ***1.5 Caractéristiques biologiques des papillomavirus***

### **1.5.1 Spécificités tissulaire et d'espèce**

Un papillomavirus donné n'infecte qu'une espèce animale. En fait, il n'y a aucun exemple connu de transmission d'un papillomavirus humain à une autre espèce. De plus, les VPHs démontrent un tropisme cellulaire marqué. Ils n'infectent que les cellules de l'épithélium squameux de la peau ou des muqueuses, produisant la plupart du temps des lésions bénignes : des verrues. Des types spécifiques du VPH peuvent montrer des préférences différentes pour l'épithélium muqueux ou l'épithélium cutané, mais ont en plus des préférences marquées pour certaines régions corporelles caractéristiques. Par exemple, le VPH-1, agent étiologique des verrues plantaires, qui n'infecte que l'épithélium cutané du pied ou bien le VPH-11, qui infecte l'épithélium muqueux de l'appareil génital, respiratoire et digestif (Fields, p.2085). Ces spécificités tissulaires et d'espèce découlent sans doute des différences notables entre les récepteurs cellulaires de deux espèces ou deux cellules différentes, mais découlent possiblement de la dépendance du virus face à des facteurs de transcription cellulaires spécifiques, facteurs nécessaires pour la complétion d'un cycle vital complet du virus, tel que décrit dans la section sur la réplication et la transcription virale.

### 1.5.2 Prolifération et transformation cellulaire

La propriété fondamentale des papillomavirus est l'induction de prolifération et de transformation cellulaire. La première réaction des cellules infectées est en effet la prolifération, d'où la manifestation clinique la plus commune des papillomavirus : l'induction de verrues (lésions bénignes). Cependant, la capacité transformante des types oncogènes (types 16, 18, 31, 33, 35, 45, 51, 58, 59 [106] ) est plus inquiétante. Cette capacité réside dans les cadres de lecture E6 et E7. L'expression des protéines E6 et E7 des types oncogènes dans des kératinocytes est suffisante pour provoquer l'immortalisation de ces cellules (ces cellules présentent aussi des anomalies dans la différenciation remarquablement similaires aux anomalies relevées dans les néoplasies du col utérin), alors que cette propriété est inexistante pour les protéines des types non-oncogènes [104]. D'autre part, les kératinocytes transfectés avec de l'ADN des VPHs oncogènes deviennent immortalisés et contiennent le génôme viral sous une forme intégrée [140]. Lorsque les cellules immortalisées sont introduites dans une souris «nude», elles deviennent tumorigéniques après l'introduction d'un oncogène actif [125]. Il est bien important de noter que tous les types de papillomavirus induisent la prolifération cellulaire, mais que seuls les types oncogènes ont la capacité transformante, la prolifération cellulaire se traduisant par l'apparition de verrues bénignes chez les types non-oncogènes [140]. Les mécanismes de cette transformation seront abordés en détail dans le chapitre suivant.



## 2. PAPILOMAVIRUS HUMAINS GÉNITAUX ET CANCER DU COL UTÉRIN

### *2.1 Association épidémiologique*

Depuis quelques années, de fortes évidences épidémiologiques, pathologiques et moléculaires se sont accumulées concernant l'implication des papillomavirus en tant qu'agent causal des néoplasies du col de l'utérus. En fait, les données épidémiologiques montrent que le cancer du col utérin partage les mêmes facteurs de risque que les maladies transmises sexuellement [135]. Les facteurs les plus convaincants sont une sexualité précoce et le nombre de partenaires sexuels [57]. L'activité sexuelle précoce constituerait un risque du fait que lors de la puberté, les tissus cervicaux subissent différents changements, rendant cette région plus vulnérable aux lésions. Le facteur du partenaire sexuel aurait aussi son importance, car les partenaires masculins des patientes atteintes du cancer du col de l'utérus ont en effet beaucoup plus de partenaires sexuelles que ceux des femmes qui n'en sont pas atteintes [12]. Or, ces hommes seront souvent des porteurs de VPHs. En fait, la détection d'ADN de VPHs provenant du pénis de l'homme est liée à un risque substantiel de développement de cancer cervical chez sa partenaire sexuelle [9].

Indépendamment des différents facteurs de risque pouvant être impliqués dans le cancer du col utérin, l'association entre le VPH et cette néoplasie est extrêmement forte. Une étude mondiale menée par «The International Agency for Research on Cancer» a démontré que la prévalence de VPHs à haut risque (surtout VPH-16) dans des spécimens de biopsies provenant de cancers cervicaux était de 99.7%, indiquant ainsi que le virus était une cause essentielle de ce cancer [40]. Cependant, il est important de mentionner que l'infection au VPH en soi n'est pas suffisante pour provoquer l'apparition de néoplasies et qu'en fait, seule une minorité de femmes infectées développeront cette pathologie [37]. Des études suggèrent aussi que des facteurs hormonaux, le tabagisme, le

nombre d'accouchements, l'usage à long terme de contraceptifs oraux, des déficiences nutritionnelles et les infections à *Chlamydia trachomatis* pourraient tous être des cofacteurs dans le développement du cancer du col utérin [176, 82, 62, 90, 146]. Évidemment, une immunosuppression associée à une greffe d'organes ou au SIDA est aussi un cofacteur déterminant dans l'apparition de lésions malignes [11].

## ***2.2 Épidémiologie du cancer du col de l'utérus***

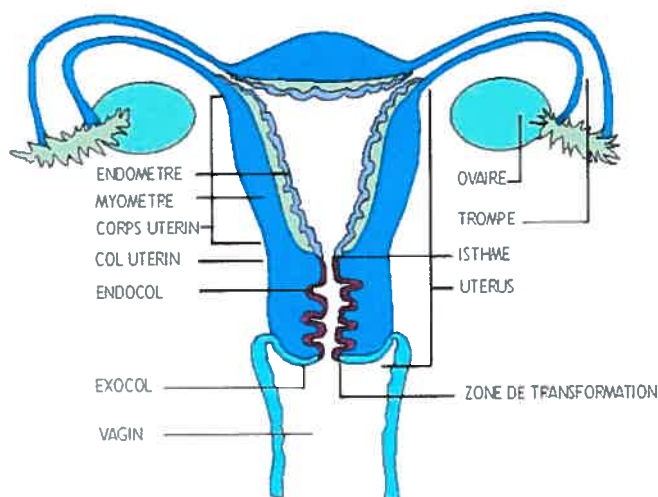
Environ 371 000 nouveaux cas de cancer invasifs du col de l'utérus seraient diagnostiqués chaque année mondialement, ce qui représente pratiquement 10% de tous les cancers chez la femme [39]. Ce serait le troisième cancer en importance chez la femme (après les cancers du sein et colorectal) et le septième en importance hommes et femmes confondus [122]. L'incidence du cancer du col utérin est plus élevée dans les pays en voie de développement où il fut le plus fréquent jusqu'au début des années 90 [120] et où il est en résurgence en ce moment [39]. En fait, dans les pays industrialisés le taux d'incidence se situe aux abords de 10 nouveaux cas par an pour 100 000 femmes, alors que dans les régions en voie de développement, comme l'Afrique, l'incidence peut grimper à plus de 40 pour 100 000 par an [37]. D'ailleurs, 4 nouveaux cas sur 5, ainsi que la plupart des morts dues au cancer du col de l'utérus surviennent dans les pays en développement [121]. Cette prévalence élevée de cancer chez les populations défavorisées est directement imputable à la non-application de moyens de dépistage efficaces, tels que le test Pap (prélèvement de cellules exfoliées du col de l'utérus). Depuis l'introduction de ce dernier, il y a 50 ans, la mortalité associée au cancer du col utérin a chuté de 73% au Canada (Santé Canada, 1998). D'ailleurs, le fait de ne pas subir régulièrement de tests Pap constitue le principal facteur de risque pour un pronostic défavorable chez les femmes qui développent un cancer du col utérin [89].

Les femmes dans la quarantaine constituent le groupe le plus à risque pour le cancer du col utérin [61]. Cet état de fait est logique étant donné que la majorité des infections au VPH ont lieu au début de la vie sexuelle active des femmes, c'est-à-dire de

15 à 20 ans. Or, en admettant que le virus persiste suffisamment longtemps (il sera éliminé en 9 mois dans la majorité des cas [41] ), le développement du cancer prendra plusieurs décennies. Bien entendu, dans la plupart des cas, les lésions régresseront avant d'atteindre la malignité après élimination du virus.

### 2.3 Pathogenèse du cancer du col utérin

Les cancers invasifs du col utérin sont précédés par une série progressive de lésions de l'épithélium cervical, que l'on appelle la zone de transformation.



*Figure 4 : Zone de transformation. La haute incidence du cancer du col utérin comparé à la basse incidence des autres cancers à d'autres sites de l'appareil génital féminin suggère une forte sensibilité de la zone de transformation aux carcinogènes.*

Ces anomalies épithéliales sont classées en LSIL («Low-Grade Intraepithelial lesion» : lésion intraépithéliale de bas grade) et en HSIL («High-Grade Intraepithelial Lesion» : lésion intraépithéliale de haut grade) [86]. Cette classification des lésions se base sur le degré de remplacement des cellules épithéliales normalement différenciées

par des cellules de type basales non-différenciées. Il semblerait que l'infection par certains types de VPHs à haut risque puisse mener sans transition aux lésions de haut grade, bien que ce mécanisme soit mal connu, mais le processus est généralement évolutif et peut prendre plusieurs décennies [81]. Dans les lésions de haut grade, pratiquement toute l'épaisseur de l'épithélium est graduellement occupée par des cellules anormales. On parle de cancer invasif lorsque ces cellules anormales migrent à travers la membrane basale pour envahir le tissu stromal et éventuellement produire des métastases à différents sites corporels. Notons ici que la plupart des anomalies cytologiques régressent d'elles-mêmes et que seule une fraction d'entre elles progressent en carcinome [111].

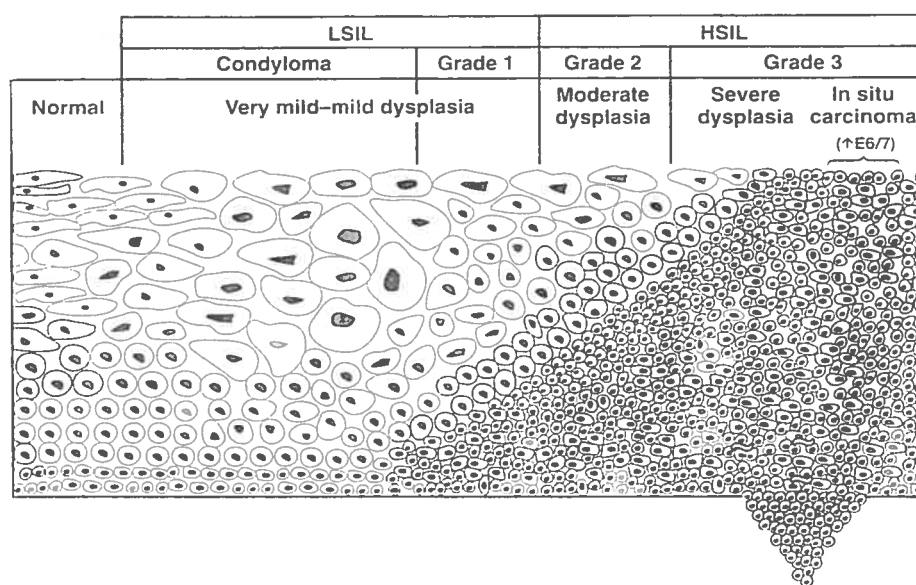


Figure 5 : Représentation schématique des lésions précurseurs du cancer du col utérin.

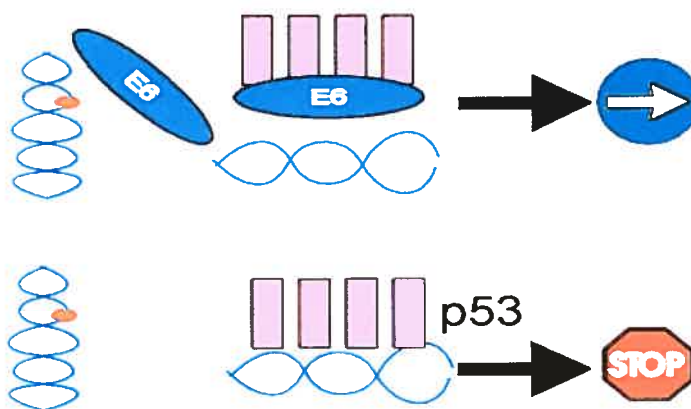
## 2.4 Mécanismes d'induction du cancer

Les propriétés oncogènes des papillomavirus à haut risque sont dues à l'expression des protéines E5, E6 et E7. Le rôle de la protéine E5 est encore nébuleux,

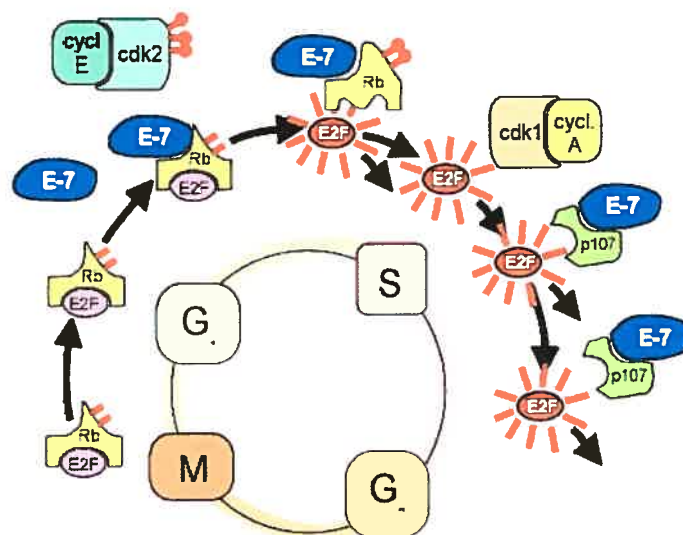
mais cette protéine interagit avec des facteurs de croissances associés à la membrane cellulaire [154] et interfère avec l'apoptose médiée par les ligands Fas (CD95) ou TRAIL («tumor necrosis factor-related apoptosis-inducing ligand») [75]. De plus, la protéine E5 interagit avec le récepteur du facteur de croissance épidermique (EGFR), dont la réponse au facteur de croissance épidermique (EGF) serait amplifiée par E5 [153]. À noter que EGF est un mitogène primaire essentiel pour les kératinocytes [131]. L'expression de E5 dans des kératinocytes humains peut aussi causer une augmentation de la quantité de EGFR sur la surface cellulaire. Ce phénomène serait causé par une inhibition de l'acidification des endosomes cellulaires par E5 [153]. Ceci a donc pour résultat une mitose cellulaire accrue, et ce de façon à augmenter le nombre de kératinocytes infectés par le VPH [130]. Cependant, le pouvoir oncogène est surtout médié par les gènes E6 et E7 [63], dont les mécanismes d'action sont mieux connus.

E6 est une petite protéine d'environ 150 acides aminés, selon le type de VPH. Elle a la propriété de lier la protéine cellulaire p53, un suppresseur de tumeur. Cette protéine est produite lorsque l'ADN cellulaire subit des dommages. Si le dommage est mineur, une augmentation dans la concentration de p53 mène à l'arrêt de la mitose en G1 pour permettre la réparation de l'ADN endommagé. Si le dommage est trop important, p53 induit la cellule à prendre la voie de l'apoptose (mort cellulaire programmée), prévenant ainsi la formation de cellules anormales [190]. Or, la liaison de E6 à p53 cible cette dernière pour la dégradation par la voie de l'ubiquitine [139]. Cette dégradation provoque une diminution considérable de la quantité de p53 actif dans la cellule avec la double conséquence de favoriser l'accumulation de mutations ou d'ADN anormal, le cycle n'étant plus stoppé en G1 après endommagement de l'ADN, tout en prévenant l'apoptose cellulaire. Ainsi, le VPH oncogène détient deux outils pour interférer avec l'apoptose, soit E5 et E6, ce qui lui confère un avantage de croissance supplémentaire. De plus, E6 peut augmenter l'activité de la télomérase cellulaire [84], dont l'activité est liée à l'immortalisation cellulaire et est caractéristique de la plupart des lignées cellulaires et des tumeurs [79]. La télomérase est la transcriptase inverse responsable de la synthèse des télomères, soit les séquences répétées d'ADN formant la

fin des chromosomes [51]. Dans la plupart des cellules humaines normales, l'absence d'activité télomérase résulte en un raccourcissement progressif des télomères à chaque division cellulaire, menant éventuellement à l'arrêt de la croissance et à la sénescence [165]. Au contraire, les cellules immortalisées ou cancéreuses possèdent une activité télomérase, voyant ainsi leur potentiel prolifératif maintenu [79]. En fait, bien que le mécanisme d'activation soit encore mal connu, E6 coopérerait avec l'oncogène Myc pour transactiver la transcription de la sous-unité catalytique de la protéine hTERT («human telomerase reverse transcriptase») par fixation sur son promoteur [166]. Cette sous-unité serait le facteur limitant dans l'absence d'activité télomérase au sein des cellules normales [166]. Or, sa transactivation par E6 aurait pour effet l'apparition de l'activité télomérase avec une immortalisation conséquente de la cellule.



*Figure 6 : Représentation schématique de l'inhibition du suppresseur de tumeur p53 par la protéine transformante E6. Lorsque des dommages à l'ADN surviennent, l'augmentation de l'expression de p53 provoque l'arrêt du cycle cellulaire ou l'apoptose. La liaison de E6 à p53 provoque la dégradation de celle-ci, entraînant une chute drastique de sa concentration et favorisant l'accumulation de mutations.*



*Figure 7 : Représentation schématique de l'inhibition de pRb par l'oncoprotéine E7. E7 lie préférentiellement pRb dans sa forme hypophosphorylée et provoque le relâchement du facteur de transcription E2F. E7 peut aussi transactiver les cyclines A et E et lier p107, qui a une fonction similaire à pRb. Ces interactions résultent en une accélération du cycle cellulaire et éventuellement en l'immortalisation de la cellule.*

La protéine E7, quant à elle, est constituée d'environ 100 acides aminés. Elle a l'habileté de lier et d'inactiver fonctionnellement la protéine cellulaire pRb [105], qui est aussi un suppresseur de tumeur ayant un rôle central dans le contrôle normal du cycle cellulaire. Il faut noter que la protéine E7 des VPHs non-oncogènes a une affinité de liaison moindre avec pRb que les types oncogènes, indiquant que cette fonction de E7 est importante pour la transformation [105]. En fait, E7 se lie préférentiellement à pRb dans sa forme hypophosphorylée, provoquant ainsi un relâchement excessif du facteur de transcription cellulaire E2F [174]. Ceci entraîne la progression accélérée du cycle cellulaire et de la synthèse d'ADN (E2F induit la progression du cycle cellulaire en phase S [23] ) et la perte de contrôle de l'expression des gènes dépendant de la régulation cyclique de E2F, tels que ceux codant pour les cyclines A et E [186, 44], qui

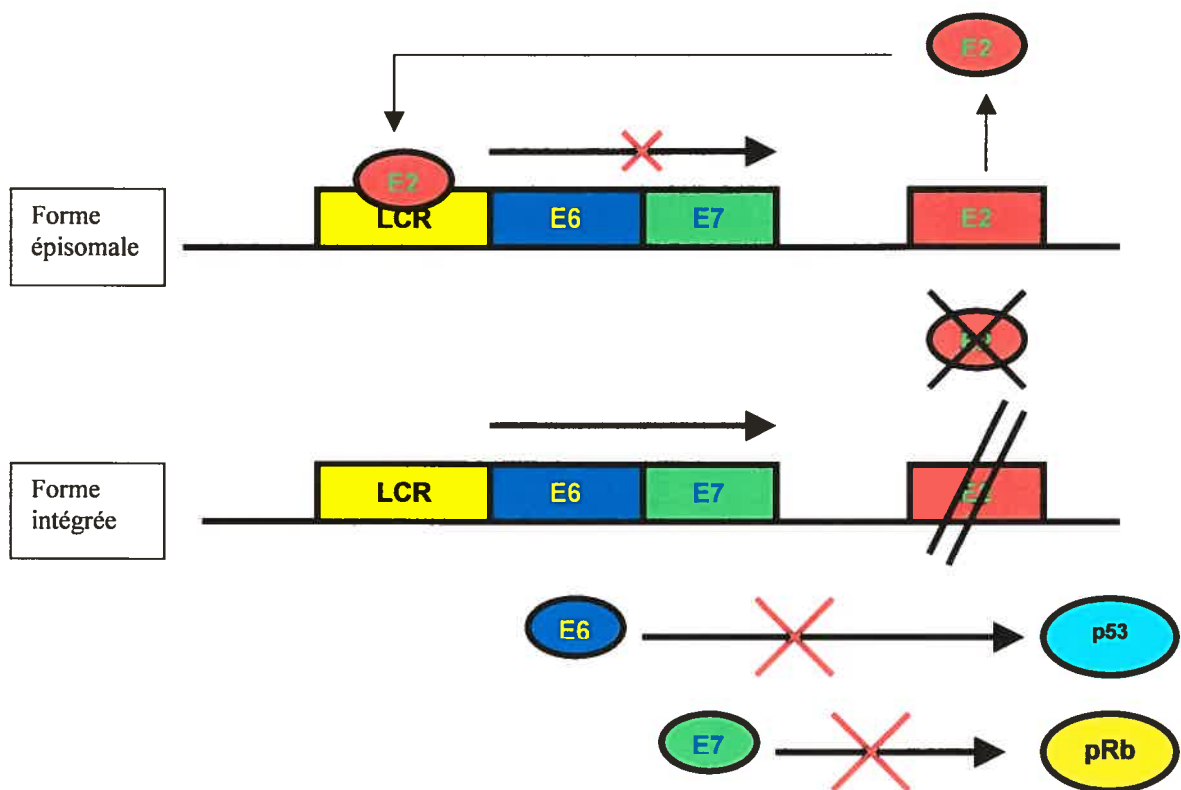
sont d'importants régulateurs positifs du cycle cellulaire. Une sur-expression de la cycline E a été observée au niveau des HSILs [147]. Or, un niveau élevé de cette cycline cause une progression accélérée du cycle cellulaire en phase S, ce qui résulte en une instabilité génétique [148]. La cycline A, quant à elle, favorise les transitions G<sub>1</sub>-S et G<sub>2</sub>-M du cycle cellulaire[56]. D'autre part, E7 lie et inhibe p107 et p130 [33], des protéines cellulaires très similaires à pRb qui sont des inhibiteurs des cyclines A et E [177]. Le résultat de la liaison de E7 à p107 mène également au relâchement de E2F [185].

En somme, il apparaît que E6 cause une dérégulation de la stabilité génomique cellulaire et que E7 s'associe avec des protéines régulatrices contrôlant la progression du cycle cellulaire. Ces oncoprotéines confèrent ainsi un fort avantage de croissance à la cellule hôte, la menant à une prolifération accrue et à une dédifférenciation. Des études *in vitro* portant sur des kératinocytes sur-exprimant les oncogènes E6 et E7 ont confirmé leurs fonctions. Des changements morphologiques similaires aux altérations précancéreuses ont été observés et ces cellules ont formé des tumeurs malignes après injection chez des souris [8]. Cependant, il est important de prendre en note le fait que les protéines E6 et E7 de différents types de VPHs ont des niveaux d'efficacité variable: les protéines des VPHs à haut risque sont très actives dans leurs fonctions oncogènes, alors que les types à bas risque expriment des protéines E6 et E7 incapables de lier ou d'inactiver p53 et pRb [20, 58].

C'est donc dire qu'une expression accrue *in vivo* de ces oncogènes pourrait avoir des conséquences néfastes. Et c'est exactement ce qui est à la source du pouvoir carcinogène des papillomavirus à haut risque. Ces VPHs à haut risque ont en fait la capacité de s'intégrer au génôme de la cellule hôte, cette intégration étant centrale pour la transformation cellulaire et le développement de néoplasies du col utérin [119]. L'intégration du génôme viral nécessite évidemment sa rupture, or cette dernière s'effectue habituellement au niveau du cadre de lecture E2 [132]. En contraste, la région du LCR et les ORFs E6 et E7 en amont de celui-ci restent généralement intacts [188].



On se rappellera que E6 et E7 sont sous contrôle du promoteur p97, celui-ci étant réprimé par la fixation de E2 sur le LCR. Donc, l'intégration du génôme viral dans l'ADN de la cellule hôte provoque la rupture de E2, entraînant ainsi la sur-expression des oncogènes E6 et E7 et conférant par le fait même un avantage de croissance à la cellule infectée [74]. Chez VPH-16 ainsi que chez VPH-33, on a remarqué une association entre une progression rapide des lésions et une haute charge virale de virus intégré [126, 137]. Dans les lésions de bas grade, le génôme viral est majoritairement sous sa forme épisomale et l'expression des oncoprotéines E6/E7 pratiquement indétectable [162]. Par contre, dans les HSILs, l'expression des protéines oncogènes est accrue, due à l'intégration de l'ADN viral [32].



*Figure 8 : Intégration du génôme des papillomavirus. L'intégration cause la rupture du cadre de lecture E2, permettant une transcription accrue des gènes E6 et E7, codant pour les oncoprotéines virales qui inactivent respectivement les suppresseurs de tumeur cellulaires p53 et pRb.*

## 2.5 *La persistance*

Le facteur clé du développement d'un cancer suite à l'infection par un VPH de haut risque est la persistance. En effet, le cancer prenant plusieurs décennies à se développer, le virus doit obligatoirement se stabiliser au sein de l'organisme de façon à y rester. Plusieurs études ont démontré que les infections aux VPHs de haut risque étaient principalement transitoires ou même asymptomatiques chez les jeunes femmes [35, 65, 140]. Cependant, bien qu'il soit maintenant accepté que la plupart des infections aux VPHs se résolvent d'elles-mêmes, les facteurs favorisant une infection transitoire ou persistante restent largement inconnus. Une étude a tenté d'éclaircir ce phénomène et a montré que la persistance se comporte un peu comme un cercle vicieux : plus l'infection persiste, plus il est difficile de l'éliminer. Ainsi, si l'infection n'était pas disparue après douze mois, la probabilité qu'elle se résolve chutait drastiquement [66]. Ce schéma de persistance de l'infection est fortement similaire au schéma de persistance des lésions intraépithéliales, qui veut en toute logique que plus les lésions pré-invasives persistent, plus le risque de voir ces lésions évoluer en cancer augmente [67]. D'autre part, l'infection ainsi que les changements cytologiques causés par celle-ci, s'il y en a, régressent habituellement après quelques mois ou années de façon spontanée, probablement suite à une réponse immunitaire cellulaire de l'hôte ciblant le VPH [65]. Cette évolution est en accord avec le concept actuel de carcinogénèse qui requiert l'action consécutive de plusieurs oncogènes. Dans les faits, la persistance est généralement définie comme étant la détection du même type de VPH dans des échantillons consécutifs obtenus pour plus de 6 mois [99], mais il n'y a pas vraiment de consensus général à ce sujet.

Les facteurs influant la persistance sont nombreux et proviennent du virus lui-même aussi bien que des comportements et du bagage génétique de l'hôte. D'abord, la charge virale infectieuse est positivement associée avec la persistance [13]. Cependant, le génotype du VPH infectieux est probablement le facteur viral le plus important. Les infections aux VPHs à haut risque perdurent plus longtemps que celles impliquant les VPHs de bas risque. Ainsi, l'intervalle de temps médian pour la perte des VPHs

oncogènes est plus grand (8.9 mois) que pour les types à bas risque (4.8 mois) [41]. De nombreuses études ont mis en évidence le fait que les VPHs persistants étaient presque toujours des VPHs à haut risque. De fait, dans une étude, les types 16, 18, 31, 33 et 35 semblent persister davantage que les autres [13]. Dans une autre étude, 95% des infections persistantes mettaient en cause des types à haut risque, soit les types 16, 18, 31, 33, 35, 39, 45, 51, 52 et 58, et 33 impliquaient VPH-16 [67]. D'ailleurs, il est connu que le VPH-16 persiste plus longtemps que les autres types [13]. Les variations intratypiques peuvent aussi avoir une influence sur la capacité à persister du virus. Par exemple, certains variants spécifiques de VPH-16 persistent mieux et plus longtemps que d'autres [37]. La notion de variant intratypique en relation avec la persistance sera largement traitée plus tard.

Chez l'hôte, le facteur de la réponse immunitaire a évidemment une grande importance pour la persistance de l'infection. Ainsi, l'association entre l'infection persistante et l'infection simultanée par plusieurs types suggère que des femmes infectées par des types multiples pourraient avoir une réponse immunitaire faible face aux VPHs, les prédisposant donc aux infections persistantes [168]. D'ailleurs, comme nous le détaillerons dans le prochain chapitre, l'infection au VIH augmente le risque d'infection multi-typique et de persistance [72, 102]. D'autre part, l'âge (les femmes plus âgées sont plus à même d'être aux prises avec une infection persistante au VPH) ainsi que l'usage de contraceptifs oraux constituent aussi des facteurs de risque pour l'infection persistante [65, 13].

### 3. VPH ET VIH

À travers le monde, on estime que 40% des millions de personnes infectées par la VIH sont des femmes (Centers for Disease Control and Prevention. 1997). C'est en 1992 que l'organisme «Centers for Disease Control and Prevention» a ajouté le cancer invasif du col de l'utérus à la liste des maladies impliquées dans la définition du SIDA (Centers for Disease Control and Prevention. 1997). Les interactions VIH-VPH peuvent être de deux natures : indirecte (ex. recrudescence de la réplication du VPH due à la réduction du système immunitaire de l'hôte induite par le VIH) ou directe, comme par exemple l'action de protéines régulatrices produites par le VIH sur l'expression génique du VPH.

#### *3.1 Interactions indirectes*

L'affaiblissement du système immunitaire induite par le VIH est au centre des effets indirects de ce virus sur l'infection au VPH. Il est connu que des femmes ayant un système immunitaire affaibli dû à une transplantation d'organe sont plus à risque pour le développement de cancers cervicaux que les femmes ayant une immunité normale [128]. D'autre part, les femmes infectées au VPH sans avoir de lésions intraépithéliales offrent une meilleure réponse (médiée par les cellules T cytotoxiques) aux antigènes du virus que celles développant des SILs (Squamous Intraepithelial Lesions) [107]. Ces faits suggèrent qu'une réponse immunitaire cellulaire efficace est associée avec une protection contre le développement de lésions dues à l'infection au VPH. Or, si l'infection au VIH mène à l'affaiblissement global de la réponse CMI (Cell-Mediated Immunity), alors la perte de cette réponse face aux antigènes du VPH peut être importante dans le développement de lésions anogénitales pré-invasives associées à l'infection aux VPHs. En somme, les VPHs jouent un rôle important dans le développement initial de ces lésions, et la réponse immunitaire atténuée associée au VIH joue un rôle en n'éliminant pas les cellules épithéliales exprimant les antigènes du VPH.

Dans les faits, la séropositivité au VIH résulte en une prévalence accrue des infections aux VPHs [168] et en un risque plus élevé de développer des SILs [83]. De plus, les lésions associées au VPH chez les femmes immunosupprimées VIH-positives sont plus grandes en taille et en nombre et ont un plus haut taux de récurrence après traitement par rapport aux lésions observées chez les femmes immunocompétentes [42]. Des évidences démontrent que les femmes VIH-positives sont plus à même d'acquérir de nouvelles infections aux VPHs de haut risque [102], que ces infections persistent généralement plus longtemps, qu'une réponse immunitaire affaiblie (mesurée par un compte de lymphocytes T CD4+) était associée à une plus grande persistance [157] ainsi qu'à une charge virale de VPH élevée [144]. Or, tel que mentionné précédemment, la persistance de l'infection à VPH est centrale au développement des lésions malignes. Mentionnons aussi que chez les femmes VIH-positives, l'expression du VPH est accrue et que ces femmes sont plus à même d'avoir des infections simultanées à plus d'un type [72]. Dans un autre ordre d'idées, le taux de dysplasies, ainsi que leur persistance et leur récurrence sont plus élevés chez les individus immunosupprimés par le VIH [95] alors qu'une basse charge virale du virus et un compte élevé de cellules CD4 sont associés à une régression des lésions [101]. De plus, la progression du cancer chez ces femmes peut être particulièrement agressive, avec un schéma métastatique inhabituel s'accompagnant d'une réponse appauvrie aux thérapies standards [138, 21].

Logiquement, les effets de l'interaction VIH-VPH devraient correspondre à une plus haute prévalence du cancer du col utérin chez les femmes VIH-positives. Une étude menée sur une cohorte de femmes de New York supporte cette hypothèse : le cancer invasif du col de l'utérus y est trois fois plus fréquents chez les femmes infectées par le VIH [34]. Cependant, les études supportant cette hypothèse sont plutôt rares. En effet, jusqu'à récemment, la plupart des femmes VIH-positives mourraient d'autres complications liées au SIDA avant que puisse s'écouler la longue période de temps que demande la progression des lésions pré-invasives en cancer invasif [115].

C'est la venue du traitement HAART («Highly Active Antiretroviral Therapy») qui pourrait changer cet état de fait. Ce traitement, en usage clinique depuis 1995, a grandement amélioré la survie des patients porteurs du VIH, cette espérance de vie allongée s'accompagnant d'une diminution de la réplication virale, d'une augmentation du compte des lymphocytes T CD4<sup>+</sup> et d'une réduction de la plupart des maladies opportunistes [116]. Le traitement HAART peut influencer la relation entre le VIH et le VPH de deux façons très paradoxales. D'abord, en prolongeant la vie de l'individu, il allonge la durée d'infection au VPH, permettant le développement de lésions malignes. En contrepartie, par sa propriété à faire chuter la charge virale du VIH et à rétablir le taux de cellules CD4, il peut restaurer partiellement la compétence immunitaire et ainsi permettre de combattre plus efficacement le développement des lésions pré-invasives associées aux VPHs à haut-risque. Dans les faits, cela s'est traduit par une légère augmentation de l'incidence du cancer du col utérin depuis la venue de l'ère de ce traitement, ce qui suggère que la reconstitution du système immunitaire résultant du traitement n'est pas suffisante pour prévenir le développement du cancer une fois le processus enclenché (Centro Operativo AIDS, 2000). D'un autre côté, un projet portant sur des femmes VIH-positives ayant des résultats Pap normaux à la base a démontré que l'incidence des lésions de bas grade est inférieure chez les patientes ayant reçues la thérapie anti-rétrovirale par rapport à celles n'ayant pas été traitées [24]. Donc, à court terme, il est possible que le nombre de cas de cancer du col utérin chez les femmes sidatiques traitées augmente légèrement, mais à long terme, la reconstitution immunitaire due au traitement pourra prévenir les infections au VPH ou permettre la régression des lésions de bas grade, permettant ainsi une possible diminution de l'incidence du cancer.

### ***3.2 Interactions directes***

Des mécanismes autres qu'une immunosuppression sévère peuvent jouer un rôle dans l'augmentation de l'expression de l'ADN des VPHs et dans les lésions qui y sont associées chez les femmes infectées par le VIH. Ainsi, des interactions locales VIH-

VPH aux niveaux tissulaire et cellulaire peuvent jouer un rôle potentiel dans l'apparition de néoplasies ano-génitales. Les cellules porteuses du VPH sont exposées à des facteurs viraux et cellulaires relâchés dans les tissus par des cellules porteuses du VIH [27].

*tat a la  
capacité de  
perforer la  
membrane  
par un  
mécanisme  
recepteur  
indépendant*

Notamment, la protéine trans-activatrice du VIH (tat) peut être sécrétée par les cellules infectées par ce virus et il a été rapporté qu'elle pouvait trans-activer des promoteurs viraux, comme celui du cytomégalovirus [69]. Quant au VPH, il a été démontré que tat avait un effet trans-activateur sur le LCR du VPH-16 [169], notamment sur le promoteur p97 (précoce), ce qui aurait pour résultat une transcription accrue des gènes E6 et E7. Le mécanisme de cette transactivation pourrait se faire via le recrutement par tat de facteurs transcriptionnels au niveau des régions activatrices du LCR du VPH [169].

La protéine Rev du VIH aurait aussi un effet sur l'infection au VPH, cet effet étant cette fois post-transcriptionnel. La protéine Rev interagit avec une séquence d'ARN appelé «Rev-responsive element» (RRE), et permet ainsi le transport des ARNm du noyau au cytoplasme de façon à ce que la traduction puisse s'accomplir [96]. Or, des travaux portant sur l'effet de cette protéine sur la régulation du VPH-1 ont mis au jour une séquence inhibitrice riche en nucléotides AU de 27 nucléotides dans la région 3' du LCR du VPH-1, séquence qui inhiberait le transport des ARNm au cytoplasme. Il se trouve que la protéine Rev du VIH était capable de compenser l'effet inhibiteur de cette séquence du VPH [159], ce qui pourrait signaler un autre mécanisme de transactivation du VPH par le VIH.

Une autre possibilité serait que l'expression aberrante de cytokines par les cellules infectées du VIH joue un rôle potentiel dans l'infection au VPH. Par exemple, une étude démontre que l'interleukine-10 induit significativement la transcription de l'oncogène E7 chez VPH-16 et suggère que cette cytokine peut causer une augmentation de la persistance et de la progression des lésions [3].

La possibilité d'une co-infection du VPH et du VIH au sein d'une même cellule n'était pas considérée jusqu'à tout récemment, car les deux virus infectaient des types cellulaires différents. Mais le doute pesant sur cette co-infection a enfin été levé. En effet, une étude récente a établi que le VIH pouvait infecter productivement les cellules épithéliales du col utérin [4]. Ainsi, il est maintenant envisageable qu'une interaction directe au sein d'une même cellule puisse intervenir entre le VIH et le VPH. En fait, cette hypothèse a déjà été examinée avec des résultats probants. Ainsi, en 85, Shwartz et al. infectèrent des cellules Hela (VPH-18 positive) avec le VIH-1 et observèrent que les transcrits du VPH-18 s'y accumulaient à un niveau 2 à 7 fois supérieur au niveau observé dans les cultures contrôles [141]. Aussi, il fut observé qu'une co-infection du VIH et du VPH-18 résultait en une activation des gènes E1 et L1 [27], suggérant une possible transactivation du VPH par le VIH.



## 4. LES VARIANTS CHEZ LES PAPILLOMAVIRUS

### 4.1 Phylogénie des VPHs

La définition de variant intratypique implique une homologie entre deux isolats d'au moins 98%. Cependant, différentes parties du génôme des VPHs évoluent à des rythmes différents. Ainsi, les domaines protéiques sont relativement bien conservés entre les différents types et nous permettent de comparer ceux-ci par alignement [16]. Par contre, la séquence impliquée dans la régulation de la transcription, le LCR, est plus variable et les séquences de LCR de différents types sont alignés avec plus de difficulté. Cette difficulté trouve notamment sa source dans la présence fréquente de délétions et d'insertions dans cette région. La variation génomique du LCR de deux isolats d'un même type peut atteindre 5%, par rapport à une diversité d'au maximum 2% pour les gènes codants [16]. Or, la séquence du LCR est suffisamment diversifiée à l'intérieur d'un même type pour permettre de reconstituer l'évolution moléculaire de ce type [15].

Plusieurs travaux ont été faits dans ce sens, sur VPH-16 et VPH-18, permettant ainsi de reconstituer l'évolution préhistorique de ces types, mais aussi de mettre cette évolution en relation avec le mouvement des anciennes populations humaines. Aujourd'hui, il est possible d'affirmer que l'*Homo sapiens* ou son ancêtre direct a évolué en Afrique et qu'il est apparu à l'extérieur du continent africain il y a 200 000 ans [114]. Or, il a été calculé à l'aide d'une estimation de la vitesse d'évolution des VPHs que leur âge phylogénétique devait être de cet ordre, suggérant une coévolution du virus avec les trois groupes ethniques majeurs chez l'homme, soit les Africains, les Caucasiens et les Asiatiques. Partant de l'hypothèse que les différents types de VPHs ont coévolué en association avec l'homme, on peut prédire que l'évolution moléculaire des VPHs indiquera un génôme précurseur d'origine africaine et que suivant les grandes migrations des populations humaines à partir de ce point, des groupes de variants

régionaux prendront forme avec l'apparition rare et sporadique de mutations, le VPH étant un virus à ADN évoluant beaucoup moins rapidement que les virus à ARN. Ainsi, des observations sur VPH-16 ont mené à la construction d'un arbre phylogénétique permettant la classification de ce type en cinq grandes branches régionales: deux branches de variants africains (Af1 et Af2), une branche Européenne (E) dérivant des variants africains, une branche asiatique (As) semblant avoir émergée de la branche européenne, et enfin une branche asiatique-américaine [68]. Pour appuyer ces faits, une étude similaire portant sur le VPH-18 a permis la création d'un autre arbre phylogénétique démontrant une évolution très similaire à celle du VPH-16 [114]. Ces arbres phylogénétiques montrent les mêmes caractéristiques principales, soit une petite diversité au sein d'un même groupe régional par rapport à une grande distance évolutionnaire entre les 5 grands branches continentales, suggérant une accumulation très lente de mutations sur une longue période de temps [15].

Considérant le fait que des types différents du VPH varient en affinité et en pathogénicité tissulaire et considérant la variabilité génétique au sein d'un même type, il est raisonnable d'envisager que cette variabilité intratypique peut se refléter dans des variations au niveau des propriétés biologiques de différents variants.

#### ***4.2 Variants et facteur de risque***

Des études expérimentales ont mis en évidence le fait que des altérations nucléotidiques pourraient affecter le potentiel oncogène du VPH-16 *in vitro* [103]. Certaines de ces altérations augmentent l'activité du promoteur gérant la transcription des oncoprotéines E6 et E7 [98]. Ces observations suggèrent que des variants naturels d'un type de VPH donné peuvent avoir des comportements biologiques différents dans une population. La persistance accrue de certains variants de VPH-16 est un exemple de ces comportements [37]. Dans un même ordre d'idées, les variants non-européens (NE) ont plus tendance à persister que les variants européens (E) [170].

Les études portant sur les variants se basent souvent sur les notions de variants «prototype-like» (PL) et «non-prototype-like» (NPL), les variants PL se définissant comme identiques au prototype défini ou ayant au maximum une mutation par rapport à celui-ci, les NPL regroupant tous les autres. Une de ces études s'est penchée sur la relation entre les variants du VPH-16 avec le risque de développer des lésions intraépithéliales de haut grade [178]. Les femmes infectées par des variants NPL étaient 6.5 fois plus à risque pour l'apparition de HSIL. De plus, ce risque était 4.5 fois plus élevé pour les femmes infectées par des variants NE (non-européen), le prototype étant un variant E (européen) [180].

À l'instar des HSILs, le cancer du col utérin lui-même montre des degrés d'association diverse selon les variants du VPH-16 [178]. Cela pourrait expliquer partiellement pourquoi certaines infections à VPH-16 progressent jusqu'au cancer alors que d'autres régressent. Ainsi, une étude démontre que l'infection aux variants NE a 11 fois plus de chance de progresser en cancer que l'infection par les variants E, ces derniers étant phylogénétiquement associés au prototype [64]. Ces données fournissent une évidence épidémiologique que différents variants de VPH-16 ont un potentiel oncogène différent.

Le risque accru observé chez les femmes porteuses de variants NPL et NE peut être redevable de différences dans le comportement biologique des variants et il est possible qu'il reflète un avantage de croissance des variants NPL sur les variants PL. Le phénomène pourrait être attribuable aux différences observées dans la régulation transcriptionnelle de différents variants, à l'habilité de l'hôte à offrir une réponse immunologique à des épitopes spécifiques encodés par les variants et à l'activité biologique des protéines encodées par le virus, telle qu'une capacité transformante accrue des oncoprotéines E6/E7.

### 4.3 Les variants du LCR

La fonction régulatrice du LCR suggère que celui-ci pourrait être responsable des différences observées dans l'infection à différents variants. Les changements nucléotidiques dans des sites permettant la liaison de facteurs de transcription cellulaire seraient tributaires de ces différences. Le déterminant majeur de la transformation cellulaire se trouve au niveau des régions LCR, E6 et E7. Romanczuk et al. [133] ont rapporté que l'ensemble LCR/E6/E7 du VPH-18 était plus efficace que celui du VPH-16 pour la transformation. Comparativement, les gènes E6 et E7 de ces deux VPHs, lorsque régulés par le même promoteur, immortalisent les kératinocytes avec la même efficacité, suggérant que les différences observées dans l'immortalisation n'étaient pas dues à des différences dans les gènes E6 et E7 eux-mêmes, mais plutôt à des différences dans la région régulatrice du LCR [133]. Le fait que le LCR des variants asiatiques-américains semble avoir une activité transcriptionnelle accrue par rapport aux variants européens va dans le sens de cette hypothèse [77]. Dans ce cas, des altérations génétiques du LCR causeraient une sur-expression de E6 et E7, avec une baisse conséquente de l'efficacité de p53 et de pRb. Par ailleurs, la diversité génétique du LCR des variants européens est plutôt faible par rapport aux variants des autres branches [87].

Chez le VPH-16, des délétions de longueurs variables ont été trouvées. Il a été rapporté que des variants portant des délétions dans la région 5' ou dans la région centrale du LCR montraient des activités transcriptionnelles accrues ou plus diminuées par rapport au VPH-16 prototypique [50]. Des délétions se retrouvent aussi fréquemment dans la régions 3', contenant la majeure partie des activateurs viraux ainsi que le promoteur précoce, là où se fixent plusieurs facteurs de transcription, notamment YY1, connu pour être un suppresseur transcriptionnel. Des altérations génétiques au niveau de tels sites de liaison peuvent résulter en l'abolition de la liaison ou en une augmentation ou diminution de l'affinité de la liaison. Ainsi, une délétion touchant un des sites de liaison de YY1 cause une augmentation conséquente de l'activité transcriptionnelle du LCR [29]. D'autres auteurs ont confirmé ces résultats après avoir isolé des mutants ayant une substitution au niveau d'un site YY1 [117]. Ces mutants

voyaient l'activité de leur promoteur p97 grimper de 2 à 4 fois. De plus, une mutation particulière, la transition G-A en position 7841, était retrouvée dans la plupart des cas de néoplasies du col utérin. Une étude portant sur le VPH-18 a mis au jour une substitution T à C (nt 104) au niveau du promoteur p105, substitution touchant aussi un site YY1, la séquence prototype ayant une liaison deux fois plus forte avec ce facteur de transcription [134]. Les auteurs de l'étude ont aussi noté que les variants portant cette mutation se retrouvaient à une forte fréquence dans les spécimens de cancer.

À la lumière de ces observations, il est plausible que des variations intratypiques au sein de la région régulatrice du LCR peuvent avoir une influence fonctionnelle dans la pathogénicité des VPHs. Cependant, certains auteurs soulèvent le point que chez les variants européens la variabilité de cette région est moindre, les changements de séquences touchant rarement des sites de liaisons à des facteurs de transcription [87]. De plus, dans une étude portant sur les variants européens, aucun variant spécifique du LCR ne fut associé à la progression des lésions, alors qu'un variant E6 se retrouvaient significativement plus souvent dans les HSILs et les spécimens cancéreux [76]. Ces états de fait supportent le concept selon lequel la variation génomique au niveau du LCR n'est pas une stratégie commune à toutes les branches phylogénétiques pour la promotion du potentiel oncogène viral et que sous cette optique, les variations au niveau des ORFs E6 et E7 seraient plus significatives. Cependant, il faut aussi considérer le phénomène de coségrégation des mutations des différentes régions du génôme. En effet, s'il y a coségrégation des mutations du LCR, de E6 et de E7, l'association qui fut observé entre des mutations du LCR et le cancer ne pourrait qu'en être le reflet.

#### ***4.4 Les variants de E6 et E7***

De nombreuses observations suggèrent que les variants E6 des VPHs possèdent des activités biologiques et biochimiques différentes. D'abord, des substitutions d'acides aminés près de la région amino-terminale de E6 sont connues pour altérer l'activité de dégradation de p53 [20]. En fait, il semble que les variants naturels de E6 diffèrent par

deux activités fonctionnelles importantes : leur habileté à coopérer avec la protéine E7 pour l'immortalisation et la transformation cellulaire et leur habileté à cibler la protéine p53 pour la dégradation [151]. En effet, le pouvoir immortalisant et transformant dépend d'une coopération des deux oncoprotéines du VPH. Par exemple, chez VPH-16 un variant de E7 (N29S) fut décrit comme ayant une activité deux fois plus forte que le prototype pour la transformation de fibroblastes de rats [43]. Cependant, les auteurs de l'étude ont remarqué que ce variant pouvait se trouver en combinaison avec la protéine E6 d'un variant particulier (z84), qui a une activité plus faible que le prototype pour la transformation [151], et que la combinaison des deux variants résultait en une activité transformatrice globale moindre que pour les variants proprement prototypiques.

À la lumière de ces résultats *in vitro*, il est tentant de croire que des variants de E6 ou E7 seraient associés au développement de lésions malignes. Pour la protéine E6 du VPH-16, un variant européen a effectivement été décrit comme ayant une relation avec l'apparition de HSILs et de cancers [184]. Ce variant porte une substitution d'acide aminé, une leucine pour une valine au résidu 83 (L83V). Par contre, une étude portant sur une population de femmes russes a obtenu un résultat contraire : ce variant E6, pas plus qu'aucun autre d'ailleurs, n'avait de relation avec un potentiel oncogène accru [71]. L'équipe du Dr Zehbe a tenté d'apporter la lumière sur ce variant en particulier. Dans un premier temps, ils ont établi une corrélation statistique nette entre le variant L83V et la sévérité des lésions du col utérin dans une population exclusivement scandinave provenant de la Finlande, de la Suède et de la Norvège [184]. Dans un second temps, une autre étude s'est penchée sur des populations italienne et tchèque et contredisait les résultats provenant de Scandinavie : le variant L83V n'était pas relié au développement de la maladie, cependant que d'autres variants y étaient associés [183]. En fait, ces contradictions inter-population permettent de souligner les faits suivants: 1) le potentiel oncogène des différents variants E6 n'est peut-être pas tant lié à des différences fonctionnelles entre ces variants mais tiendrait plutôt au fait qu'il y a des différences immunogénétiques entre des populations distinctes; 2) il y a des populations plus à risque pour l'infection par certains variants. Étant donné que la persistance du VPH est

un pré-requis majeur pour le développement de lésions malignes, le polymorphisme des gènes du HLA («Human Leukocyte Antigen») pourrait être un bon candidat pour de telles différences. Ainsi, il a été rapporté que l'haplotype DRB1\*04-DQB1\*03, qui est approximativement trois fois plus abondant dans la population suédoise qu'en Italie ou qu'en République Tchèque, était positivement associé avec le variant L83V du VPH-16 [183]. Une étude plus poussée portant sur les gènes HLA de classe I a permis de conclure que certains allèles, en concert avec des variants spécifiques du VPH-16, pouvaient présenter des prédispositions au développement du cancer, alors que d'autres avaient un effet protecteur [182].

Cependant, il existe une autre interprétation des divergences génétiques inter-populations en relation avec les variants du VPH. Ainsi, l'oncoprotéine E6 interagissant principalement avec p53, des équipes se sont penchées sur le polymorphisme du gène codant pour ce suppresseur de tumeurs. Ces études ont porté sur le polymorphisme de celui-ci au niveau du codon 72, lequel peut coder soit pour une arginine ou une proline, les deux allèles variant selon les populations [6]. Or, le variant L83V du VPH-16 était significativement sur-représenté chez les femmes homozygotes pour p53 Arg ayant le cancer du col utérin, suggérant que chez les femmes p53 Arg/Arg, l'infection avec ce variant particulier conférait un plus grand risque pour le cancer [164]. En fait, il a été démontré dans une importante étude que la forme Arg du p53 était plus susceptible à la dégradation *in vitro* [152]. La même étude indique que les individus homozygotes pour l'arginine en position 72 sont 7 fois plus à risque que les individus hétérozygotes pour le développement de tumeurs associées aux VPHs oncogènes. L'allèle p53 Arg constituerait donc aussi un facteur de risque significatif pour l'apparition du cancer associé aux VPHs.

En ce qui a trait à l'oncoprotéine E7, on peut aussi croire que des variations intratypiques au sein de celle-ci aura des effets sur ses fonctions biologiques. Il a été observé dans des expériences de mutagenèse que certaines mutations à des sites très spécifiques, tel que le domaine de liaison à pRb, avaient des effets négatifs sur la

capacité de cette protéine à immortaliser des cellules *in vitro* [5]. Cependant, ces mutations n'ont encore jamais été observées *in vivo*. D'autre part, aucune corrélation claire entre les variations intratypiques de E7 et la pathogénicité des VPHs n'a été établie. D'ailleurs, il a été rapporté que E7 était moins variable que E6 [113]. Cette moindre variabilité pourrait trouver son explication dans la sélection imposée par le système immunitaire. En effet, il semble que la réponse T-cytotoxique à E6 est beaucoup plus importante que celle à E7 lors de la contre-attaque du système immunitaire suivant l'infection au VPH-16 [183]. Nakagawa et ses collègues ont observé que l'infection au VPH disparaissait chez les femmes montant une réponse CTL contre E6, alors que pour les femmes chez qui l'infection était persistante, cette réponse n'était pas déclenchée [108]. De plus, la réponse CTL à E7 ne montrait pas une telle corrélation, car cette réponse était détectée chez les femmes ayant des lésions progressives aussi bien que régressives. Ainsi, au cours de son évolution, la protéine E7 n'aurait pas été soumise à une pression sélective aussi forte.

#### ***4.5 Autres variants***

Le cadre de lecture E2 est un bon candidat pour l'étude de variants, étant donné le rôle régulateur de la protéine codée par celui-ci pour la transcription des gènes précoces. La dite protéine est constituée d'un domaine d'activation transcriptionnelle à son extrémité N-terminale, d'une région charnière centrale et d'un domaine de liaison à l'ADN en C-terminal [47]. Des expériences de mutagenèse dirigée ont permis de démontrer que des mutations au niveau de la région trans-activatrice pouvaient induire un potentiel d'immortalisation et de d'activation accru [132]. Cependant, les études portant sur les variants naturels de E2 sont rares. Une de celles-ci a démontré que différents variants E2 du VPH-16 ne différaient que très légèrement entre eux au niveau de l'activité transcriptionnelle, et que les différences observées semblaient être indépendantes de la position taxonomique des isolats [167]. Plus récemment, on a isolé un variant E2 portant une mutation au niveau du domaine de liaison à l'ADN [46]. Ce



variant est positivement associé aux HSILs et pourrait avoir une importance pour la progression des lésions.

Pour la protéine E4, il semblerait que les mutations soient rarissimes. L'absence de variation dans la région amino-terminale de la protéine, qui est responsable de la destruction du réseau de cytokératine, suggère que cette région de E4 est intolérante à tout changement d'acides aminés [173]. Le gène E4 ne serait donc pas un bon candidat pour l'étude de variants.

Le gène E5 fut aussi étudié. Ce dernier est connu pour être le plus variable des gènes chez les VPHs [16]. Bien que les domaines fonctionnels de la protéine n'aient pas encore été définis, une de ces études est parvenue à établir qu'un variant particulier de E5 était associé à un facteur de risque accru pour le développement de néoplasies du col utérin [7].

Enfin, des variations au niveau des protéines de structure L1/L2 pourraient avoir grande une influence sur le potentiel infectieux de différents variants. Évidemment, cette influence serait tributaire de la réponse immunitaire à ces deux protéines. Les études portant sur ces protéines se tournent donc vers cet aspect de l'infection aux VPHs. Une de celles-ci a notamment observé que le sérotype de différents variants L1 de VPH-16 était le même [123], ce qui fait de cette protéine une excellente cible pour une vaccination effective. De fait, des vaccins prophylactiques basés sur des VLPs constituées de la protéine L1 sont en essai clinique, le but principal de ces vaccins étant de réduire à long terme l'incidence du cancer du col utérin et de ses lésions précurseurs [85].

## OBJECTIFS ET HYPOTHÈSES

En comparaison avec les VPH-16 et 18, très peu d'études ont porté sur le polymorphisme des VPHs d'autres types, tel qu les types 31, 33 et 35. En fait, les variants E6 du VPH-33 ont déjà été abordés de façon peu exhaustive dans une étude portant aussi sur les types 52 et 58 [179]. À notre connaissance, aucune étude de variants n'a été publiée sur les types 31 et 35. Ces types sont phylogénétiquement très près du VPH-16, qui est sans doute le type le plus étudié et le mieux caractérisé, et du VPH-52, qui a déjà été décrit dans une étude précédente menée dans notre laboratoire. Ces deux derniers types sont très polymorphiques et de nombreux variants de ceux-ci ont été décrits. De plus, une association entre leur polymorphisme et la persistance a déjà été établie.

De par leur proche parenté avec les VPH-16 et 52, il nous est permis de croire que les VPH-31, 33 et 35 démontreront aussi une grande variabilité et que de cette variabilité sera tributaire des différences notables dans la capacité de persistance de différents variants, la persistance étant un déterminant majeur du développement de lésions pré-cancéreuses du col utérin.

*preuve  
à  
revoir*

*les différences notables au niveau de  
la persistance, sur le résultat de la  
variabilité génétique*

Donc, dans un premier temps, notre objectif sera de décrire les polymorphismes des VPH-31, 33 et 35 et de les comparer entre eux au niveau des régions LCR, E6 et E7. La région LCR étant la plus variable chez les VPHs étudiés jusqu'ici, son séquençage devrait permettre une classification précise des variants. Cependant, de par sa variabilité même, les régions LCR de différents types sont peu susceptibles de pouvoir être alignées et comparées. Ainsi, nous pourrions nous servir à cette fin des variants obtenus pour les régions codantes E6 et E7.

Dans un second temps, nous évaluerons l'association épidémiologique entre la persistance virale et les variants eux-mêmes. Nous tenterons de voir si les variants non-

prototypiques sont plus à même de persister, comme c'est le cas pour VPH-16 et 52. Nous espérons aussi trouver des mutations significatives dans le LCR (mutations touchant des sites de liaison aux facteurs de transcription), dans E6 et dans E7, mutations qui pourraient avoir un impact sur la persistance virale.

**ARTICLES**

**Déclaration de l'étudiant concernant sa participation aux 2 articles inclus dans ce mémoire:**

J'ai participé au développement du projet, notamment aux niveaux de la sélection des échantillons à analyser, de la conception des diverses paires d'amorces et de l'optimisation des conditions de PCR. J'ai effectué toutes les manipulations, à la suite desquelles j'ai participé au traitement et analyse des résultats puis à la rédaction des articles.

## **Article 1**

### **Polymorphism of Human Papillomavirus type 31 isolates infecting the genital tract of women who have or who are at risk of having HIV infection.**

Running Head: HPV-31 polymorphism in HIV-infected women

Simon Gagnon<sup>1,2</sup>, Catherine Hankins<sup>3,4</sup>, Cécile Tremblay<sup>2,5</sup>, Pierre Forest<sup>1</sup>, Karina Pourreaux<sup>4</sup>, Fabrice Rouah<sup>4</sup>, The Canadian Women's HIV Study Group, and François Coutlée<sup>1,2,5</sup>.

1- Laboratoire de Virologie Moléculaire, Centre de Recherche du Centre Hospitalier de l'Université de Montréal, Montréal, Qué., Canada.

2- Département de Microbiologie et Immunologie, Université de Montréal, Montréal, Qué., Canada.

3- Departments of Epidemiology , Biostatistics and Occupational Medicine, and Oncology, McGill University, Montreal, Que., Canada.

4- Direction de la Santé Publique de Montréal-Centre, Institut National de Santé Publique du Québec, Montréal, Qué., Canada.

5- Département de Microbiologie et Infectiologie, Centre Hospitalier de l'Université de Montréal, Montréal, Qué., Canada.

The Canadian Women's HIV Study Group includes the following investigators from across Canada: Principal investigator: Catherine Hankins. Halifax: Janet Conners\*, Rob Grimshaw, David Haase, Lynn Johnston\*, Wally Schlech\*, Arlo Yuzicappi-Fayant\*. Hamilton: Stephen Landis, Fiona Smaill. London: Tom Austin, Ole Hammerberg, Ted Ralph\*. Montréal: François Coutlée\*, Julian Falutz, Alex Ferenczy, Marina Klein\*, Louise Labrecque, Normand Lapointe, Richard Lalonde, John Macleod, Grégoire Noël\*, Chantal Perron\*, Jean-Pierre Routy, and Emil Toma. Ottawa: Claire Touchie, Garry Victor\*. Québec: Louise Côté, Hélène Senay, Sylvie Trottier. Saskatoon: Kurt Williams\*. Sherbrooke: Alain Piché\*. Sudbury: Roger Sandre. Toronto: Louise Binder\*, Donna Keystone, Anne Phillips, Anita Rachlis\*, Irving Salit, Cheryl Wagner, Sharon Walmsley\*. Vancouver: Paula Braitstein\*, David Burdge\*, Marianne Harris, Deborah Money\*, Julio Montaner\*. \* study co-investigators.

Address correspondence to: François Coutlée, Département de Microbiologie et Infectiologie, Hôpital Notre-Dame du Centre Hospitalier de l'Université de Montréal, 1560 Sherbrooke est, Montréal (Québec), H2L 4M1, Canada. Tel. 514-890-8000, 25162; fax: 514-412-7512

e-mail: [francois.coutlee@ssss.gouv.qc.ca](mailto:francois.coutlee@ssss.gouv.qc.ca).

All participants provided written informed consent to participate. Ethics committees of each participating institution approved the Canadian Women's HIV Study protocol.

This study was supported by CANFAR. The Canadian Institutes of Health Research supports The Canadian Women's HIV Study. F.C. is a chercheur national supported by the Fonds de la Recherche en Santé du Québec and by le Réseau FRSQ-SIDA.



**Abstract (248 words)**

The genomic polymorphism of high-risk human papillomavirus (HPV) genotypes other than 16 has not been extensively investigated. The polymorphism of HPV-31 is described in a cohort of sexually active Canadian women. HPV-31 isolates detected in cervicovaginal lavages collected from 62 HIV-seropositive and 23 HIV-seronegative women were further analysed by PCR-sequencing of the long control region (LCR), E6 and E7 genes. Twenty-five nucleotide variation sites over 399 nucleotides in the LCR, 13 variation sites over 496 nucleotides in E6 and 9 variation sites over 296 nucleotides in E7 defined 18, 10 and 8 variants, respectively. The median number of variation sites per E6 or E7 variants (five sites per variant) compared to the prototype was significantly smaller than the number of variation sites per variant in the LCR (ten sites per variant) ( $p < 0.01$  for both comparisons). Non-synonymous variations were not randomly distributed in E6 and E7. Non-synonymous E6 mutations were not located in putative sites for oncoprotein activity but three variations were within cysteine arrays. The LCR prototype was significantly overrepresented in Caucasian women (14 (25%) of 56) compared to women of African descent (0 (0%) of 15 women,  $p = 0.03$ ). LCR variant 31.4 was detected in 8 (53.3%) of 15 women of African descent and in 2 (3.6%) of 56 Caucasian women ( $p = 0.003$ ). Non-prototypic variants did not increase the risk for persistent HPV-31 infection. HPV-31 polymorphism was greater in the LCR than in the oncoprotein genes, and was related to the ethnic origin of women.

## Introduction

High-risk human papillomaviruses (HPV) types are associated with squamous intraepithelial lesions (SIL) and cancer of the uterine cervix [1]. HIV-seropositive women are at greater risk than HIV-seronegative women for HPV infection, persistence of HPV infection and cervical SIL [2-5]. Nearly 80% of HIV-infected women followed prospectively over two years became infected at the uterine cervix by HPV [6]. Furthermore, the rate of incident HPV infection is reported to be at least three times higher in HIV-seropositive women [7]. Nearly 20% of HIV-seropositive women will develop SIL, those with low T lymphocyte CD4+ cell counts being at highest risk [8,9].

Mucosal HPV types are classified into low-risk types that are found predominantly in benign lesions, and high-risk types that are associated with pre-malignant or malignant lesions [10]. HPV-31 has been classified as a high-risk type based on phylogenetic analysis of HPV sequence alignments and on epidemiological studies [10-12]. HPV-31 immortalizes epithelial cells and human keratinocytes [13,14]. HPV-31 causes 4% of cervical cancers worldwide [10]. HPV-31 is the fourth most common type in cervical cancer after types 16, 18 and 45 [10]. Up to 6% of HIV-seropositive women with cervical LSIL or HSIL are infected by HPV-31 [15]. However, HPV-31 is also frequently detected in women without disease. Up to 6% of HPV-infected women without cancer harbor HPV-31 DNA sequences [4,10]. Type 31 has a high incidence rate in HIV-seronegative (1.43 per 100 person-years) and in HIV-seropositive women (2.44 per 100 person-years). In one study on incident HPV infection

in HIV-seronegative women, type 31 was in second place compared to other oncogenic types [5].

HPV-16 polymorphism of the long control region (LCR) and of the E6 open reading frame (ORF) have been associated with HPV persistence and anogenital SIL [16,17]. However, the genomic variations of high-risk type 31 have never been investigated [16]. A better understanding of the impact of mutations which define variants and confer a greater risk for the development of cancer could help explain why only a subset of women infected by HPV develop cervical cancer.

This work is the most extensive description up to now of HPV-31 polymorphism. The polymorphism of the regulatory region and transforming open reading frames of HPV-31 is described in a population of Canadian women infected or at risk of being infected by HIV. The association between HPV-31 polymorphism and ethnic origin of participants was investigated. The role of HPV-31 polymorphism in persistence of infection was assessed.

## **Materials and methods**

*Study population and study design.* Subjects were selected from participants in the Canadian Women's HIV Study infected in at least one specimen with HPV type 31. The latter study investigates determinants of HPV and persistence of HPV infection in HIV-seropositive women. The original study design and characteristics of the cohort have been described elsewhere [3,19]. In the Canadian Women's HIV Study, 1055 women were recruited from 1993 to 2000 across Canada from outpatient sexually transmitted

diseases (STD) clinics, family practices or infectious disease clinics involved in HIV and STD management [3,19,20]. Women were eligible to participate to the study if they were seropositive for HIV-1 or if they were seronegative for HIV but at risk for STD and accepted to have annual serological tests for HIV antibodies [3,19]. All participants provided written informed consent to participate. Ethics committees of each participating institution approved the Canadian Women's HIV Study protocol.

For all participants, a cervicovaginal lavage using 10 ml of phosphate-buffered saline [pH 7.4] was obtained at least once during pelvic examination [20]. Specimens were refrigerated within one hour and transported to a central laboratory on wet ice. A standardized questionnaire was administered upon study entry and six monthly thereafter [3]. For all HIV-seropositive women, T lymphocyte CD4 blood counts, a Pap smear obtained with a cytobrush and Ayre spatula, vaginal tampon specimens and cervicovaginal lavages were obtained and processed as described previously at inclusion and at 6-month intervals [20]. For HIV-seronegative women, vaginal tampons were obtained at 6-month intervals while cervicovaginal lavages and Pap smears were collected at one-year intervals.

*Sample processing.* Cell suspensions were pelleted after centrifugation at 2,500 rpm for 10 min at 4°C and resuspended in 500 µl of 10 mM Tris [pH 8.3]. Cell suspensions were lysed with Tween 20 at a final concentration of 0.8% (v/v), and digested with 250 µg/ml of proteinase K for 2 hours at 45°C. Cell lysates were boiled for 10 min and stored at -70°C until tested. The delay between sampling and processing never extended beyond 5 days.

*β-globin and HPV DNA detection.* An aliquot of 5 microliters from each sample was amplified for β-globin DNA with PC04 and GH20 primers to control for DNA integrity, for the presence of an adequate number of epithelial cells, and for amplification inhibitors. β-globin-negative samples were extracted with phenol-chloroform and precipitated with ethanol. Five hundred ng of extracted nucleic acids were then amplified for β-globin. β-globin-positive samples (lysate or extracted DNA) were tested for the detection and typing of HPV DNA using the MY09-MY11-HMB01 consensus L1 PCR as described previously [20]. Negative, weak positive (ten HPV18 DNA copies), and strong positive controls (HPV types 6/11, 16, 31, 33, 35, 39, 45) were included in each amplification run to monitor contamination and overall endpoint sensitivity of each PCR run. The amplification mixture contained 6.5 mM MgCl<sub>2</sub>, 50 mM KCl, 2.5 units of *Taq* DNA polymerase (*AmpliTaq*, Roche Molecular Diagnostics, Mississauga, Ont.), 200 μM each of dATP, dCTP, dGTP, dTTP, and 50 pmoles of each primer. Amplifications were performed in a TC9600 Thermocycler (*AmpliTaq*, Roche Molecular Diagnostics, Mississauga, Ont.), for 40 cycles with the following cycling parameters: 95°C for 1 min, 55°C for 1 min and 72°C for 1 min. Amplified products were spotted onto nylon membranes and reacted under stringent conditions with <sup>32</sup>P-labelled oligonucleotide probes for types 6, 11, 16, 18, 31, 33, 35, 39, 45, 51, 52, 53, 56 and 58 [20,21].

*Amplification of HPV-31 LCR, E6 and E7 for molecular variant analyses.* Several HPV-31-specific primer pairs were designed to amplify the 3' end of LCR, the complete E6 and E7 genomic segments of isolates detected in cervicovaginal lavages for PCR-

sequencing analysis. A 434 bp fragment at the 3' end of the LCR and (LCR-31-sense, location nucleotide (nt) 7381-7403, 5'-TTAGGTGTCACGCCATAGTAAAA-3', and LCR-31-antisense, location nt 7795-7814, 5'-CATGACACAACCTTGGCAGT-3'), a 641 bp fragment (E6-31-sense, location nt 40-57, 5'-CGAAAGTGGTGAACCGAA-3' and E6-31-antisense, location nt 659-680, 5'-GACTGTCTATGACATCCTCCTC-3') containing the complete 450 bp E6 open reading frame and a 542 bp fragment (E7-31-sense, location nt 510-529, 5'-CGTTGCATAGCATGTTGGAG-3' and E7-31-antisense, location nt 1032-1051, 5'-CTGCTTCTGCCTGATTGTTG-3') containing the complete 296 bp E7 open reading frame, were selected from the published HPV-31 sequence [22]. The E6 and E7 amplicons overlapped over 171 bp.

Amplification reactions for each HPV-31 region were performed with 2 µl of processed sample in a 100 µl reaction volume containing 10 mM Tris-HCl [pH 8.3], 50 mM KCl, 2.5 units of *AmpliTaq* Gold PCR enzyme (*AmpliTaq*, Roche Molecular Diagnostics, Mississauga, Ont.), 0.5 µM of each primer, 0.25 mM each dCTP, dTTP, dGTP and dATP. MgCl<sub>2</sub> concentration was adjusted at 2.0 mM for LCR and E6 primer pairs and at 1.5mM for E7 primers. Amplifications were completed in a 9600 Thermal Cycler (Roche Diagnostic System, Mississauga, Ont.). The amplification profiles for the LCR reagents included an initial step at 94°C for 120 sec followed by 10 cycles at 94°C for 15 sec, at 55°C for 30 sec and at 72°C for 45 seconds. Amplification was done with 35 cycles at 94°C for 15 sec, at 55°C for 30 sec and at 72°C for 45 seconds. Each cycle was increased by five seconds at the extension step. Amplification was completed by an extension step at 72°C for 7 min. The amplification of E6 DNA sequences was initiated

at 94°C for 120 sec followed by 10 cycles at 94°C for 15 sec, at 60°C for 30 sec and at 72°C for 45 seconds. Amplification was accomplished by 35 cycles at 94°C for 15 sec, at 60°C for 30 sec and at 72°C for 45 seconds. Each cycle was increased by five seconds at the extension step. Amplification was completed by an extension step at 72°C for 7 min. The amplification profile for E7 reagents included an initial step at 94°C for 120 sec followed by 10 cycles at 94°C for 15 sec, at 63°C for 30 sec and at 72°C for 45 seconds. Amplification was completed by 35 cycles at 94°C for 15 sec, at 63°C for 30 sec and at 72°C for 45 seconds. The extension step of each cycle was increased by five seconds. Amplification was completed by an extension step at 72°C for 7 min.

*Molecular variant analysis by PCR-sequencing.* PCR-amplified HPV-31 DNA fragments were purified with the QIAquick PCR purification kit (Quiagen Inc., Mississauga, Ont) when a single band was visualized after gel electrophoresis or with the QIAquick gel extraction kit protocol (Quiagen Inc.) when several bands were visualized. Direct double-stranded PCR-sequencing was done since most amplified products contain the appropriate nucleotide at any given position and the derived sequence will be a true representation of the HPV DNA template [24]. Twenty ng of purified amplicons were sequenced with the fluorescent cycle-sequencing method (BigDye terminator ready reaction kit, Perkin-Elmer) with 25 cycles at 96°C for 10 sec, 50°C for 5 sec and 62°C for 4 min. Sequence analysis was performed on an ABI Prism 3100 Genetic Analyzer system at the Laboratoire d'Épidémiologie moléculaire, CHUM. If a non-prototype variant or ambiguities were obtained, PCR-sequencing was repeated once. If results between these 2 sequencing reactions were discrepant, the lysate was

sequenced three times to investigate the presence of multiple variant infections or *Taq*-induced errors. Unconfirmed mutations were considered as PCR artifacts. In contrast, variants had mutations in non-random nucleotide positions.

*Data analysis.* To classify our isolates initially, HPV-31 LCR DNA sequences were compared using the BLAST sequence analysis software from the Genetic Computer Group [25] to sequences of known HPV-31 variants and the prototype (<http://hpv-web.lanl.gov>) and to the complete GenBank database [11,22]. Isolates with a DNA sequence different from the prototype were classified as non-prototype variant [26]. Sequences from non-prototype variants were aligned for classification using Clustal W 1.8 [27]. The multiple alignments were further refined by manual intervention. To assess the effect of variations in the LCR on putative binding sites for cellular proteins, we used the TFSEARCH [28]. The TFSEARCH searches highly correlated sequence fragments against TFMATRIX transcription factor binding site profile database in the 'TRANSFAC' databases by GBF-Braunschweig.

Proportions were compared using the z statistical test. Categorical variables were compared with the Pearson's chi-square test or the Fisher's exact test, depending on the sample size. The number of variations per variant or per isolate was compared using a non-parametric method, the Mann-Whitney rank sum test. In these comparisons, the number of variations was adjusted for the length of amplicon as described in the result section. Double-sided p values < 0.05 were considered significant. The Mann-Whitney rank sum test was used to compare years since immigration between immigrant women infected by prototype and by non-prototype variants.



## Results

Overall, 732 HIV-seropositive and 323 HIV-seronegative women were screened for cervical infection by HPV in the Canadian Women's HIV Study. Combining prevalent and incident cases of HPV-31 infection, 79 (7.5%) of 1055 participants had one or more genital specimens containing HPV-31 sequences DNA detected by consensus PCR. Of the 487 genital samples provided throughout the study by these HPV-31-positive 79 participants, 13 (2.7%) were inadequate for PCR analysis. The latter samples had no impact on molecular variant analysis and on establishing persistence HPV-31 infection defined as detection of HPV-31 in at least three consecutive samples for more than nine months (see below). Five (2.2%) of the 225 cervicovaginal lavages and 8 (3.1%) of 262 vaginal tampon specimens were inadequate for PCR analysis ( $p = 0.57$ ). HPV-31-positive women provided a median of 6 specimens (mean of  $6.4 \pm 4.2$ , range of 2 to 17) collected at a mean of  $3.6 \pm 2.2$  visits (median of 4 visits, range of 1 to 8 visits). Overall, 62 (8.5%) of the 732 HIV-seropositive women and 17 (5.3%) of the 323 HIV-seronegative were infected by HPV-31 (Odds ratio 1.67, 95% confidence interval 0.96-2.90,  $p = 0.09$ ).

Women infected by HPV-31 had a median age of 31 years (mean of  $31.4 \pm 9.1$  range of 17 to 63 years). The mean and the median CD4 counts closest to the first HPV-31-positive sample were measured for 61 HIV-seropositive women and reached  $370 \pm 281 \times 10^6$  and  $298 \times 10^6$  cells/L, respectively (range 6 – 1297 cells/L). Of the 61 HIV-infected women who completed the study questionnaire, 23 (37.8%) had a

diagnosis of AIDS (13 women were in clinical category C of HIV infection, 4 women in clinical category B and 6 in clinical category A). Of the 38 HIV-seropositive women without AIDS, 11 were in clinical category B and 27 had asymptomatic HIV infection with CD4 counts above 200 cells per  $\mu\text{l}$  of blood.

*Polymorphism of the HPV-31 LCR.* The 3' end of the LCR of the first isolate from each participant was analyzed with PCR-sequencing. HPV types 16, 18, 31, 33, 35, 45, and 66 did not react with the LCR primer set (data not shown). The frequency and nature of variations defining each of the 17 non-prototype variants compared to the prototype (for a total of 18 variants) are displayed on Figure 1. Each variant contained from 2 to 13 variations (mean of  $9.4 \pm 3.3$ , median of 10) compared to the prototype. Isolates 31-LCR-16 or 31-LCR-17 and the prototype were the most divergent variant (Fig. 1). Excluding primer sequences from the LCR amplicon sequence, there was a total of 13 (3.3%) variation sites over 391 bases between the prototype and the latter variants. The LCR polymorphism was generated by 25 single-nucleotide point variations over 391 bases (bases 7404 to 7794). Thus 6.4% of nucleotide positions were implicated in HPV-31 polymorphism. We did not encounter single-nucleotide or fragment deletion or insertion in the 79 isolates (Fig 1). The most frequent variation sites at positions 7477 and 7457 were shared by all non-prototypic variants while sites 7449, 7525 and 7575 were shared by 15, 15 and 14 variants, respectively (Fig. 1). Only seven (28%) of the 25 single-nucleotide variations were found in only one variant while the other variations were shared by at least two HPV-31 variants.

7474?

We assessed the effect on the putative binding sites for cellular proteins of each variation in the LCR. Loss or addition of putative binding sites for cellular factors were found in 10 of the 25 variation sites (Fig. 1). The variation at nucleotide 7474 resulted in the putative loss of a binding site for the E2 protein that was however gained by variation at nucleotide 7477. Several frequent variations resulted in the loss of a putative binding site for catabolite activator protein (cap) (Fig. 1). The variation at site 7659 resulted in the loss of a putative binding site for C/EBP as we have described in HPV-52 [29,30]. Only 4 variants had this variation for a total of 17 isolates (Fig 1). The other variations introduced additional putative binding sites for cellular proteins such as TATA, Brn-2, AML-2a and GATA or resulted in a loss for binding SRY. The impact of variations on putative binding sites of cellular factors for HPV-31 was less than what reported for type 52 [29].

The prototype was the second most frequent variant being detected in 14 isolates. The LCR-31-3 was the most frequent of the 18 LCR variants and was identified in 20 (25.3%) of 79 isolates (Fig 1). Twelve HPV-31 variants were detected in at most two participants. The evolutionary relationship among the 18 HPV-31 variants was explored by subjecting the sequence data to phylogenetic analysis (Fig. 2). HPV-31 variants clustered into lineages.

We investigated the relationship between ethnicity and HPV-31 LCR variations for women for whom the questionnaires were completed. Of the 79 participants, 56 (70.9%) were Caucasians (49 women were born in North America, 4 in Europe, 2 in South America, 1 in Asia), 15 (19.0%) women were of African descent (9 women born

in Haiti, 4 in Africa, 1 in North America and 1 in Europe), 4 (5.1%) were Asian and 2 (2.5%) were Aboriginal women. The ethnicity for two women was not obtained. All women were living permanently in Canada when recruited into the study. We first evaluated the association between variant classification and ethnic origin by considering only Caucasian women and women from African descent as well as variants that infected at least 5 participants including the prototype (n=14), 31-LCR-3 (n=17), 31-LCR-4 (n=11), 31-LCR-13 (n=10), and 31-LCR-15 (n=5). The distribution of these variants is presented in Table 1. All prototype isolates were detected in Caucasian women. The prototype and variant 31-LCR-3 were significantly over represented in Caucasian women while variant 31-LCR-4 was detected in more than half of women from African descent ( $p \leq 0.05$  for each comparison). Two 31-LCR-13 isolates were detected in women from African descent born in North America. We could not analyse the association between country of birth and each variant because of the small distribution of participants born in Europe (n=5), Latin America (n=3), Africa (n=4) and Asia (n=3).

Women from African descent were more frequently infected with the LCR variant 31-LCR-4 variant. The 12 women infected by the latter variant were born in Haiti (n=6), African countries (n=2), Canada (n = 2), Asia and unknown (n = 1 each). Six (66.7%) of 9 women born in Haiti, 2 (3.8%) of the 53 women born in Canada and 0 (0%) of the 3 women born in Latin America were infected by variant 31-LCR-4 variant ( $p < 0.0001$ , Chi-square test). When the rate of detection of variant 31-LCR-4 variant was compared by region of birth, a significant difference was found, after the Bonferroni

correction for a total of 3 comparisons (setting a significant p value below 0.017), between women born in Haiti and women born in Canada ( $p=0.0001$ ) but not between women born in Haiti and Latin America ( $p = 0.18$ ) or between women born in Canada and Latin America ( $p = 1.0$ ). When the detection rate of the latter variant in women born in Canada was compared to the detection rate in women who had immigrated to Canada, the difference was also statistically significant (2 (3.8%) of 53 versus 9 (37.5%) of 24,  $p = 0.0003$ ). Women who were born outside of Canada and were infected with variant 31-LCR-4 ( $n=9$ ) had been living in Canada for  $7.3 \pm 4.1$  years (median of 8.5 years, range of 1 to 12 years), had been sexually active in Canada ( $n=8$ ) for  $13 \pm 7$  years (median of 16 years, range of 5 to 18 years) and had been sexually active before immigration ( $n=3$ ) for  $7.1 \pm 5.6$  yr. (median of 6 years, range of 1 to 19 years). When women from African descent infected by variant 31-LCR-4 were compared to women infected with other variants, there was no difference in number of years in Canada since immigration (median of 8 years (range 1-10) versus 9 years (range 3-18)), of sexual activity since immigration (median of 6 years (range 1-19) versus 13 years (range 3-22)) or of sexual activity before immigration (median of 16 years (range 5-18) versus 4 years (range 1 – 16) ( $p>0.10$ ), but the number of participants in both categories was low.

We also assessed whether HIV-seropositive and HIV-seronegative women were susceptible to the same HPV-31 variants (Table 1). The distribution of variants was similar between HIV-seropositive and HIV-seronegative women living in Canada. There was also no significant association between HPV-31 variant and age of participant or level of blood CD4 counts (data not shown).

*Association between HPV-31 LCR variant and persistence of HPV-31 infection.* Sixteen women were considered to have persistent HPV-31 infection. All had obtained three consecutive specimens containing HPV-31 DNA over a period greater than 9 months. PCR-sequencing of the LCR of HPV-31 isolates detected in two samples collected at least 9 months apart demonstrated the presence of the same variant in each pair of samples obtained from the same participant. The average period of HPV-31 positivity was  $785 \pm 618$  days (median 702 days; range 269 to 2784 days). Twenty-seven women were infected by HPV-31 for less than 9 months, and were considered to have transient HPV-31 infection. Thirty six (45.6%) of the 79 HPV-31-infected women could not be classified into transient or persistent infection: 20 women provided samples at one visit only, 13 were infected by HPV-31 at their last visit, 2 were followed for only two visits less than 9 months apart and one woman had HPV-31 infection detected in vaginal tampon specimens only and not in cervicovaginal lavages.

The 16 women with persistent infections were compared to 27 women with transient infections to assess the association between HPV-31 variants and persistence of infection. The two most frequent variants were considered in the analysis: the prototype and variant 31-LCR-3. Considering the prototype, 4 (25%) of 16 women with persistent infection versus 6 (22%) of 27 women with transient infection were infected by the prototype ( $p = 0.99$ ). For the 31-LCR-3 variant, 3 (19%) women with persistent infection and 6 (18.5%) women with transient infection were infected with the latter variant ( $p = 0.99$ ). Seven and 12 different variants were found in women with persistent and transient infections, respectively. Cytology smears were obtained for 64 participants

(33 were normal, 7 were suggestive of ASCUS, 14 of LSIL, one of HSIL, and 9 could not be interpreted). Cytology smear results were identical between HIV-seropositive and seronegative women ( $p > 0.1$  for all comparisons). Women with persistent HPV-31 infections were slightly more likely to have an abnormal cytology result (9 (60%) of 15 women) than women with transient infections (6 (30%) of 20 women), although the difference did not reach statistical significance ( $p = 0.08$ ). There was no significant difference in the proportion of prototypic variants found in women with normal smears (6 (18.2%) of 33 women) and those with abnormal smears (4 (18.2%) of 22) or with LSIL only (2 (14.3%) of 14;  $p > 0.99$ ).

*Polymorphism of the HPV-31 E6 protein.* The analysis of HPV-31 variants was then extended by sequencing the complete 450-bp E6 gene (nucleotide 108 to 557) in the subset of 41 women with persistent or transient infections. Ten molecular variants including the prototype were identified in our cohort and were defined by 13 single-nucleotide variation sites provided in Fig 2. Deletion or insertion were not a source of polymorphism in E6. We compared the polymorphism of E6 and LCR in the 42 isolates analyzed for both regions. The latter 41 isolates did not include variants 31-LCR-9, 31-LCR-12, and 31-LCR-17 that had been identified in the complete bank of isolates. All the variations detected in these variants were shared by other variants except for variation at nucleotide 7565 that was unique to variant 31-LCR-17. Genetic variability was lower in E6 than in the LCR. Thirteen variation sites were found over a 450 bp fragment for E6 compared to 24 (6.2%) over a 390 bp fragment for the LCR ( $p = 0.032$ ). Despite the fact that the E6 fragment sequenced was longer than the LCR segment, the

median number of 5 variation sites per E6 variant compared to the prototype (mean of  $4.51 \pm 0.78$ , range of 4 to 6 mutations per variant) was smaller than the median number of 10 LCR variations per variant (mean of  $9.64 \pm 2.53$ , range of 2 to 13 variations per variant;  $p = 0.0002$ ). The number of mutations in E6 per isolates ( $3.68 \pm 2.16$  mutation sites per variant, median of 5, range of 0 to 6) was smaller to that found in the LCR ( $7.60 \pm 4.39$  mutation sites, median of 9, range of 0 to 13). Considering only the 41 HPV-31 isolates evaluated for both the LCR and E6, the number of LCR variants was not different from the number of E6 variants (15 variants in 41 isolates versus 10 variants in 41 isolates;  $p = 0.34$ ).

Non-synonymous nucleotide variations were found in 31 (75.6%) of 41 isolates and were observed at six positions in E6 (Fig 3). The non-synonymous variations at nucleotides 520 and 285 were shared by a majority of non-prototypic variants. The non-synonymous mutations at sites 285 and 297 resulted in a change from a hydrophilic to hydrophobic amino acid (Fig 3). Four cysteine arrays for predicted zinc fingers (...CXXC...) could be located at amino acid positions 30 to 33, 63 to 66, 103 to 106 and 136 to 139. Amino acid substitutions resulting from non-synonymous variations were not distributed evenly over the E6 protein (Fig 2). Four of the six non-synonymous variations (from amino acids 52 to 65) occurred near the second cysteine array with two amino acid variations located between the two cysteine residues. Another significant variation was located in the carboxy-terminal end of the protein inside the last predicted zinc finger. In contrast, no significant nucleic acid variation occurred at the amino-terminal end of E6 from nucleotides 108 (amino acid 1) to 258 (amino acid 51).



By aligning the E6 proteins from HPV-16 and HPV-31 prototypes, we found a level of similarity of 65% in the amino acid sequence (data not shown). We deduced the potential functional impact of variations of HPV-31 E6 from the mapping of HPV-16 E6 functions on the E6 amino acid sequence [11,31-35]. The genomic region with the cluster of significant variations was located after the putative p53 degradation site (amino acids FAF located at positions 45 to 47 on E6) [36]. One variation occurred at nucleotide position 475 (amino acid 123) located in one of the putative site for p53 binding and degradation [11,36,37].

We further tested whether HIV-seropositive and HIV-seronegative women were infected by the same HPV-31 E6 variants. Twenty-three (74.2%) of 31 HIV-seropositive women and 8 (80.0%) of 10 HIV-seronegative women were infected by isolates with non-prototype E6 variants ( $p = 1.00$ ). There was no difference between HIV-seropositive and HIV-seronegative women for the number of isolates with non-synonymous variations (23 (74.2%) of 31 women versus 8 (80%) of 10,  $p = 1.00$ ). The greatest difference between HIV-seropositive and seronegative women in our cohort was found for 31-E6-1 variant that was detected only in HIV-seropositive women ( $n = 4$ ). The difference in detection rate of the latter variant between HIV-seropositive and seronegative women was not statistically significant however ( $p = 0.56$ ). Due to the small number of participants and number of variants, we lacked the power to conclude that the latter variants were preferentially infecting HIV-seropositive women. There was no difference in the rate of detection of HPV-31 E6 prototypes between normal women and those with LSIL (5 of 19 women versus 3 of 8 women,  $p = 0.38$ ) or between women

with persistent (4 of 18 women) and those with transient infections (6 of 23 women;  $p = 1.00$ ). Significant variations or individual variations were not associated with persistence or cytological anomalies (data not shown).

*Genomic polymorphism of the HPV-31 E7 protein.* HPV-31 polymorphism in E7 is summarised in Fig 2 in the subset of 41 women with persistent or transient infections. Each of the 8 E7 variants contained from 2 to 5 mutations (mean of  $4.29 \pm 1.11$ , median of 5) compared to the prototype. Deletion or insertion was not demonstrated in E7. There was a tendency for greater genetic variability in the LCR compared to E7. Nine (3.0%) mutation sites were found over 296 nucleotides compared to 24 (6.2%) over 390 bp nucleotides for the LCR ( $p = 0.06$ ). The number of mutation sites per variant in the LCR was reduced by a factor of 0.76 to take into account that the sequence analysed in the LCR was longer than that for E7 (390 bases in comparison to 296 bases). The median number of 4 variation sites per E7 isolate compared to the prototype (mean of  $3.17 \pm 1.75$ , range of 0 to 5 mutations per variant) was smaller than the median number of 6.84 LCR variations per isolate (mean of  $5.78 \pm 3.32$ , range of 0 to 9.12 variations;  $p < 0.0001$ ). In the 41 HPV-31 isolates evaluated, the number of LCR variants was similar to the number of E7 variants (15 variants versus 8 variants;  $p = 0.140$ ). The number of mutation sites per non-prototypic variant in E7 (mean of  $4.29 \pm 1.11$ , median of 5, range of 2 to 5 mutations) was then compared to that of non-prototypic variants in the LCR (median of 10 variations, mean of  $9.64 \pm 2.53$ , range of 2 to 13 variations) taking into account that the LCR fragment that was sequenced was slightly longer than that for E7. The number of variation sites per LCR variant was greater than the number

of variation sites per E7 variant ( $p < 0.001$ ). The number of variation sites for the 41 isolates analysed was equivalent between E7 and E6 (9 (3.0%) over 296 nucleotides versus 13 (2.9%) over 449 nucleotides,  $p = 0.91$ ).

Of the 9 variation sites in the E7 ORF, 5 were non-synonymous (Fig 223). E7 functions have been mapped for HPV-16 but not for HPV-31 [31,32,38]. To investigate the possibility that these variations could result in a change in protein function, we aligned the HPV-16 and HPV-31 E7 proteins. We found a 78% (76 over 98 amino acids) level of similarity in the amino acid sequence (data not shown). The CR1 region of E7 was not polymorphic in our isolates. Two significant variations at nucleotides 626 and 645 were located in the CR2 region of E7 while three other significant variations occurred in the CR3 region. No mutation was found in the regions interacting with casein kinase II but one (at nucleotide 626, amino acid 23) was located in the putative region binding with Rb. For all variants, the cysteine doublet in the CR3 region was conserved.

*Cosegregation of HPV-31 Variations.* The prototype-E6-31, 31-E6-1, 31-E6-2, 31-E6-4, 31-E6-5, and 31-E6-6 variants, and 31-E7-2 variant were further classified by LCR analysis (Table 2). As shown in table 2 however, three HPV-31 LCR variants were further classified by E6 or E7 analysis (variants 31-LCR -3, 31-LCR-4, 31-LCR-13).

## **Discussion**

Nearly 8% of women participating in our cohort study were infected at some point in time by HPV-31. HPV-31 was frequently detected in a recent cohort study of

HIV-seronegative women living in Brazil [39]. In another cohort study recruiting HIV-infected women, HPV-31 had the highest incidence rate [40]. HPV infection in general has been found to be more frequent in HIV-seropositive than HIV-seronegative women. In opposite to our report on HPV-52 infection, the detection rate in our cohort of HPV-31 was similar in HIV-seropositive and HIV-seronegative women [29].

Phylogenetic studies of HPV-16 and 18 have established the stability of HPV variants and suggested coevolution of HPV with humans [41-45]. We analyzed HPV-31 sequence variations in a collection of cervical samples from Canadian women at risk of being infected or infected by HIV. Our report is the first extensive work on the genetic polymorphism of HPV-31 *in vivo*. Very few studies have examined the genetic polymorphism of high-risk types other than 16 or 18 and none have described HPV-31 polymorphism [18,46-48].

We first studied the genetic polymorphism of the LCR region, since this region is a hypervariable non-coding segment of the HPV genome [16,43,44]. The amplicon studied included the region of the transcriptional enhancer [43,45,49]. In the LCR, we observed that HPV-31 variants did not evolve from random mutational since variation sites were shared by several isolates and by several variants. If mutations occurred randomly, each nucleotide would have an equivalent likelihood of varying from the prototype sequence in any given isolate. Our data indicate the opposite, since many nucleotides were mutated in several isolates and very few variations occurred only once. Combinations of variations were found in different isolates. Variations in all three genomic areas studied are likely biological entities rather than PCR artifacts. All

variations from the prototype were confirmed. The same pattern of genomic variation has been described for HPV types 6, 16, 18, 33, 35 and 52 [24,29,44,45,50-52].

The genetic variation of 3.3% found between the most divergent LCR non-prototypic variants and the prototype corresponds to the expected degree of diversity within a type [16,42,53]. No deletion or insertion were identified in the 79 HPV-31 isolates in opposite to our findings for HPV type 33 in which 13 (35.1%) of 37 isolates had insertions or deletions (OR of 42.8, 95% CI of 5.3 – 344,  $p < 0.001$ ) [54]. We had also observed a greater occurrence of deletions for HPV-52 (31.5% of 111 HPV-52 isolates) compared to HPV-16 (0% of 144 HPV-16 isolates) in a previous study on the same cohort of women [29,55]. These differences in the polymorphism between types stresses the importance to investigate types other than 16 or 18 to better understand the molecular epidemiology of HPV infection and HPV-related factors predisposing to progressive SIL.

Ethnicity is associated with lineage of HPV-16 variants and also with cervical cancer risk [16,17]. One study found a small degree of diversity among HPV-16 isolates within geographic regions compared a greater diversity of isolates from different geographic regions, suggesting a slow accumulation of mutations over a prolonged period of time [43]. Isolates in the current were detected from women living in Canada although of different ethnic origin. We found that women of African descent were more likely to be infected by non-prototype HPV-31 variants. The prototype strain of an HPV type is usually the first isolate described and sequenced. The HPV-31 prototype was detected in a biopsy sample from a woman living in the United States with cervical

intraepithelial neoplasia {Goldsborough, DiSilvestre, et al. 1989 ID: 786}. Understandably, Caucasian women were more frequently infected by the prototype than non-prototype variants. One variant preferentially infected women from African descent. By analyzing time since immigration and years of sexual activity before and after immigration, we could not demonstrate a difference between women from African descent infected with the latter variant and those infected with other variants. Knowledge on the sexual partner(s) would have been useful to further assess the association between this variant and ethnic origin, but this information was not investigated in our cohort study. Further study on acquisition and transmission of HPV should evaluate the status of partner in terms of HPV infection as well as HPV molecular variants involved.

As found by others, E6 and E7 were less variable than the LCR [16]. A mutation site reported in the E6 ORF of type 16 was also found in HPV-31 at nucleotide 148 and resulted in an amino acid change [56]. All other mutation sites in E6 and E7 were not shared between HPV-16 and HPV-31 [57]. Significant variations were identified in the CR2 and CR3 regions of E7. CR2 in HPV-16 E7 is involved in binding with the pRB and is essential for transformation [11,31,32,38,58]. However, their functional impact was not investigated in this report.

Intergene variation between LCR and E6 or E7 was demonstrated here as has been reported for type 16 [24,59]. Recombination events between variants are very rare [53]. For HPV-52, we found that E6 and E7 analysis did not help in further classifying LCR variants [29]. Wheeler et al. had found that E6 analysis did classify variants that

were not discriminated by LCR analysis [59]. We found here that analysis of E6 or E7 sequences was useful to discriminate between variants that were not different by LCR variant analysis. E5 was also reported to be as rich in genomic variations as the LCR in HPV-16 [43]. Sequencing several genomic areas of HPVs may help refine variant classification. Complete analysis of the HPV genome may also provide more information on the association between variants and disease. In opposite to previous reports on HPV-16 [17], HPV-52 [29] and HPV-33/35 [52], non-prototypic HPV-31 variants were not associated with persistence of HPV infection. The power of this evaluation was however limited by the small number of participants with persistent and transient infections. We can not rule out an association between some HPV-31 molecular variants and persistence. However, when we consider the associations that we found in our cohort for types 33/35 and 52 with similar number of participants, the strength of the association between HPV-31 variants and disease or persistence will likely be lower than for the other types.

This study demonstrates clearly that HPV-31 is polymorphic. The polymorphism of HPV-31 was different from HPV-52 studied in the same cohort of women. HPV-31 polymorphism was associated with the ethnic origin of participants as non-prototype variants were more frequently encountered in women of African descent. Variations in the LCR were more frequent than in coding genes. However, several variations in E6 or E7 introduced a change in amino acid sequence and allowed to discriminate further between variants than only the LCR analysis. The role of HPV-31 polymorphism should

now be investigated in women with and without SIL, some significant mutations described in the HPV-16 E6 being also detected in the E6 of type 31.

### **Acknowledgments**

We would like to thank Mme Diane Gaudreault and Mme Diane Bronsard for processing genital samples. We would like to thank M<sup>me</sup> Julie Lacaille and France Dion for the sequencing of HPV-31. This study was supported by CANFAR. The Canadian Institutes of Health Research supports The Canadian Women's HIV Study. F.C. is a chercheur national supported by the Fonds de la Recherche en Santé du Québec and by le Réseau FRSQ-SIDA.





**E6 mutation at nucleotide position:**

E6 Variants	2	2	2	2	2	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	5	Frequency
reference	t	a	c	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	c	
mutation	c	c	t	g	c	t	g	c	a	g	g	g	t									
31-Prototype-E6	T	A	C	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	C	10
31-E6-1	-	-	G	-	-	T	-	-	-	-	-	-	-	-	-	G	T	-	-	-	-	4
31-E6-2	-	-	T	-	-	T	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	3
31-E6-3	-	-	C	T	-	T	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	2
31-E6-4	C	-	-	G	-	T	-	-	-	-	-	-	G	T	-	-	-	-	-	-	-	6
31-E6-5	-	-	T	-	-	T	-	-	A	G	-	-	-	-	-	-	-	-	-	-	-	5
31-E6-6	-	-	T	-	-	T	-	-	A	-	G	-	-	-	-	-	-	-	-	-	-	8
31-E6-7	-	-	T	-	C	T	-	-	A	-	G	-	-	-	-	-	-	-	-	-	-	1
31-E6-8	-	-	T	-	-	T	-	-	C	A	-	G	-	-	-	-	-	-	-	-	-	1
31-E6-9	-	-	T	-	-	T	G	-	A	-	G	-	-	-	-	-	-	-	-	-	-	1

Codon position	47	52	60	64	65	71	73	77	99	100	107	123	138
Wild-Type AA	F	I	H	T	K	S	V	R	L	L	Q	K	A
Mutated AA	F	L	Y	A	N	S	V	R	L	L	Q	R	V

a)

**E7 mutation at nucleotide position:**

E7 Variants	5	6	6	6	6	6	6	6	6	6	6	6	6	6	7	7	7	7	7	7	7	Frequency
reference	g	t	t	c	c	c	c	c	c	c	c	c	c	c	g	a						
mutation	a	c	g	t	t	t	t	t	t	t	t	t	t	t	a	g						
31-Prototype-E7	G	T	T	C	C	C	C	C	C	C	C	C	C	C	G	A						9
31-E7-1	-	-	-	T	-	-	-	-	T	-	-	-	-	-	A	G						9
31-E7-2	A	-	-	-	-	-	-	-	T	-	-	-	-	-	A	G						18
31-E7-3	A	-	-	-	-	-	-	-	T	-	-	-	-	-	A	G						1
31-E7-4	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	G						1
31-E7-5	A	C	-	-	-	-	-	-	T	-	-	-	-	-	A	G						1
31-E7-6	-	-	-	-	-	-	-	-	T	-	-	-	-	-	T	A	G					1
31-E7-7	A	-	G	-	-	-	-	-	T	-	-	-	-	-	A	G						1

Codon position	7	15	17	23	29	37	41	46	62
Wild-Type AA	T	L	P	H	P	V	P	E	K
Mutated AA	T	L	P	Y	L	V	L	K	E

b)

**Figure 2- 2a)** Sequence variations of 79 HPV-31 isolates in open reading frame E6. **2b)** Sequence variations of 79 HPV-31 isolates in open reading frame E7.

**Figure Legends.**

**Figure 1- Sequence variations in the 3'end of LCR of 79 HPV-31 isolates.** LCR nucleotide positions where variations were detected are written across the top vertically. Positions for which no variation was encountered compared to the prototype are marked with a '-', whereas a letter showing the mutated base indicates a variation site for that variant. Designation of variants was arbitrary. Frequency is for the number of isolates identified for each variant. The loss or gain in putative transcription factor binding site is provided under the last variant.

**Figure 2. 2a) Sequence variations in E6 of 79 HPV-31 isolates. 2b) Sequence variations in the E7 of 111 HPV-31 isolates.** See figure legend to figure 1. The change in amino acid is provided: Aa is for amino acid.

**Table 1.** Association between HPV-31 variants, ethnicity ad HIV infection

HPV-31 LCR variants	No. of women infected (% of total)					
	Caucasian (n= 56)	Black (n=15)	p value	HIV pos. (n=62)	HIV neg. (n=17)	p value
Prototype	14 (25.0)	0 (0)	0.03	12 (19.4)	2 (11.8)	0.72
31.3	17 (30.4)	1 (6.7)	0.05	16 (25.8)	4 (23.6)	1.00
31.4	2 (3.5)	8 (53.3)	0.001	11 (17.7)	1 (5.9)	1.00
31.13	7 (12.5)	2 (13.3)	1.00	8 (12.9)	2 (11.8)	0.98
31.15	4 (7.1)	1 (6.7)	1.00	4 (6.5)	1 (5.9)	0.95
other variants	12 (21.4)	3 (20.0)	0.72	9 (14.5)	5 (29.4)	0.06

**Table 2.** HPV-31 variant classification in function of genomic region analyzed for 41 isolates.

LCR variants	No. of isolates	E6 variants	E7 variants
LCR-P	9	E6 31-P	E7-P
LCR 31-1	1	E6 31-P	E7 31-4
LCR 31-2	1	E6 31-4	E7 31-1
LCR 31-3	3	E6 31-1	E7 31-1
	4	E6 31-4	E7 31-1
LCR 31-4	1	E6 31.3	E7 31-2
	1	E6 31.3	E7 31-5
LCR 31-5	1	E6 31-5	E7 31-2
LCR 31-6	1	E6 31-2	E7 31-2
LCR 31-7	1	E6 31-5	E7 31-2
LCR 31-8	1	E6 31-2	E7 31-2
LCR 31-10	1	E6 31-1	E7 31-6
LCR 31-11	1	E6 31-4	E7 31-1
LCR 31-13	1	E6 31-6	E7 31-7

	4	E6 31-6	E7 31-2
	1	E6 31-7	E7 31-2
	1	E6 31-8	E7 31-2
LCR 31-14	1	E6 31-9	E7 31-2
LCR 31-15	1	E6 31-6	E7 31-2
LCR 31-16	1	E6 31-6	E7 31-2

## REFERENCES

1. Bosch FX, Lorincz A, Munoz N, Meijer CJLM, and Shah KV. The causal relation between human papillomavirus and cervical cancer. *J Clin Pathol* 2002; 55:244-265.
2. Palefsky JM. Human papillomavirus infection and anogenital neoplasia in human immunodeficiency virus-positive men and women. *J Natl Cancer Inst* 1998; Monographs.:15-20.
3. Hankins C, Coutlee F, Lapointe N, Simard P, Tran T, Samson J et al. Prevalence of risk factors associated with human papillomavirus infection in women living with HIV. *Can Med Ass J* 1999; 160:185-191.
4. Jamieson DJ, Duerr A, Burk R, Klein RS, Paramsothy P, Schuman P et al. Characterization of genital human papillomavirus infection in women who have or who are at risk of having HIV infection. *American Journal of Obstetrics & Gynecology* 2002; 186:21-27.
5. Ahdieh L, Klein RS, Burk R, Cu-Uvin S, Schuman P, Duerr A et al. Prevalence, incidence, and type-specific persistence of human papillomavirus in human immunodeficiency virus (HIV)-positive and HIV-negative women. *J Inf Dis* 2001; 184:682-690.
6. Sun XW, Kuhn L, Ellerbrock TV, Chiasson MA, Bush TJ, and Wright TC, Jr. Human papillomavirus infection in women infected with the human immunodeficiency virus. *New Engl J Med* 1997; 337:1343-1349.

7. Minkoff HL, Eisenberger-Matityahu D, Feldman J, Burk R, and Clarke L. Prevalence and incidence of gynecologic disorders among women infected with human immunodeficiency virus. *American Journal of Obstetrics & Gynecology* 1999; 180:824-836.
8. Duerr A, Kieke B, Warren D, Shah K, Burk R, Peipert JF et al. Human papillomavirus-associated cervical cytologic abnormalities among women with or at risk of infection with human immunodeficiency virus. *American Journal of Obstetrics & Gynecology* 2001; 184:584-590.
9. Ellerbrock TV, Chiasson MA, Bush TJ, Sun XW, Sawo D, Brudney K et al. Incidence of cervical squamous intraepithelial lesions in HIV-infected women. *JAMA* 2000; 283:1031-1037.
10. Munoz N, Bosch FX, de Sanjose S, Herrero R, Castellsague X, Shah K et al. Epidemiologic classification of human papillomavirus types associated with cervical cancer. *New England Journal of Medicine* 2003; 348:518-527.
11. Myers G, B, Bernard H-U, Delius H, Baker CC, Icenogle J, Halpern AL, and Wheeler C. Human papillomavirus. A compilation of analysis of nucleic acid and amino acid sequences. Los Alamos National Laboratory. Publication LA-UR 95-3675. 1995. Los Alamos, N.M.  
Ref Type: Book, Whole
12. Chan S-Y, Bernard H-U, Ong C-K, Chan S-P, Hofmann B, and Delius H. Phylogenetic analysis of 48 papillomavirus types and 28 subtypes and



- variants: a showcase for the molecular evolution of DNA viruses. *J Virol* 1992; 66:5714-5725.
13. McDougall JK. Immortalization and transformation of human cells by human papillomavirus. [Review] [104 refs]. *Current Topics in Microbiology & Immunology* 1994; 186:101-119.
  14. Storey A, Pim D, Murray A, Osborn K, Banks L, and Crawford L. Comparison of the in vitro transforming activities of human papillomavirus types. *EMBO Journal* 1988; 7:1815-1820.
  15. La Ruche G, You B, Mensah-Ado I, Bergeron C, Montcho C, Ramon R et al. Human papillomavirus and human immunodeficiency virus infections: relation with cervical dysplasia-neoplasia in African women. *Int J Cancer* 1998; 76:480-486.
  16. Giannoudis AHerrington CS. Human papillomavirus variants and squamous neoplasia of the cervix. *J Pathol* 2001; 193:295-302.
  17. Villa LL, Sichero L, Rahal P, Caballero O, Ferenczy A, Rohan T et al. Molecular variants of human papillomavirus types 16 and 18 preferentially associated with cervical neoplasia. *J Gen Virol* 2000; 81 Pt 12:2959-2968.
  18. Stewart AC, Eriksson AM, Manos MM, Munoz N, Bosch FX, Peto J et al. Intratype variation in 12 human papillomavirus types: a worldwide perspective. *J Virol* 1996; 70:3127-3136.

19. Hankins C, Lapointe N, Walmsley S, and The Canadian Women's HIV study Grou. Participation in clinical trials among women living with HIV in Canada. *Can Med Ass J* 1998; 159:1359-1365.
20. Coutlée F, Hankins C, Lapointe N, Gill J, Romanowski B, Shafran S et al. Comparison between vaginal tampon and cervicovaginal lavage specimens collection for detection of human papillomavirus DNA by the polymerase chain reaction. *J Med Virol* 1997; 51:42-47.
21. Hildesheim A, Schiffman MH, Gravitt PE, Glass AG, Greer CE, Zhang T et al. Persistence of type-specific human papillomavirus infection among cytologically normal women. *J Inf Dis* 1994; 169:235-240.
22. Goldsborough MD, DiSilvestre D, Temple GF, and Lorincz AT. Nucleotide sequence of human papillomavirus type 31: a cervical neoplasia-associated virus. *Virology* 1989; 171:306-311.
23. Mayrand MH, Coutlee F, Hankins C, Lapointe N, Forest P, De et al. Detection of human papillomavirus type 16 DNA in consecutive genital samples does not always represent persistent infection as determined by molecular variant analysis. *J Clin Microbiol* 2000; 38:3388-3393.
24. Yamada T, Manos MM, Peto J, Greer CE, Munoz N, Bosch FX et al. Human papillomavirus type 16 sequence variation in cervical cancers - a worldwide perspective. *J Virol* 1997; 71:2463-2472.
25. Altschul SF, Gish W, Miller W, Myers EW, and Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990; 215:403-410.

27. Thompson JD, Higgins DG, and Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl Acids Res* 1994; 22:4673-4680.
28. Heinemeyer T, Wingender E, Reuter I, Hermjakob H, Kel AE, Kel OV et al. Databases on transcriptional regulation: TRANSFAC, TRRD and COMPEL. *Nucl Acids Res* 1998; 26:362-367.
29. Aho J, Hankins C, Tremblay C, Lang F, Forest P, Pourceaux K, Rouah F, The Canadian Women's HIV study Group, and Coutlée F. Molecular analysis of human papillomavirus type 52 isolates detected in the genital tract of HIV-seropositive and seronegative women. *Journal of Infectious Diseases* . 2003.  
Ref Type: In Press
30. Landschulz WH, Johnson PF, Adashi EY, Graves BJ, and McKnight SL. Isolation of a recombinant copy of the gene encoding C/EBP. [erratum appears in *Genes Dev* 1994 May 1;8(9):1131]. *Genes & Development* 1988; 2:786-800.
31. Munger K, Scheffner M, Huibregste JM, and Howley PM. Interactions of HPV type 16 oncoproteins with tumour suppressor gene products. *Cancer Surv* 1992; 12:197-217.
32. Werness BA, Munger K, and Howley PM. Role of the human papillomavirus oncoproteins in transformation and carcinogenic progression. de Vita, Hellman, and Rosenberg. *Important advances in oncology*. 1[1], 3-18. 1991.

New York, Academic Press.

Ref Type: Book Chapter

33. Crook TVousden KH. Interaction of HPV E6 with p53 and associated proteins. *Biochemical Society Transactions* 1994; 22:52-55.
34. Chen JJ, Reid CE, Band V, and Androphy EJ. Interaction of papillomavirus E6 oncoproteins with a putative calcium-binding protein. *Science* 1995; 269:529-531.
35. Farthing AJVousden KH. Functions of human papillomavirus E6 and E7 oncoproteins. *Trends in Microbiology* 1994; 2:170-174.
36. Foster SA, Demers GW, Etscheid BG, and Galloway DA. The ability of human papillomavirus E6 proteins to target p53 for degradation in vivo correlates with their ability to abrogate actinomycin D-induced growth arrest. *J Virol* 1994; 68:5698-5705.
37. Crook T, Tidy JA, and Vousden KH. Degradation of p53 can be targeted by HPV E6 sequences distinct from those required for p53 binding and trans-activation. *Cell* 1991; 67:547-556.
38. Barbosa MS, Edmonds C, Fisher C, Schiller JT, Lowy DR, and Vousden KH. The region of the HPV E7 oncoprotein homologous to adenovirus E1a and Sv40 large T antigen contains separate domains for Rb binding and casein kinase II phosphorylation. *EMBO Journal* 1990; 9:153-160.
39. Franco EL, Villa LL, Sobrinho JP, Prado JM, Rousseau M-C, Désy M et al. Epidemiology of acquisition and clearance of cervical Human Papillomavirus

- infection in women from a high-risk area for cervical cancer. *J Inf Dis* 1999; 180:1415-1423.
40. Ahdieh L, Klein RS, Burk R, Cu-Uvin S, Schuman P, Duerr A et al. Prevalence, incidence, and type-specific persistence of human papillomavirus in human immunodeficiency virus (HIV)-positive and HIV-negative women. *J Inf Dis* 2001; 184:682-690.
  41. Bernard HU, Chan SY, and Delius H. Evolution of papillomaviruses. *Current Topics in Microbiology & Immunology* 1994; 186:33-54.
  42. Ho L, Chan SY, Burk RD, Das BC, Fujinaga K, Icenogle JP 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* 1993; 67:6413-6423.
  43. Chan S-Y, Ho L, Ong C-K, Chow V, Drescher B, Durst M et al. Molecular variants of human papillomavirus type 16 from four continents suggest ancient pandemic spread of the virus and its coevolution with humankind. *J Virol* 1992; 66:2057-2066.
  44. Ho L, Chan S-Y, Chow V, Chong T, Tay S-K, Villa LL et al. Sequence variants of human papillomavirus type 16 in clinical samples permit verification and extension of epidemiological studies and construction of phylogenetic tree. *J Clin Microbiol* 1991; 29:1765-1772.
  45. Ong CK, Chan SY, Campo MS, Fujinaga K, Mavromara-Nazos P, Labropoulou et al. Evolution of human papillomavirus type 18: an ancient phylogenetic

root in Africa and intratype diversity reflect coevolution with human ethnic groups. *J Virol* 1993; 67:6424-6431.

46. Nindl I. HPV variants and risk of cervical cancer. *Papillomavirus Report* 2002; 13:1-6.
47. Xin CY, Matsumoto K, Yoshikawa H, Yasugi T, Onda T, Nakagawa S et al. Analysis of E6 variants of human papillomavirus type 33, 52 and 58 in Japanese women with cervical intraepithelial neoplasia/cervical cancer in relation to their oncogenic potential. *Cancer Letters* 2001; 170:19-24.
48. Chan KS, Lam C-W, Cheung TK, Li WWH, Lo KWK, Chan MYM et al. Association of human papillomavirus type 58 variant with the risk of cervical cancer. *J Natl Cancer Inst* 2002; 94:1249-1253.
49. Chong T, Chan WK, and Bernard HU. 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. *Nucl Acids Res* 1990; 18:465-470.
50. Heinzl PA, Chan SY, Ho L, O'Connor M, Balaram P, Campo MS et al. Variation of human papillomavirus type 6 (HPV-6) and HPV-11 genomes sampled throughout the world. *J Clin Microbiol* 1995; 33:1746-1754.
51. Ho L, Chan SY, Burk RD, Das BC, Fujinaga K, Icenogle JP 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* 1993; 67:6413-6423.

52. Gagnon S, Hankins C, Tremblay C, Pourreaux K, Forest P, The Canadian Women's HIV study Group et al. Persistence of human papillomavirus types 33 and 35 infections and viral polymorphism in HIV-seropositive and HIV-seronegative women. Persistence of human papillomavirus types 33 and 35 infections and viral polymorphism in HIV-seropositive and HIV-seronegative women. submitted 2003;
53. Yamada T, Wheeler CM, Halpern AL, Stewart AC, Hildesheim A, and Jenison SA. Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, L2, and L1 coding segments. *J Virol* 1995; 69:7743-7753.
54. Bergeron S, Hankins C, Tremblay C, Pourreaux K, The Canadian Women's HIV study Group, and Coutlée F. Polymorphism of Human papillomavirus types 33 and 35 in HIV-seropositive and seronegative women. in preparation 2003;
55. David-Cordonnier MH, Hamdane M, and D'Halluin JC. c-Myb protein binds to the EP element of the HBV enhancer and regulates transcription in synergy with NF-M. *Biochimica et Biophysica Acta* 1999; 1446:82-92.
56. van Duin M, Snijders PJF, Vossen MTM, Klaassen E, Voorhorst F, Verheijen RHM et al. Analysis of human papillomavirus type 16 E6 variants in relation to p53 codon 72 polymorphism genotypes in cervical carcinogenesis. *J Gen Virol* 2000; 81:317-325.

57. Hu XR, Guo ZM, Tianyun P, Ponten F, Wilander E, Andersson S et al. HPV typing and HPV16 E6-sequence variations in synchronous lesions of cervical squamous-cell carcinoma from Swedish patients. *Int J Cancer* 1999; 83:34-37.
58. Phelps WC, Yee CL, Munger K, and Howley PM. The human papillomavirus type E7 gene encodes transactivation and transformation functions similar to those of adenovirus E1A. *Cell* 1988; 53:539-547.
59. Wheeler CM, Yamada T, Hildesheim A, and Jenison SA. Human papillomavirus type 16 sequence variants- identification by E6 and L1 lineage-specific hybridization. *J Clin Microbiol* 1997; 35:11-19.



## **Article 2**

**Persistence of human papillomavirus types 33 and 35 infections and viral polymorphism in human immunodeficiency virus-seropositive and human immunodeficiency virus-seronegative women.**

Running Head: HPV-33 and 35 polymorphism in HIV-infected women

Simon Gagnon<sup>1,2</sup>, Catherine Hankins<sup>3,4</sup>, Cécile Tremblay<sup>2,5</sup>, Pierre Forest<sup>1</sup>, Karina Pourreaux<sup>4</sup>, The Canadian Women's HIV Study Group, and François Coutlée<sup>1,2,3,5</sup>.

1- Laboratoire de Virologie Moléculaire, Centre de Recherche du Centre Hospitalier de l'Université de Montréal, Montréal, Qué., Canada.

2- Département de Microbiologie et Immunologie, Université de Montréal, Montréal, Qué., Canada.

3- Departments of Epidemiology , Biostatistics and Occupational Medicine, and Oncology, McGill University, Montreal, Que., Canada.

4- Direction de la Santé Publique de Montréal-Centre, Institut National de Santé Publique du Québec, Montréal, Qué., Canada.

5- Département de Microbiologie et Infectiologie, Centre Hospitalier de l'Université de Montréal, Montréal, Qué., Canada.

The Canadian Women's HIV Study Group includes the following investigators from across Canada: Principal investigator: Catherine Hankins. Halifax: Janet Conners\*, Rob Grimshaw, David Haase, Lynn Johnston\*, Wally Schlech\*, Arlo Yuzicappi-Fayant\*. Hamilton: Stephen Landis, Fiona Smaill. London: Tom Austin, Ole Hammerberg, Ted Ralph\*. Montréal: François Coutlée\*, Julian Falutz, Alex Ferency, Marina Klein\*, Louise Labrecque, Normand Lapointe, Richard Lalonde, John Macleod, Grégoire Noël\*, Chantal Perron\*, Jean-Pierre Routy, and Emil Toma. Ottawa: Claire Touchie, Garry Victor\*. Québec: Louise Côté, Hélène Senay, Sylvie Trottier. Saskatoon: Kurt Williams\*. Sherbrooke: Alain Piché\*. Sudbury: Roger Sandre. Toronto: Louise Binder\*, Donna Keystone, Anne Phillips, Anita Rachlis\*, Irving Salit, Cheryl Wagner, Sharon Walmsley\*. Vancouver: Paula Braitstein\*, David Burdge\*, Marianne Harris, Deborah Money\*, Julio Montaner\*. \* study co-investigators.

Address correspondence to: François Coutlée, Département de Microbiologie et Infectiologie, Hôpital Notre-Dame du Centre Hospitalier de l'Université de Montréal, 1560 Sherbrooke est, Montréal (Québec), H2L 4M1, Canada. Tel. 514-890-8000, 25162; fax: 514-412-7512

e-mail: [francois.coutlee@ssss.gouv.qc.ca](mailto:francois.coutlee@ssss.gouv.qc.ca).

All participants provided written informed consent to participate. Ethics committees of each participating institution approved the Canadian Women's HIV Study protocol.

This study was supported by CANFAR. The Canadian Institutes of Health Research supports The Canadian Women's HIV Study. F.C. is a chercheur national supported by the Fonds de la Recherche en Santé du Québec and by le Réseau FRSQ-SIDA. CT is a chercheur-boursier supported by the FRSQ

**Abstract (241 words)**

The role of genomic polymorphism of human papillomavirus (HPV) types other than 16 has not been extensively documented. We have shown previously that non-prototypic HPV-52 variants were associated with persistent HPV infection, a surrogate predictor for cervical lesions. The genetic polymorphism of HPV types 33 and 35 is described here in a cohort of sexually active women. Consecutive cervicovaginal lavages obtained at 6-month intervals from 1055 participants were screened for HPV-33 and 35 with a L1 consensus PCR assay. One isolate from each of the 95 women infected by HPV-33 or 35 was analysed by PCR-sequencing of the long control region (LCR), E6 and E7 genes. For HPV-33, 101 nucleotide variation sites in the LCR, 10 variation sites in E6 and 6 variation sites in E7 defined 11 variants, 6 variants and 6 variants, respectively. For HPV-35, 8 variation sites in the LCR, 10 variation sites in E6 and 5 variation sites in E7 defined 8 variants, 8 variants and 7 variants, respectively. The most variable genomic area in these two HPV types was the LCR. Overall, infection with an E7 non-prototype-like isolate was associated with HPV persistence (relative risk (RR) of 2.0, 95% confidence interval (CI) 1.35-2.97) while HIV infection (95% CI 0.46-4.17), CD4 cell count > 200 cells per  $\mu\text{l}$  in HIV-seropositive women (95% CI 0.64-2.41), ethnic origin

(95% CI 0.68-1.53) and age ( $p=0.90$ ) were not associated with persistence. HPV polymorphism of types other than 16 is also associated with persistence of infection.

## Introduction

High-risk types of human papillomaviruses (HPV) are associated with squamous intraepithelial lesions (SIL) and cancer of the uterine cervix [1]. Mucosal HPV types are classified into low-risk types and high-risk types depending on their association with pre-malignant or malignant proliferations. HPV types 31, 33, 35, and 52 are closely related genetically to HPV-16, the most prevalent type in cervical cancer [2]. The latter types segregate into the same phylogenetic branch of group A of HPV types [3]. HPV types 31, 33 and 35 are amongst the eight most frequent genotypes detected in cancer of the uterine cervix and are classified as oncogenic types [2].

There are less than 2% and 5% genetic variations in coding and non-coding regions within an HPV type, respectively [4,5]. The intratype polymorphism defines molecular HPV variants of one type. Intratypic HPV-16 polymorphism is associated with HPV persistence and anogenital SIL or cancer [4]. Polymorphisms in HPV-16 LCR and/or E6 ORF confer a higher risk for developing anogenital SIL or cancerous lesions [4], or HPV persistence [6]. A better understanding of the impact of mutations which confer a greater risk for the development of cancer could help explain why only a subset of women infected by HPV develop cervical cancer.

However, the description and the impact of genetic variations of high-risk types other than 16 or 18 on oncogenic potential have not been investigated extensively [4,7-10]. There have been few published reports to date describing HPV-33 and 35 molecular variants [7,9]. One report described limited nucleotide variations in types 33 and 35 [7].

The latter evaluation was accomplished on a limited number of isolates detected mainly in cervical cancer specimens.

In this work, we present results on the polymorphism of the regulatory region and transforming open reading frames of HPV types 33 and 35 in a population of Canadian women infected by HIV or at risk for HIV infection. The polymorphism of HPV types 33 and 35 exhibited more variations than type 31. Variations in E7 were associated with persistence of HPV infection with types 33 and 35.

### **Materials and methods**

*Study population and study design.* Subjects were selected from participants in the Canadian Women's HIV Study who were found to be infected with HPV types 33 and 35 in at least one genital specimen. The latter study investigates determinants of HPV infection in HIV-seropositive women. The original study design, characteristics of the cohort and epidemiology of HPV infection have been described elsewhere [11,12]. 1055 women recruited from 1993 to 2000 across Canada from outpatient sexually transmitted diseases (STD) clinics, family practices or infectious disease clinics involved in HIV and STD management [11-13]. Women were eligible to participate if they were seropositive for HIV or if they were seronegative for HIV but at risk for STD and accepted to have annual serological tests for HIV antibodies [11,12]. All participants provided written informed consent to participate. Ethics committees of each participating institution approved the Canadian Women's HIV Study protocol. For all participants, a cervicovaginal lavage using 10 ml of phosphate-buffered saline [pH 7.4] was obtained at

least once during pelvic examination [13]. Specimens were refrigerated within one hour and transported to a central laboratory on wet ice.

A standardized questionnaire was administered upon study entry and six months thereafter [11]. For all HIV-seropositive women, blood CD4 cell count, a Pap smear obtained with a cytobrush and Ayre spatula, vaginal tampon specimens and cervicovaginal lavages were obtained and processed as described previously at inclusion and 6-month intervals [13]. For HIV-seronegative women, vaginal tampons were obtained at 6-month intervals while cervicovaginal lavages and Pap smears were collected at one-year intervals. Cell suspensions from genital specimens were processed as previously described [13].

*Sample processing.* Cell suspensions were pelleted after centrifugation at 2,500 rpm for 10 min at 4°C and resuspended in 500 µl of 10 mM Tris [pH 8.3]. Cell suspensions were lysed with Tween 20 at a final concentration of 0.8% (v/v), and digested with 250 µg/ml of proteinase K for 2 hours at 45°C. Cell lysates were boiled for 10 min and stored at -70°C until tested. The delay between sampling and processing never extended beyond 5 days.

*β-globin and HPV DNA detection.* An aliquot of 5 microliters from each sample was amplified for β-globin DNA with PC04 and GH20 primers to control for DNA integrity, for the presence of an adequate number of epithelial cells, and for amplification inhibitors [13]. β-globin-negative samples were extracted with phenol-chloroform and precipitated with ethanol. Five hundred ng of extracted nucleic acids were then amplified for β-globin. β-globin-positive samples (lysate or extracted DNA) were tested

for the detection and typing of HPV DNA using the MY09-MY11-HMB01 consensus L1 PCR as described previously [13,14]. Negative, weak positive (ten HPV18 DNA copies), and strong positive controls (HPV types 6/11, 16, 31, 33, 35, 39, 45) were included in each amplification run to monitor contamination and overall endpoint sensitivity of each PCR run. The amplification mixture contained 6.5 mM MgCl<sub>2</sub>, 50 mM KCl, 2.5 units of *Taq* DNA polymerase (AmpliTaq, Roche Molecular Diagnostics, Mississauga, Ont.), 200 μM each of dATP, dCTP, dGTP, dTTP, and 50 pmoles of each primer. Amplifications were performed in a TC9600 Thermocycler (Perkin-Elmer Cetus, Montréal, Canada) for 40 cycles with the following cycling parameters: 95°C for 1 min, 55°C for 1 min and 72°C for 1 min. Amplified products were spotted onto nylon membranes and reacted under stringent conditions with <sup>32</sup>P-labelled oligonucleotide probes for types 6, 11, 16, 18, 31, 33, 35, 39, 45, 51, 52, 53, 56 and 58 [13,15]. Measures used to avoid false positive reactions due to contamination have been described elsewhere [13].

*Amplification of HPV- 33 and 35 LCR, E6 and E7 for molecular variant analyses.*

Several type-specific primer pairs were designed to further characterize by PCR-sequencing the 3' end of LCR, the complete E6 and E7 genomic segments from HPV-33 and 35 isolates detected in cervicovaginal lavages. All isolates were analyzed for LCR polymorphism while only isolates involved in transient or persistent infections (see below) were sequenced in the E6 and E7 open reading frames. For type 33, a 542 bp fragment at the 3' end of the LCR (LCR-33-sense, location nucleotide (nt) 7337-7354, 5'-TGGGTGTACCTATATGAG-3' and LCR-33-antisense, location nt 7861-7878, 5'-



GACCTAAAACGGTTAGTC-3'), a 562 bp fragment (E6-33-sense, location nt 35-54, 5'-TGTAACCGAAAGCGGTTCAA-3' and E6-33-antisense, location nt 575-596, 5'-TAACGTTGGCTTGTGTCCTCTC-3') containing the complete 483 bp E6 open reading frame and a 412 bp fragment (E7-33-sense, location nt 529-548, 5'-AGGTCCCGACGTAGAGAAAC-3' and E7-33-antisense, location nt 920-940, 5'-TCTACCTCAAACCAACCAGTA-3') containing the complete 324 bp E7 open reading frame, were selected from the published HPV-33 sequence [16]. The E6 and E7 amplicons overlapped over 68 bp.

For type 35, a 696 bp fragment at the 3' end of the LCR (LCR-35-sense, location nucleotide (nt) 7208-7228, 5'-GTGCAATGTGTTGTACGTGGG-3' and LCR-35-antisense, location nt 33-52, 5'-CAACCGTTTTTCGGTACGACC-3'), a 588 bp fragment (E6-35-sense, location nt 33-52, 5'-GGTCGTACCGAAAACGGTTG-3' and E6-35-antisense, location nt 601-620, 5'-GTTGCCTCGGGTTCCAAATC-3') containing the complete 501 bp E6 open reading frame and a 513 bp fragment (E7-35-sense, location nt 488-508, 5'-AACATCGGTGGACGGTGGACA-3' and E7-35-antisense, location nt 979-1000, 5'-CCCCCTGTACAGTCATCTTC-3') containing the complete 300 bp E7 open reading frame, were selected from the published HPV-35 sequence [17]. The E6 and E7 amplicons overlapped over 133 bp.

Amplification reactions for each HPV region were performed with 2 µl of processed sample in a 100 µl reaction volume containing 10 mM Tris-HCl [pH 8.3], 50 mM KCl, 2.5 units of *AmpliTaq* Gold enzyme (Roche Diagnostic System, Mississauga, Ont.), 2.0 mM MgCl<sub>2</sub> concentration, 0.5 µM of each primer, 0.25 mM each dCTP,

dTTP, dGTP and dATP. Amplification reactions were performed with 5 µl of processed sample in a 100 µl reaction volume containing 10 mM Tris-HCl [pH 8.3], 50 mM KCl, 2.5 units of AmpliTaq Gold DNA polymerase, 0.5 µM of each primer, 0.25 mM each dCTP, dTTP, dGTP and dATP. Amplifications were completed in a 9600 Thermal Cycler (Roche Diagnostic System, Mississauga, Ont.). The amplification profiles with the LCR reagents included an initial step at 94°C for 120 sec followed by 10 cycles at 94°C for 15 sec, at 55°C for 30 sec and at 72°C for 45 seconds. Amplification was completed with 35 cycles at 94°C for 15 sec, at 55°C for 30 sec and at 72°C for 45 seconds. Each cycle was increased by five seconds at the extension step. A last step was performed to complete the amplification reaction at 72°C for 7 min. The amplification of E6 DNA sequences was initiated at 94°C for 120 sec followed by 10 cycles at 94°C for 15 sec, at 60°C for 30 sec and at 72°C for 45 seconds. Amplification was done by 35 cycles at 94°C for 15 sec, at 60°C for 30 sec and at 72°C for 45 seconds. Each cycle was increased by five seconds at the extension step. Amplification was completed by an extension step at 72°C for 7 min. The amplification profile for E7 reagents included an initial step at 94°C for 120 sec followed by 10 cycles at 94°C for 15 sec, at 63°C for 30 sec and at 72°C for 45 seconds. Amplification was completed by 35 cycles at 94°C for 15 sec, at 63°C for 30 sec and at 72°C for 45 seconds. Each cycle was increased by five seconds at the extension step. Amplification was completed by an extension step at 72°C for 7 min.

*Molecular variant analysis by PCR-sequencing.* PCR-amplified HPV-33 and 35 DNA fragments were purified with the QIAquick PCR purification kit (Quiagen Inc.,

Mississauga, Ont) when a single band was visualized after gel electrophoresis or with the QIAquick gel extraction kit protocol (Quiagen Inc.) when several bands were visualized. Direct double-stranded PCR-sequencing was done since most amplified products contain the appropriate nucleotide at any given position [18]. Twenty ng of purified amplicons were sequenced with the fluorescent cycle-sequencing method (BigDye terminator ready reaction kit, Perkin-Elmer) with 25 cycles at 96°C for 10 sec, 50°C for 5 sec and 62°C for 4 min. Sequence analysis was performed on an ABI Prism 3100 Genetic Analyzer system at the Laboratoire d'Épidémiologie moléculaire, CHUM. If a non-prototype variant or ambiguities were obtained, PCR-sequencing was repeated once. If results between these 2 sequencing reactions were discrepant, the lysate was sequenced three times to investigate the presence of multiple variant infections or *Taq*-induced errors. Unconfirmed mutations were considered as PCR artifacts. In contrast, variants have mutations in non-random nucleotide positions.

*Data analysis.* To classify our isolates initially, HPV-33 and 35 LCR DNA sequences were compared using the BLAST sequence analysis software from the Genetic Computer Group [19] to sequences of the prototypes (<http://hvp-web.lanl.gov>) and to the complete GenBank database [20]. Isolates with a DNA sequence different from the prototype were classified as non-prototype variant [16,17]. Sequences from non-prototype variants were aligned for classification using Clustal W 1.8 [21]. The multiple alignments were further refined by manual intervention. To assess the effect of mutations in the LCR on putative binding sites for cellular proteins, the TFSEARCH program was used [22]. The TFSEARCH searches highly correlated sequence fragments

against TFMATRIX transcription factor binding site profile database in the 'TRANSFAC' databases by GBF-Braunschweig.

Proportions were compared using the z statistical test. Categorical variables were compared with the Pearson's chi-square test or the Fisher's exact test, depending on the sample size. The variations per variant or per isolate by genomic area were compared using a non-parametric method, the Mann-Whitney rank sum test. The number of mutation was adjusted for the length of amplicon as described in the result section. The comparison of proportion of prototypic variant or isolates with deletions or insertions identified for each genomic area of each type was done with a Fisher's exact test.

Women were considered to have persistent HPV-33 or 35 infection if they provided consecutive genital specimens containing the same variant and spanning at least a nine-month period [23,24]. Women were classified as having transient HPV-33 or 35 infection if they were infected for less than 9 months and had negative specimens following HPV-31 or 35-positive samples. Women who entered the study without detectable HPV-33 or 35 DNA but became positive at a subsequent visit were considered to have incident infection. Because the prevalence of each molecular variant of HPV-33 or 35 was often low, their association with persistence could not be evaluated for each variant individually. The isolates were thus classified into two broad categories: prototype-like and non-prototype strains. Isolates different from the prototype by at most one variation were designated prototype-like variants. Univariate analysis was first performed to identify factors significantly associated with persistent infection using Fisher's exact or Pearson's chi-square tests for proportions, and the Mann-Whitney rank

sum test for continuous variables (age and CD4 cell counts). Double-sided p values < 0.05 were considered significant. The magnitude of the associations between risk factors and persistent HPV-33 and 35 infections was assessed by calculating relative risks (RR) and 95% confidence intervals (CI).

## Results

Overall, 1055 participants in the Canadian Women's HIV study (732 HIV-seropositive and 323 HIV-seronegative women) were screened for cervical infection by HPV. Combining prevalent and incident cases of infection, 37 (3.5%) and 58 (5.5%) women had one or more genital specimens containing HPV types 33 or 35 DNA. Of the 682 genital specimens collected at enrollment and follow-up visits from these HPV-33 or 35-positive women, 5 (0.7%) failed to amplify  $\beta$ -globin. The lack of results for these samples inadequate for PCR analysis had no impact on classifying participants into transient or persistent infections (see below). The median age of women infected by HPV-33 and HPV-35 was 31.0 years (mean  $32.1 \pm 7.8$  years; range 18 – 47 years) and 32.0 years (mean  $32.5 \pm 7.4$  years; range 19 – 53 years), respectively ( $p = 0.93$ ). Thirty-four (91.9%) of 37 HPV-33-infected women were HIV-seropositive compared to 52 (89.7%) of 58 HPV-35-positive women ( $p = 1.00$ ). The median and the mean CD4 cell count closest to the first HPV-33 and HPV-35-positive sample were measured for 34 and 51 HIV-seropositive women, respectively, and reached  $220 \times 10^6$  cells/L (mean  $310 \pm 301 \times 10^6$  cells/L; range 9 –  $1311 \times 10^6$  cells/L) and  $206 \times 10^6$  cells/L (mean  $281 \pm 231 \times 10^6$  cells/L; range 6 –  $1311 \times 10^6$  cells/L), respectively ( $p = 0.96$ ). AIDS had been

diagnosed in 41 (47.7%) of 86 HIV-infected women (9 in clinical category A, 16 in clinical category B and 16 in clinical category C). Cytology testing with Pap smears was conducted concurrently with HPV detection for 23 women infected by HPV-33 (13 had a normal smear, 6 had low grade SIL (LSIL), 1 had high grade SIL (HSIL) and 3 smears were acellular) and for 38 women infected by HPV-35 (18 had normal smears, 11 had LSIL, 3 had HSIL and 6 smears were acellular).

*Genomic polymorphism of HPV-33.* The frequency as well as the site and nature of variations defining each of non-prototype variants compared to the prototype in the LCR are displayed in Figure 1a. The LCR sequence from nucleotide position 7400 to 7800 was studied in this report to match the area analyzed for HPV-31 and 52 in our previous work [8,25]. The prototype was the most frequent of the HPV-33 LCR variants being detected in 18 (48.6%) of 37 isolates (Fig. 1a). Each variant compared to the prototype strain contained from 1 to 15 variations (mean of  $6.8 \pm 4.0$ , median of 7) in the LCR (Table 1). Excluding primer sequences from the LCR amplicon sequence, there were 101 (25.2%) nucleotide variation sites over 401 nucleotides. The maximal number of sequence variations for any variant compared to the prototype was 15 nucleotides (3.7%) over 401 nucleotides. As shown in Fig. 1a, LCR variations were not distributed randomly. Some variations were shared by several isolates and by several variants. Combinations of variations were also shared by different isolates.

HPV-33 variant 33-LCR-8 contained an insertion of 8 nucleotides at position 7528 (Fig. 1a). Loss of the 78-nucleotide duplication unique to type 33 was encountered for 13 isolates classified into 7 variants (Fig 1a) [16]. The effects on the binding sites for

cellular proteins of each variation in the LCR, or combination of variations when they were less than 10 bases apart, are also provided at the bottom of Figure 1a. Only five LCR variations resulted in the loss or gain of putative cellular factor binding sites. The loss of the 78-nucleotide deletion resulted in the loss of putative binding sites for cellular proteins SRY and c-Myb. Binding sites for the HPV E2 protein were not polymorphic.

The site and nature of variations in E6 and E7 open reading frames defining each non-prototype variants compared to the prototype are depicted in Figure 1b. While deletions or insertions were demonstrated in the LCR region, they were not a source of polymorphism in E6 or E7. The median number of variation sites per variant and per isolate are summarised in Table 1. The maximal number of nucleotides involved in sequence variation for any variant compared to the prototype in E6 was 8 (1.7%) over 483 nucleotides (Table 1). Non-synonymous nucleotide variations in E6 were identified in 8 (29.6%) of 27 isolates and were observed at 6 (60%) of 10 variation sites (Fig. 1b). Non-synonymous variations were shared by several variants and isolates, indicating the presence of variable amino acid positions in HPV-33 E6. Due to the small number of HIV-seronegative participants, we lacked power to investigate if individual variants infected preferentially HIV-seropositive women.

Six of the 10 variations defining 6 E6 variants were non-synonymous. Amino acid substitutions resulting from non-synonymous variations were not distributed evenly over the E6 protein (Fig 1b). The non-synonymous variation A138V at nucleotide position 521 was within the last of the four cysteine arrays for predicted zinc fingers (...CXXC...). By aligning the E6 proteins from HPV-16 and HPV-33 prototypes, we

found a level of similarity of 65.2% in the amino acid sequence (data not shown). We deduced the potential functional impact of variations of HPV-33 E6 from the mapping of HPV-16 E6 functions on the E6 amino acid sequence [20,26-30]. None of the mutations occurred in E6 regions involved in binding or degradation of p53 [20,31,32]. The mutation at nucleotide 355 (L83V) resulted in a change in amino acid as described for HPV-16 and 52 in other studies [4,8].

Of the 6 variation in E7, 4 were non-synonymous (Fig 1b). All non-prototypic isolates carried at least one non-synonymous variation. E7 functions have been mapped for HPV-16 but not for HPV-33 or 35 [26,27,33]. To determine if these variations could be located in an important region for protein functions, we aligned the HPV-16 and HPV-33 E7 proteins. We found a 60.2% level of similarity in the amino acid sequence (data not shown). Two significant variations at nucleotides 586 (K5R) and 606 (V12I) were located in the CR1 region while the two other significant variations occurred in the CR3 region near the carboxyl-terminus end of E7 at amino acid positions 93 and 97.

*Genomic polymorphism of HPV-35.* The frequency, site and nature of variations in the LCR defining each of non-prototype variants compared to the prototype are displayed in Figure 2a. As for HPV-33, only the LCR sequences from nucleotide 7400 to nucleotide 7800 were analyzed. The prototype was not detected in the 58 HPV-35-infected participants. Each of the 6 LCR variants compared to the prototype strain contained from 2 to 6 variations (mean of  $3.9 \pm 1.4$ , median of 4) (Fig. 2a). There were 8 nucleotides (2%) involved in variation over 401 nucleotides in the LCR. The maximal number of sequence variations for any variant compared to the prototype was 6



nucleotides (1.5%) over 401 nucleotides. Insertion was not encountered in these isolates. Single-nucleotide deletions were identified at two positions. Five of 8 variations resulted in the loss or gain of putative binding sites for cellular proteins (Figure 2a) but did not alter binding sites for HPV E2 protein.

Variations defining each of non-prototype variants compared to the prototype in E6 and E7 genes are presented in Figure 2b. The median number of variation sites per variant and per isolate are summarised in Table 1. Nucleotide deletion was not a source of polymorphism in E6 or E7. Non-synonymous nucleotide variations in E6 were found in 17 (45.9%) of 37 isolates and were observed at 2 (20%) of 10 variation sites in E6 (Fig 2b). Amino acid substitutions resulting from non-synonymous variations were not distributed evenly over the E6 protein (Fig 2b). The two non-synonymous E6 variations were located between the second and third of the four cysteine arrays for predicted zinc fingers. As for HPV-33, E6 proteins from HPV-16 and HPV-35 prototype were aligned (level of similarity of 69.0% in the amino acid sequence) (data not shown). None of the significant variations occurred in the E6 region involved in binding or degradation of p53.

Of the 5 variation sites in E7, 3 were non-synonymous (Fig 2b). Only one non-prototypic variant (6 isolates) did not carry a significant variation. As for HPV-33, we aligned the HPV-16 and HPV-35 E7 proteins. We found a 74.7% level of similarity in the amino acid sequence (data not shown). The significant variation Y23H at nucleotide 628 was located in a putative pRb binding site. The pRb binding site is identical in HPV-16 E7 and HPV-35 E7(amino acids DLYCYE) [20,26,27,33,34]. This variation occurred

in two E7 variants. The two other significant variation sites were located in the CR3 region (Fig 2b).

*Differences in the polymorphism of LCR, E6 and E7 in types 31, 33 and 35.* Table 1 presents the variation range and the median number of variations per variant, the median number of variations per isolate, as well as the total genomic variation sites in the three areas of the HPV genomes studied in this work. Results obtained for HPV-31 have been described previously in detail by us and are summarized here to allow comparison between these three related types [25]. Firstly, the variability of each HPV genomic region was compared within each type. Genetic variability was greater for the LCR than for coding sequences for HPV-33. In HPV-33 E6 and E7 (Table 1), variability of 10 (2.1%) nucleotide variation sites over 483 nucleotides and 6 (1.9%) nucleotide variation sites over 324 nucleotides was lower than the 101 (25.3%) nucleotide variation sites identified over 401 nucleotides for the LCR ( $p < 0.0001$  for both comparisons). The differences in the proportion of variation sites between the LCR, E6 and E7 of HPV-35 (Table 1) were not statistically different ( $p = 0.811$  and  $0.993$  for each comparison). The number of variation sites in the LCR of type 31 was greater than in E6 or E7 as reported previously [25]. We had already shown for HPV-31 that the total number of variations per variant (Table 1) was greater for the LCR than for E6 ( $p = 0.0003$ ) or E7 ( $p = 0.0004$ ) while there was no difference between E6 and E7 ( $p = 0.18$ ) [25]. When the number of variations per variant were compared between the LCR, E6 and E7 within types 33 or 35, the differences were not statistically significant ( $p > 0.20$  for all comparisons). In these calculations, the number of variation sites per variant was

increased by a factor of 1.21 for the HPV-33 LCR, 1.49 for HPV-33 E7, 1.25 for the HPV-35 LCR and by 1.67 for the HPV-35 E7 to adjust for the different length of amplicons sequenced (Table 1). The ratio of number of variants over number of isolates tested for the LCR, E6 and E7 were similar for each type (Figures 1a, 1b, 2a, 2b). The proportions of non-prototypic LCR, E6 and E7 variants were similar for types 31 and 33 ( $p > 0.20$  for all comparisons). However, the proportion of non-prototypic HPV-35 E7 variants (14 (37.8%) of 37) was significantly smaller than the proportion of non-prototypic LCR (58 (100%) of 58,  $p = 0.0001$ ) or E6 (37 (100%) of 37,  $p = 0.0001$ ) variants (Fig. 2a and 2b).

*Differences in the HPV genomic polymorphism between types 31, 33 and 35.* For each genomic region, the median number of variations per variant and per isolate were compared between types (Table 1). The length of DNA sequence was not taken into account since they were similar between types. In the LCR, comparisons of number of variations per variant was significant between types 31 and 33 ( $p = 0.002$ ), and between 31 and 35 ( $p = 0.0008$ ), while there was no difference between types 33 and 35 ( $p = 0.14$ ). When the number of variations per isolate were compared, the differences were steeper for the first two comparisons ( $p = 0.0001$ ) and a small significant difference was found between types 33 and 35 ( $p = 0.04$ ). For the E6 open reading frame, the differences between median number of variations per variant or per isolate between types 31 and 33, and between types 33 and 35 were not significant ( $p = 0.11$  for all comparisons). However, the number of variations per variant was greater for type 31 than 35 ( $p = 0.04$ ). This level of significance was lost when the number of variations per

isolate was considered, possibly because of a widespread distribution of isolates between several variants while 32 of 37 isolates were classified into 2 variants only for type 35 (Fig. 2b). When the number of variations per E7 variants or isolates were compared, HPV-35 had significantly less E7 variations than type 31 ( $p < 0.01$ ) and type 33 ( $p < 0.01$ ) although the number of isolates tested for all three genomic areas were similar.

While 13 (35%) of 37 HPV-33 isolates and 58 (100%) of 58 HPV-35 isolates had deletions or insertions in the LCR, none (0%) of the 79 HPV-31 isolates had deletions or insertions ( $p < 0.0001$ ). For each genomic region (Fig 1 and 2), the proportion of non-prototypic isolates was greater for type 35 than for types 31 and 33, and was greater for type 31 than 33 ( $p < 0.005$  for each comparison). As an example for the LCR, non-prototypes were identified for 58 (100%) of 58 isolates while non-prototypes were identified for 65 (82.3%) of 79 HPV-31 isolates ( $p = 0.0001$ ) and for 19 (51.4%) of 37 HPV-33 isolates ( $p = 0.0001$ ). The only exception was the comparison between types 33 and 35 for E7 non-prototypic isolates (9 (33%) of 27 versus 14 (38%) of 37,  $p = 0.80$ )

*HPV-33 and 35 polymorphism and persistence of HPV infection.* Because of the small number of participants infected with types 33 and 35 in our cohort, results obtained with these types were combined to evaluate the role of HPV polymorphism in persistence. Thirty-four women had persistent HPV-33 or 35 infection. All had at least three consecutive specimens containing HPV-33 or -35 DNA over a period greater than 9 months. To refine the definition of persistence, PCR-sequencing of the LCR was applied on isolates detected in two samples collected at least 9 months apart. The same variants

were identified in all pairs of specimens, demonstrating true persistence. Twenty-seven women had transient HPV-31 or 35 infection, being infected by one of these types for less than 9 months (HPV-33/35-positive cervicovaginal lavages at one or two visits) and having at least two visits without HPV-33 or 35 infection after the last positive sample. The nature of HPV-33 and 35 infections could not be determined for 33 (34.7%) of the 95 women infected by one of these types: 17 women participated only once, 8 were infected at their last visit only, 6 had only completed two visits with less than 9 months between each visit, one woman had HPV infection detected in vaginal tampon specimens only and not in cervicovaginal lavages, and for one isolate E6 and E7 could not be sequenced.

Single variants for each type were not associated with persistence (data not shown). However, the power of this analysis was limited by the small number of isolates per variant that did not allow to assess the role of variants individually. Considering all the HPV-33 LCR variants that lost the 78-nucleotide duplication: 8 (44.4%) of 18 women with persistent infection versus 0 (0%) of 8 women with transient infection had lost the duplication (RR of 1.8 (95% CI 1.19 - 2.72,  $p = 0.03$ ).

For types 16, 18 and 52 in other studies as discussed below, non-prototypic LCR variants have been associated with persistence. However, HPV-35 LCR and E6 prototypic variants were not detected in our population. We thus investigated the role of HPV 33 and 35 E7 polymorphism considering two categories of variants: prototype-like and non-prototypes. We also considered other risk factors for persistence, by comparing 34 women with persistent infection to 28 women with transient infection. As shown in

table 2, infection by a non-prototypic E7 variant was significantly associated with persistence of HPV-33 and 35 infections with a relative risk of 2 (95% CI 1.35-2.97). The other factors we not significantly associated with persistence. Since our cohort was initiated in 1993, only a few HIV-seropositive women were under highly active antiretroviral therapy (HAART) while infected by HPV, precluding the analysis of the impact of HAART on HPV persistence. When E7 variants were classified according to the presence of non-synonymous variation(s) compared to the prototype, persistence of infection was associated with infection by variants containing non-synonymous variations compared to the prototype. For type 35, 6 (37.5%) of 16 women with persistent infection were infected by variants with non-synonymous variations compared to 1 (4.8%) of 21 women with transient infection ( $p = 0.03$ ). For type 33, 8 (44.4%) of 18 women with persistent infection were infected by variants with non-synonymous variations compared to none (0%) of 7 women with transient infection ( $p = 0.06$ ). When both types were combined, variants carrying non-synonymous variations were significantly associated with persistence (Table 2). The presence of non-synonymous E6 variations was not associated with persistence (15 (44.1%) of 34 women with persistent infection had an isolate with at least one significant E6 mutation as opposed to 9 (32.1%) of 28 women with transient infection (RR of 1.25, 95% CI of 0.80 - 1.95,  $p = 0.43$ )).

A second definition for transient infection included only incident transient infections, thus excluding cases infected at baseline with HPV-33 and 35 who could represent either true transient infection or the end of a persistent infection. The same

association between non-prototypic variants and persistence was uncovered: 17 (50%) of 34 women with persistent infection and 2 (15%) of 13 women with transient infections were infected by non-prototypic isolates (RR of 1.47, 95% CI 1.05 – 2.06,  $p = 0.046$ ). Because the cohort has just been established, we could not evaluate the association between HPV polymorphism and long term persistence measured in years of HPV infection.

## **Discussion**

In this work, HPV-33 and 35 sequence variations were assessed in a collection of cervical samples obtained from Canadian women infected or at risk of being infected by HIV. Our report is the first extensive work on the genetic polymorphism of infection by these HPV types. Very few studies have examined the genetic polymorphism of high-risk types other than 16 or 18 [7,9,10,35]. Genetic variations within 12 genital HPV types were assessed by analysis of the HPV L1 gene in isolates from invasive cervical cancer [7]. In a limited number of isolates, these types showed limited total nucleotide variation in L1. One single-nucleotide and four nucleotide variations were demonstrated in 8 isolates of HPV-35 and 10 isolates of HVP-33. Another evaluation of the polymorphism of HPV-33 in E6 was more extensive [9]. The latter report analyzed 20 isolates from women with SIL or invasive cancer and described 14 variation sites in E6 while we found 10 sites. None of our participants had invasive cancer. Considering that non-prototype variants were associated with persistence in our study, cervical cancers may contain preferentially non-prototype variants.

We first studied the genetic polymorphism of the LCR region, a hypervariable non-coding segment of the HPV genome, that allows to refine our analysis of molecular variants [4,36]. Non-coding sequences are more divergent than open reading frame sequences, since mutations in coding genes could be counterselected by leading to loss of function. Also functional elements in the enhancer, the transcription-factor binding sites, are often widely spaced by nucleotide stretches of little functional significance in which mutations have no impact. As for studies on HPV-16, 18 and 52 LCR polymorphism, we focused on the region of the transcriptional enhancer [8,36,37]. HPV-16 variants have been shown to have different functional enhancer activity [38].

The genetic variation of 3.5 % found in the LCR for both types corresponds to the expected degree of diversity within a type [4]. HPV-33 and 35 variations were not randomly distributed and variation sites were shared by several isolates and variants, a pattern already reported for types 6, 16, 18, 31 and 52 [8,18,25,37-39]. The presence of a perfect 78-bp tandem repeat after the putative origin of replication in HPV-33 LCR is a unique feature for this HPV type [16]. Our work confirmed the presence of this duplication in the majority of isolates but disclosed the loss of this tandem repeat in several non-prototypic variants. By now, HPV-33 is the only type for which such a large size duplication has been found. It was at first suggested that the 78-bp repeats could act as enhancers of viral transcription being located upstream from the promoter [16]. This conclusion derived from the presence of large tandem repeats in the other polyomaviruses SV40 and BK that were shown to enhance transcription from promoters [40-42]. The structure of the 78-bp repeats suggested that it could attract cellular



regulatory proteins in the enhancer region of HPV-33 through a Z-DNA-forming region [16]. However, others have shown an opposite effect of the presence of tandem repeats in SV40 [43]. Tandem repeats of a sequence containing the MBS-II site repressed transcription possibly by competition between *c-myb* with transactivators. Experiments on the hepatitis B virus have also demonstrated that *c-myb* could activate or repress the hepatitis B enhancer [44]. The loss of the duplication in HPV-33 resulted in loss of a putative binding site for *c-myb*. The impact of the loss of the HPV-33 duplication on HPV-33 transcription level should be further investigated in vitro as well as the association between the loss of tandem repeats and grade of cervical lesions in ongoing case-control and cohort studies. The association between loss of duplication and persistence of HPV-33 infection suggests a suppressive role of the duplication in vivo.

In several studies, the HPV-16 E6 variant 350G (L83V) was associated with persistence and CIN, although this association remains controversial [4]. Interestingly, the same mutation site was found for HPV-33 but not for type 35. It has been reported also in type 52 [8]. Non-synonymous mutations were found within cysteine arrays for zinc fingers in types 33 and 35, a phenomenon that was not reported for HPV-16 or HPV-52. One of the non-synonymous variations in E7 was located within the putative pRb binding site. The role of this mutation in persistent infection could not be investigated individually because of the small number of participants. As we found for HPV-33 E6, another report on HPV-33 polymorphism found that more than half of the variation sites resulted in a change in amino acid [9]. The nucleotide variations described in the current study at nucleotide position 214, 315, 325, 355, 387, 521 and

534 had been described in the latter publication [9]. As we had found in our study, the HPV-33 E6 prototype was the most frequently detected variant.

As reported for HPV types 16, 31 and 52, E6 and E7 were less variable than the LCR for HPV-33 but not for HPV-35 [8,10,25]. In opposite to reports on HPV-58 and 52, the rate of mutations in our study was similar in E6 and E7 for both types [35]. While deletions or insertions were demonstrated in the LCR region, they were not a source of polymorphism in the coding regions. Insertions and deletions in the LCR were found frequently in HPV-33 as we had reported for HPV-52 [8]. They were not as important in types 31 and 35. Insertions in the LCR of HPV-16 have been reported to date in only variant Bb1 despite the fact that over 500 HPV-16 isolates from various international studies have been examined [5,18,45-47]. The phylogenetic analysis of human papillomaviruses revealed that types 16, 31 and 35 derived from the same sub-branch while types 33 and 52 were close in the distance matrix tree but more distant from the first three types [3]. The LCR of type 33 was more variable than that of types 31 and 35. In the LCR, there was no difference in the number of variation sites between types 33 and 35 when variants were compared but there was a difference when isolates were compared. The absence for HPV-35 of prototypic isolates and the important number of HPV-33 prototypic isolates may explain this difference. HPV-35 E7 was less polymorphic than the E7 open reading frames of types 31 and 33. The proportion of non-prototypic isolates was different between types 31, 33 and 35. These differences underscore that the polymorphism of HPV genomes is not the same across related types.

As we have previously shown for HPV-16 and HPV-52, all women with persistent infection had the same variant in two specimens collected at least 9 months apart, indicating true persistence [48,49]. Only women with the same variant of HPV-33/35 detected in consecutive samples for at least 9 months were classified as having persistent HPV infection in our study. A cut-off for persistence of infection of 9 months was chosen for this study, corresponding to three positive samples obtained at 6-month intervals. In a recent study of 1425 women living in Brazil, the mean duration of infection with high-risk types was 8.9 months [50]. Our data provide epidemiologic evidence that HPV-33 and 35 variants differ in their capacity to induce persistent infection. Women with persistent HPV-33 and 35 infection were more frequently infected with non-prototypic E7 variants. Ethnicity has been associated with HPV variance and also with cervical cancer or SIL [4]. However in our study, race was not associated with HPV persistence.

Non-prototypic HPV-16 variants have been associated with persistence of infection [6,51] and anogenital cancer [52-55]. We found the same association between non-prototypic E7 variants and persistence of HPV-33/35 infection. Due to the small number of isolates and participants, we could not identify in E7 the individual variation or combination of variations that are responsible for this association. The association with persistence of non-prototypic variants contrasts with our previous report on HPV-31 [25].

This study illustrates that the polymorphism of oncogenic HPV types is not similar between types. Variations in the LCR were more frequent than in coding genes.

All these differences in the polymorphism between these related types stresses the importance to investigate types other than 16 or 18 to better understand the molecular epidemiology of HPV infection. We identified several variations in E6 or E7 that introduced a change in amino acid sequence. As found for other types, HPV-33 and 35 polymorphism was associated with persistence. We now need to investigate the polymorphism of these types in cervical cancer specimens.

### **Acknowledgments**

We would like to thank Mme Diane Gaudreault and Mme Diane Bronsard for processing genital samples. We would like to thank M<sup>me</sup> Julie Lacaille and France Dion for the sequencing of HPV-. This study was supported by CANFAR. The Canadian Institutes of Health Research supports The Canadian Women's HIV Study. F.C. is a chercheur-national supported by the Fonds de la Recherche en Santé du Québec and by le Réseau FRSQ-SIDA.



**E6 mutation at nucleotide position:** **E7 mutation at nucleotide position:**

E6 Variants	1	2	3	3	3	3	3	3	3	3	5	5	3	3	3	5	6	8	8	8	8	8	6	6	2	3	5	6	Frequency	
reference mutation	c	a	c	c	a	a	g	a	t	c	c	t	c	c	t	c	t	a	g	t	g	c	a	g	a	c	t	g	t	Frequency
33-Prototype-E6	C	A	C	C	A	G	A	T	C	C								A	G	T	G	C	A	A	G	T	G	C	A	18
33-E6-1	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	G	-	-	-	-	-	G	-	-	-	-	1	
33-E6-2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
33-E6-3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2
33-E6-4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
33-E6-5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4

Codon position	18	32	36	69	73	83	93	94	138	142
Wild-type AA	A	E	P	F	I	V	K	P	A	S
Mutated AA	V	E	T	F	L	L	N	P	V	S

**a)**

E7 Variants	5	6	7	8	8	8	8	6	2	3	0	2	3	0	2	Frequency
reference mutation	a	g	t	g	c	a	g	t	a	c	t	g	t	a	c	Frequency
33-Prototype-E7	A	G	T	G	C	A	A	G	T	G	C	A	A	G	T	18
33-E7-1	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
33-E7-2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2
33-E7-3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
33-E7-4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4
33-E7-5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1

Codon position	5	12	50	87	93	97
Wild-type AA	K	V	A	V	T	Q
Mutated AA	R	I	A	V	S	L

**b)**

**Figure 2- 2a)** Sequence variations in the E6 open reading frame of HPV-33. **2b)** Sequence variations in the E7 open reading frame of HPV-33.

## Mutation at nucleotide position:

LCR Variants	7	d	7	7	7	7	7	7	7	7	7	Frequency
35-Prototype-LCR	T	T	A	T	T	T	C	T	C	T	A	0
35-LCR-1	C	-	-	-	-	-	X	-	X	-	-	6
35-LCR-2	C	-	G	-	-	-	X	-	X	-	-	1
35-LCR-3	C	-	-	C	-	-	X	-	X	-	-	20
35-LCR-4	C	-	-	C	C	-	X	-	X	-	-	1
35-LCR-5	C	-	-	C	C	-	X	-	X	G	G	6
35-LCR-6	C	X	-	C	C	-	X	-	X	-	G	13
35-LCR-7	C	X	-	C	C	-	X	-	X	G	G	11

reference  
mutation

## Transcription factor binding site

Gain	-	-	HSF-1	-	-	cap	-
			HSF-2				
Loss	-	-	C/EBPa	SRY	.Oct-1	-	SRY

**Figure 3-**Sequence variation in the 3' end of the LCR of HPV-35

**E6 mutation at nucleotide position:**

E6 Variants	1	1	1	1	1	1	1	1	1	2	3	3	3	3	3	3	7	1	6	3	7	5	3	6	1	0	Frequency
reference mutation	T	C	T	C	T	G	T	A	A	A	A	T	G	C	C	C											
35-Prototype-E6	T	C	T	G	T	A	A	A	A	T	G																0
35-E6-1	-	A	-	-	-	-	-	T	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	18
35-E6-2	-	A	-	-	-	-	-	T	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
35-E6-3	-	A	-	-	-	-	-	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
35-E6-4	-	A	-	-	-	-	-	T	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
35-E6-5	-	A	-	-	-	-	-	T	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
35-E6-6	G	A	C	-	-	-	-	T	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	14	
35-E6-7	C	A	C	-	-	-	-	T	-	-	-	-	-	C	C	-	-	-	-	-	-	-	-	-	-	1	
Codon position	6	8	9	18	26	62	68	73	78	87																	
Wild-type AA	A	R	P	E	I	V	K	I	W	T																	
Mutated AA	A	R	P	E	I	V	K	V	R	T																	

a)

**Figure 4- 4a) Sequence variations in the E6 open reading frame of HPV-35.4b) Sequence variations in the E7 open reading frame of HPV-35.**

**E7 mutation at nucleotide position:**

E7 Variants	6	6	6	6	6	6	7	7	7	7	7	7	7	7	7	7	8	8	8	8	8	8	8	8	8	8	8	Frequency
reference mutation	C	T	T	C	T	T	C	T	C	T	C	T	C	T	C	T	C	G	G	A								
35-Prototype-E7	C	T	T	C	T	T	C	T	C	T	C	T	C	T	C	T	C	G	G	A								23
35-E6-1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	6
35-E6-2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
35-E6-3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
35-E6-4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
35-E6-5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4
35-E6-6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
Codon position	17	23	38	51	63																							
Wild-type AA	P	Y	T	S	E																							
Mutated AA	P	H	T	C	K																							

b)



## Figure Legends.

**Figure 1. Sequence variations in the 3'end of LCR of HPV-33.** LCR nucleotide positions where variations were detected are written across the top vertically. Positions for which no variation was encountered compared to the prototype are marked with a '- ', whereas a letter showing the mutated base indicates a variation site for that variant. Designation of variants was arbitrary. Frequency is for the number of isolates identified for each variant. In 1a, the loss or gain in putative transcription factor binding site is provided under the last variant. D is for deletion, I is for insertion. In 1b, the change in amino acid is provided: Aa is for amino acid.

**Figure 2. 2a) Sequence variations in the E6 open reading frame of HPV-33. 2b) Sequence variations in the E7 open reading frames of HPV-33.** See legend for figure 1.

**Figure 3. Sequence variations in the 3'end of LCR of HPV-35.** See legend for figure 1.

**Figure 4. 4a) Sequence variations in the E6 open reading frame of HPV-35. 4b) Sequence variations in the E7 open reading frames of HPV-35.** See legend for figure 1.

Table 1. Genomic variability in the LCR, E6 and E7 of HPV types 31, 33 and 35

Genomic area	No. isolates	<u>Median no. of variations</u> per variant per isolates		% variation	<u>no. variations</u> length (nt.)
<b><u>LCR</u></b>					
HPV-31	79	10 (2-13)	9	6.4 %	25/391
HPV-33	37	7 (1-15)	1	20.0 %	101/506
HPV-35	57	4 (2-6)	4	1.2 %	8/655
<b><u>E6</u></b>					
HPV-31	41	5 (4-6)	5	2.9 %	13/450
HPV-33	27	7 (1-8)	0	2.1 %	10/483
HPV-35	37	3 (2-6)	3	2.0 %	10/501
<b><u>E7</u></b>					
HPV-31	41	5 (4-5)	4	3.0 %	9/296
HPV-33	27	4 (1-5)	0	1.9 %	6/324
HPV-35	37	2 (1-3)	0	1.7 %	5/300

Table 2. Univariate analysis of risk factors for persistence of HPV infection in 34 women with persistent HPV-33 and 35 infections and 28 women with transient HPV-33 and 35 infections.

	<b>transient (n = 28)</b>	<b>persistent (n = 34)</b>	<b>p<sup>1</sup></b>
<b>Age (years, median)</b>	31.0	29.5	0.90
<b>HIV status</b>			
seronegative	3 (11)	2 (6)	0.65
seropositive	25 (89)	32 (94)	
<b>CD4 counts (median)<sup>2</sup></b>	310	210	0.26
<b>CD4 counts</b>			
>200	8 (50)	11 (61)	0.78
<200	8 (50)	7 (39)	
<b>Ethnicity</b>			
Caucasian	8 (53)	17 (55)	1.00
African	7 (47)	14 (45)	
<b>E7 variants</b>			
non-prototype	2 (7)	14 (41)	0.005
Prototype-like	26 (93)	20 (59)	
<b>E6 variants</b>			
non-synonymous variations	9 (32)	15 (45)	0.43
synonymous or prototype	19 (64)	19 (55)	

Data are number (percentage) unless otherwise stated. **1:** test of association using Fishers exact test except for the comparison of age and CD4 counts between groups which was evaluated by a Mann-Whitney rank sum test. **2:** Means were determined on 34 and 10 women with persistent and transient HPV-52 infections.

## REFERENCES

1. Bosch FX, Lorincz A, Munoz N, Meijer CJLM, and Shah KV. The causal relation between human papillomavirus and cervical cancer. *J Clin Pathol* 2002; 55:244-265.
2. Munoz N, Bosch FX, de Sanjose S, Herrero R, Castellsague X, Shah K et al. Epidemiologic classification of human papillomavirus types associated with cervical cancer. *New England Journal of Medicine* 2003; 348:518-527.
3. Chan S-Y, Bernard H-U, Ong C-K, Chan S-P, Hofmann B, and Delius H. Phylogenetic analysis of 48 papillomavirus types and 28 subtypes and variants: a showcase for the molecular evolution of DNA viruses. *J Virol* 1992; 66:5714-5725.
4. Giannoudis AHerrington CS. Human papillomavirus variants and squamous neoplasia of the cervix. *J Pathol* 2001; 193:295-302.
5. Ho L, Chan S-Y, Chow V, Chong T, Tay S-K, Villa LL et al. Sequence variants of human papillomavirus type 16 in clinical samples permit verification and extension of epidemiological studies and construction of phylogenetic tree. *J Clin Microbiol* 1991; 29:1765-1772.
6. Villa LL, Sichero L, Rahal P, Caballero O, Ferenczy A, Rohan T et al. Molecular variants of human papillomavirus types 16 and 18 preferentially associated with cervical neoplasia. *J Gen Virol* 2000; 81 Pt 12:2959-2968.

7. Stewart AC, Eriksson AM, Manos MM, Munoz N, Bosch FX, Peto J et al. Intratype variation in 12 human papillomavirus types: a worldwide perspective. *J Virol* 1996; 70:3127-3136.
8. Aho J, Hankins C, Tremblay C, Lang F, Forest P, Pourreaux K, Rouah F, The Canadian Women's HIV study Group, and Coutlée F. Molecular analysis of human papillomavirus type 52 isolates detected in the genital tract of HIV-seropositive and seronegative women. *Journal of Infectious Diseases* . 2003.  
Ref Type: In Press
9. Xin CY, Matsumoto K, Yoshikawa H, Yasugi T, Onda T, Nakagawa S et al. Analysis of E6 variants of human papillomavirus type 33, 52 and 58 in Japanese women with cervical intraepithelial neoplasia/cervical cancer in relation to their oncogenic potential. *Cancer Letters* 2001; 170:19-24.
10. Chan KS, Lam C-W, Cheung TK, Li WWH, Lo KWK, Chan MYM et al. Association of human papillomavirus type 58 variant with the risk of cervical cancer. *J Natl Cancer Inst* 2002; 94:1249-1253.
11. Hankins C, Coutlee F, Lapointe N, Simard P, Tran T, Samson J et al. Prevalence of risk factors associated with human papillomavirus infection in women living with HIV. *Can Med Ass J* 1999; 160:185-191.
12. Hankins C, Lapointe N, Walmsley S, and The Canadian Women's HIV study Group. Participation in clinical trials among women living with HIV in Canada. *Can Med Ass J* 1998; 159:1359-1365.

13. Coutlée F, Hankins C, Lapointe N, Gill J, Romanowski B, Shafran S et al. Comparison between vaginal tampon and cervicovaginal lavage specimens collection for detection of human papillomavirus DNA by the polymerase chain reaction. *J Med Virol* 1997; 51:42-47.
14. Coutlée F, Gravitt P, Richardson H, Hankins C, Franco E, Lapointe N et al. Nonisotopic detection and typing of human papillomavirus DNA in genital samples by the line blot assay. *J Clin Microbiol* 1999; 37:1852-1857.
15. Hildesheim A, Schiffman MH, Gravitt PE, Glass AG, Greer CE, Zhang T et al. Persistence of type-specific human papillomavirus infection among cytologically normal women. *J Inf Dis* 1994; 169:235-240.
16. Cole ST, Streek RE. Genome organization and nucleotide sequence of human papillomavirus type 33, which is associated with cervical cancer. *J Virol* 1986; 58:991-995.
17. Marich JE, Ponsler AV, Rice SM, McGraw KA, and Dubemsky TW. The phylogenetic relationship and complete nucleotide sequence of human papillomavirus type 35. *Virol* 1992; 186:770-776.
18. Yamada T, Manos MM, Peto J, Greer CE, Munoz N, Bosch FX et al. Human papillomavirus type 16 sequence variation in cervical cancers - a worldwide perspective. *J Virol* 1997; 71:2463-2472.
19. Altschul SF, Gish W, Miller W, Myers EW, and Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990; 215:403-410.

20. Myers G, B, Bernard H-U, Delius H, Baker CC, Icenogle J, Halpern AL, and Wheeler C. Human papillomavirus. A compilation of analysis of nucleic acid and amino acid sequences. Los Alamos National Laboratory. Publication LA-UR 95-3675. 1995. Los Alamos, N.M.  
Ref Type: Book, Whole
21. Thompson JD, Higgins DG, and Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl Acids Res 1994; 22:4673-4680.
22. Heinemeyer T, Wingender E, Reuter I, Hermjakob H, Kel AE, Kel OV et al. Databases on transcriptional regulation: TRANSFAC, TRRD and COMPEL. Nucl Acids Res 1998; 26:362-367.
23. Schlecht NF, Kulaga S, Robitaille J, Ferreira S, Santos M, Miyamura RA et al. Persistent human papillomavirus infection as a predictor of cervical intraepithelial neoplasia. Jama: Journal of the American Medical Association 2001; 286:3106-3114.
24. Ahdieh L, Klein RS, Burk R, Cu-Uvin S, Schuman P, Duerr A et al. Prevalence, incidence, and type-specific persistence of human papillomavirus in human immunodeficiency virus (HIV)-positive and HIV-negative women. J Inf Dis 2001; 184:682-690.
25. Gagnon S, Hankins C, Tremblay C, Pourceaux K, Forest P, The Canadian Women's HIV study Group et al. Polymorphism of Human Papillomavirus



- type 31 isolates infecting the genital tract of HIV-seropositive and HIV-seronegative women. submitted 2003;
26. Munger K, Scheffner M, Huibregste JM, and Howley PM. Interactions of HPV type 16 oncoproteins with tumour suppressor gene products. *Cancer Surv* 1992; 12:197-217.
  27. Werness BA, Munger K, and Howley PM. Role of the human papillomavirus oncoproteins in transformation and carcinogenic progression. de Vita, Hellman, and Rosenberg. *Important advances in oncology*. 1[1], 3-18. 1991. New York, Academic Press.  
Ref Type: Book Chapter
  28. Crook TVousden KH. Interaction of HPV E6 with p53 and associated proteins. *Biochemical Society Transactions* 1994; 22:52-55.
  29. Chen JJ, Reid CE, Band V, and Androphy EJ. Interaction of papillomavirus E6 oncoproteins with a putative calcium-binding protein. *Science* 1995; 269:529-531.
  30. Farthing AJVousden KH. Functions of human papillomavirus E6 and E7 oncoproteins. *Trends in Microbiology* 1994; 2:170-174.
  31. Foster SA, Demers GW, Etscheid BG, and Galloway DA. The ability of human papillomavirus E6 proteins to target p53 for degradation in vivo correlates with their ability to abrogate actinomycin D-induced growth arrest. *J Virol* 1994; 68:5698-5705.

32. Crook T, Tidy JA, and Vousden KH. Degradation of p53 can be targeted by HPV E6 sequences distinct from those required for p53 binding and transactivation. *Cell* 1991; 67:547-556.
33. Barbosa MS, Edmonds C, Fisher C, Schiller JT, Lowy DR, and Vousden KH. The region of the HPV E7 oncoprotein homologous to adenovirus E1a and Sv40 large T antigen contains separate domains for Rb binding and casein kinase II phosphorylation. *EMBO Journal* 1990; 9:153-160.
34. Phelps WC, Yee CL, Munger K, and Howley PM. The human papillomavirus type E7 gene encodes transactivation and transformation functions similar to those of adenovirus E1A. *Cell* 1988; 53:539-547.
35. Nindl I. HPV variants and risk of cervical cancer. *Papillomavirus Report* 2002; 13:1-6.
36. Chan S-Y, Ho L, Ong C-K, Chow V, Drescher B, Durst M et al. Molecular variants of human papillomavirus type 16 from four continents suggest ancient pandemic spread of the virus and its coevolution with humankind. *J Virol* 1992; 66:2057-2066.
37. Ong CK, Chan SY, Campo MS, Fujinaga K, Mavromara-Nazos P, Labropoulou et al. Evolution of human papillomavirus type 18: an ancient phylogenetic root in Africa and intratype diversity reflect coevolution with human ethnic groups. *J Virol* 1993; 67:6424-6431.

38. Heinzl PA, Chan SY, Ho L, O'Connor M, Balaram P, Campo MS et al. Variation of human papillomavirus type 6 (HPV-6) and HPV-11 genomes sampled throughout the world. *J Clin Microbiol* 1995; 33:1746-1754.
39. Ho L, Chan SY, Burk RD, Das BC, Fujinaga K, Icenogle JP 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* 1993; 67:6413-6423.
40. Benoist C, Chambon P. In vivo sequence requirements of the SV40 early promoter region. *Nature* 1981; 290:304-310.
41. Seif I, Khoury G, and Dhar R. The genome of human papovavirus BKV. *Cell* 1979; 18:963-977.
42. Gruss P, Dhar R, and Khoury G. Simian virus 40 tandem repeated sequences as an element of the early promoter. *Proc Natl Acad Sci USA* 1981; 78:943-947.
43. Nakagoshi H, Nagase T, Kanei-Ishii C, Ueno Y, and Ishii S. Binding of the c-myc proto-oncogene product to the simian virus 40 enhancer stimulates transcription. *J Biol Chem* 1990; 265:3479-3483.
44. David-Cordonnier MH, Hamdane M, and D'Halluin JC. c-Myb protein binds to the EP element of the HBV enhancer and regulates transcription in synergy with NF-M. *Biochimica et Biophysica Acta* 1999; 1446:82-92.
45. Schmidt M, Kedzia W, and Gozdicka-Jozefiak A. Intratype HPV16 sequence variation within LCR of isolates from asymptomatic carriers and cervical cancers. *Journal of Clinical Virology* 2001 Dec;23 (1-2 ):65 -77 23:65-77.

46. Ho L, Chan SY, Burk RD, Das BC, Fujinaga K, Icenogle JP 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* 1993; 67:6413-6423.
47. Buonaguro FM, Tornesello ML, Salatiello I, Okong P, Buonaguro L, Beth-Giraldo E et al. The uganda study on HPV variants and genital cancers. *Journal of Clinical Virology* 2000; 19:31-41.
48. Mayrand MH, Coutlee F, Hankins C, Lapointe N, Forest P, De et al. Detection of human papillomavirus type 16 DNA in consecutive genital samples does not always represent persistent infection as determined by molecular variant analysis. *J Clin Microbiol* 2000; 38:3388-3393.
49. Aho J, Hankins C, Tremblay C, Forest P, Pourreaux K, Rouah F et al. Genomic polymorphism of Human Papillomavirus type 52 predisposes to persistent infection, a surrogate marker for cervical cancer. submitted 2003;
50. Franco EL, Villa LL, Sobrinho JP, Prado JM, Rousseau M-C, Désy M et al. Epidemiology of acquisition and clearance of cervical Human Papillomavirus infection in women from a high-risk area for cervical cancer. *J Inf Dis* 1999; 180:1415-1423.
51. Londesborough P, Ho L, Terry G, Cuzick J, Wheeler C, and Singer A. Human papillomavirus genotype as a predictor of persistence and development of high-grade lesions in women with minor cervical abnormalities. *Int J Cancer* 1996; 69:364-368.

52. Hildesheim A, Schiffman M, Bromley C, Wacholder S, Herrero R, Rodriguez A et al. Human papillomavirus type 16 variants and risk of cervical cancer. *J Natl Cancer Inst* 2001; 93:315-318.
53. Da Costa MM, Hogeboom CJ, Holly EA, and Palefsky JM. Increased risk of high-grade anal neoplasia associated with a human papillomavirus type 16 E6 sequence variant. *J Inf Dis* 2002; 185:1229-1237.
54. Xi LF, Koutsky LA, Galloway DA, Kuypers J, Hughes JP, Wheeler CM et al. Genomic variation of human papillomavirus types 16 and risk for high grade cervical intraepithelial neoplasia. *J Natl Cancer Inst* 1997; 89:796-802.
55. Xi LF, Critchlow CW, Wheeler CM, Koutsky LA, Galloway DA, Kuypers J et al. Risk of anal carcinoma in situ in relation to human papillomavirus type 16 variants. *Cancer Res* 1998; 58:3839-3844.

## **DISCUSSION**

À ce jour, il n'y a jamais eu d'étude aussi exhaustive sur le polymorphisme des VPH-31, 33 et 35 que celle-ci. Ainsi, ces études n'ont pas permis de définir un grand nombre de variants, et ce dans un nombre de cibles limité, soit dans L1 et E6. Dans les faits, aucune étude du polymorphisme du VPH-31 n'a jamais été publiée. Quant à lui, le VPH-33 a été examiné au niveau des cadres de lecture L1 et E6 dans le cadre de deux projets [149, 179] alors qu'une seule étude portait sur le VPH-35, au niveau de L1 [179].

Notre projet incluait aussi une analyse de la persistance de ces types ainsi qu'une mise en relation entre l'infection au VIH et l'infection aux VPH-31, 33 et 35. Une telle analyse des variants moléculaires de ces VPHs à haut risque est importante, dans la mesure où il a déjà été démontré chez d'autres types qu'il existait une association entre certains polymorphisme et la persistance virale de même qu'avec le développement de cancers du col utérin. Considérant le fait que certains variants peuvent constituer des marqueurs pour la progression et la promotion du développement du cancer du col utérin, une analyse de variants est d'une grande importance pour la prévention, le diagnostic et la thérapie des lésions associées à l'infection par les VPHs oncogènes. En effet, des variations au sein du LCR peuvent avoir une influence sur la persistance, et ce par l'entremise de la perte ou du gain de sites de liaisons à des facteurs de transcription cellulaires pouvant moduler la transcription et la réplication des VPHs. D'autre part, des variations génétiques dans les cadres de lecture E6 et E7 peuvent résulter en des changements d'acides aminés dans les protéines correspondantes. Ces mutations significatives pourraient être situées dans des domaines protéiques importants pour la transformation et l'immortalisation cellulaire, tels que les domaines de liaison à p53 de E6 et de liaison à pRb de E7, et causer un changement dans les capacités oncogènes de ces protéines.

Au sein de notre cohorte de 1055 femmes, le VPH-31 était le plus prévalent des trois types étudiés. Dans plusieurs autres études, le VPH-31 était également plus fréquemment retrouvé que les types 33 et 35 [1, 41, 171, 106]. Les infections au VPH sont généralement plus fréquemment retrouvées chez les individus infectés par le VIH

que chez les individus VIH-négatifs. Or, la prévalence du VPH-31 chez les femmes VIH-positives n'était pas significativement plus élevée que pour les femmes séro-négatives, contrairement à ce qui est observé pour les VPH-33 et 35 (présentation CAHR, 2003), ainsi que pour le type 52 [2], qui fut étudié précédemment. De plus, il est intéressant de constater dans une étude [1] que le VPH-31 était le seul avec le VPH-16 (sur 27 types étudiés) dont l'incidence n'augmentait pratiquement pas avec l'infection au VIH. Du fait que l'infection au VIH ne semble pas avoir d'influence sur l'incidence de l'infection au VPH-31, il est ici tentant de poser l'hypothèse que ce type, de même que le VPH-16, serait plus apte à s'évader du système immunitaire. Cependant, une autre étude a déterminé que la prévalence du VPH-31 corrélait avec un décompte de lymphocytes T CD4+ inférieur à 200 cellules par  $\mu\text{L}$  de sang, contrairement aux VPH-16 [155]. D'autre part, cette corrélation n'était pas non plus retrouvée pour les types 33 et 35, ce qui va à l'inverse de ce que nous avons obtenu. Cette discordance pour les types 33/35 pourrait s'expliquer par le petit nombre d'isolats positifs pour dont nous disposions pour ces types. Quoi qu'il en soit, il a été suggéré que le VPH-16 ainsi que les types 33 et 35, qui lui sont phylogénétiquement reliés, présenteraient des épitopes particuliers qui leur confèreraient une protection contre la réponse immunitaire lors de l'infection ou la réinfection [155]. L'absence de corrélation entre l'infection au VIH et l'infection par le VPH-31 dans notre cohorte s'accorderait avec cette hypothèse, ce type s'avérant lui-aussi phylogénétiquement très proche du VPH-16. ✓

Le polymorphisme des VPH-31/33/35 a été étudié au niveau de trois régions génomiques, soit le LCR et les cadres de lecture E6 et E7. Le LCR est une région hypervariable non-codante qui montre une variation intratypique pouvant aller jusqu'à 5% [16]. En comparaison, le taux de variation intratypique dans les régions codantes ne peut atteindre que 2% [16], dû au fait que les mutations entraînant des pertes de fonction ne sont pas favorisées. La région LCR était pour nous une cible intéressante, étant donné que l'étude sur le VPH-52 effectuée dans notre laboratoire [2] avait permis de découvrir une relation significative entre la persistance de ce type et son polymorphisme. Pour leur part, les cadres de lecture E6 et E7 formaient également d'excellentes cibles, ces



derniers codant pour les principales oncoprotéines virales. D'autre part, une relation entre la persistance et le polymorphisme de E6 avait déjà été démontré ultérieurement chez le VPH-16 [37], permettant ainsi d'entrevoir la possibilité d'un phénomène similaire chez les VPH31/33/35.

Nous avons donc étudié la région «enhancer» en 3' du LCR, région contenant plusieurs sites de liaison à des facteurs de transcription, ces sites étant souvent séparés par des séquences nucléotidiques sans fonction. Les LCR des VPH-31, 33 et 35 ont respectivement montré des variations intratypiques du LCR de 3.3%, 3.5% et 3.5%, ce qui n'excède pas le degré maximum de diversité qui est de 5%. Le type 33 était plus variable que le type 31, et ce dernier était plus variable que le type 35. Pour les trois types à l'étude, on remarque que la plupart des sites de mutations sont partagés par plusieurs isolats et plusieurs variants. Très peu de mutations sont uniques à un seul variant, et lorsque ce fut le cas, ces mutations ont été de toute façon confirmées. Ceci indique d'abord que les variations observées ne sont pas des artefacts imputables à des erreurs commises par la polymérase utilisée, en l'occurrence la polymérase Tgo ayant une haute stabilité thermique et une activité «proofreading» (taux d'erreur de  $4.9 \times 10^{-7}$ ). Ensuite, cela indique clairement que les variants de ces VPHs n'évoluent pas au hasard, mais suivent un «patron» de variation génomique. D'ailleurs, ce même «patron» a déjà été démontré pour les VPHs des types 6, 16, 18 et 52 [59, 180, 114].

Considérant le fait que les mutations du LCR, de E6 et de E7 sont non-aléatoires et que le phénomène des variations intergènes a déjà été rapporté chez le VPH-16 [175], nous avons vérifié, pour le VPH-31, s'il y avait co-ségrégation des variants des différentes régions ciblées. En d'autres termes, il s'agissait de déterminer si les variants classifiés selon une région du génôme du papillomavirus permettait la classification des variants des autres régions. Le degré de diversité inter-variant du LCR pouvant atteindre 5%, en opposition à 2% pour les cadres de lecture E6 et E7, il était permis de croire que la région non-codante du virus pourrait jouer le rôle de «classificateur». Ainsi, pour le VPH-52, les variants de E6 et de E7 ne permettait pas la classification des variants du

LCR [2]. En opposition, il a été rapporté pour le VPH-16 que la séquence de E6 permettait la classification de variants qui n'étaient pas classifiés par le LCR [175]. Il est important de mentionner que les événements de recombinaison entre différents variants d'un VPH sont extrêmement rares [181], au point de pouvoir éliminer cet élément biaisant de notre analyse. Pour le VPH-31, nous avons pu déterminer que le LCR permettait effectivement la classification de plusieurs variants de E6 et E7. Cependant, trois variants du LCR étaient classifiés par l'analyse de E6 et E7. Ainsi, nous offrons un support aux faits que l'analyse des séquences des cadres de lectures E6 et E7 peuvent être utiles pour discriminer des variants identiques au niveau de la séquence du LCR et que le séquençage de plusieurs régions génomiques des VPHs peut aider à préciser la classification des variants.

Dans un autre ordre d'idées, nous avons observé des différences notables entre les polymorphismes des VPHs-31, 33 et 35. Par exemple, pour le LCR, le nombre médian de variations par variant du type 31 était significativement plus élevé que ceux des types 33 et 35, cependant que ce nombre était similaire chez ces deux derniers. Pour le nombre médian de variations par variant du cadre de lecture E6, il n'y avait pas de différences lors des comparaisons des types 31 et 33 et des types 33 et 35. Cependant, un plus grand nombre de variations était observé chez le VPH-31 par rapport au VPH-35. Quant à E7, on observe qu'il y a moins de variations par variant pour le type 35 en comparaison avec les types 31 et 33.

Alors que des délétions et des insertions étaient retrouvées dans le LCR du type 33, ce type de mutation était plus rare chez le VPH-35 et n'était pas retrouvé chez le VPH-31. Les événements de délétions ne semblent donc pas être une grande source de polymorphisme pour ces derniers types. Des délétions et insertions étaient aussi fréquemment retrouvées dans les LCR du VPH-52 [2]. Cependant, pour le VPH-16, les événements d'insertion dans la région non-codante semblent plus rares et n'ont été rapportés que pour un seul variant, le variant Bb1 [180]. Or, les analyses phylogénétiques révèlent que les VPHs-16, 31 et 35 dérivent de la même sous-branche

phylogénétique, alors que les types 33 et 52 forment une autre sous-branche [16]. Les insertions et délétions comme source de polymorphisme varieraient donc en importance selon la parenté phylogénétique de différents types de VPHs.

Pour toutes les régions analysées, à une exception près, la proportion d'isolats non-prototypique était plus grande pour le VPH-35 que pour les types 31 et 33, et cette proportion était plus grande pour le type 31 que pour le type 33. L'exception se trouve au niveau de la région E7, pour laquelle la proportion d'isolats non-prototypiques ne démontrait pas de différence significative entre les types 33 et 35.

En conséquence des mutations au niveau du LCR, des sites potentiels de liaison à des facteurs de transcription furent gagnés ou perdus pour les trois types. Pour le VPH-31, les éléments les plus notables sont le gain (mutation 7477) et la perte (mutation 7474) de sites de liaison à E2 et la perte d'un site de liaison à C/EBP $\beta$  (mutation 7659). E2 est le répresseur viral de la transcription précoce alors que C/EBP $\beta$  a été décrit comme étant un répresseur cellulaire de la transcription et de la réplication des VPHs [88, 172]. Donc théoriquement, la perte de ces sites pourrait avoir pour conséquence une transcription et une réplication accrue. Pour le VPH-35, le seul élément notoire à ce niveau est la perte d'un site de liaison au facteur de transcription Oct-1 (mutation 7636), connu pour être un répresseur de la transcription du VPH-18 [70].

C'est au niveau du VPH-33 que nous retrouvons la variation génétique la plus étonnante. Le VPH-33 prototypique contient une parfaite répétition en tandem de 78 paires de base dans la région 3' du LCR [19]. Aucune répétition de cette taille n'a jamais été décrite pour aucun autre papillomavirus. Il est suggéré que cette répétition pourrait jouer un rôle comme activateur en *cis* de la transcription virale [19]. Des répétitions en tandem se retrouvent aussi près du promoteur précoce du virus SV40 (72 pb) [52], chez le virus murin du sarcome de Moloney (73 pb) [26] et chez le virus BK (68 pb) [142] et peuvent activer la transcription de ces virus en *cis*. Chez le VPH-33, une telle activation pourrait être médiée par la formation potentielle d'une structure «Z-DNA» dans la séquence répétée [19]. Une telle structure pourrait avoir la propriété et la fonction

d'attirer les facteurs transcriptionnels au niveau des promoteurs eucaryotiques [19]. Or, 7 des variants que nous avons décrit pour le LCR du VPH-33 n'ont pas la duplication en question (13 isolats). De plus, nous avons obtenu une association entre l'absence de la duplication et la persistance du VPH-33, ce qui suggère un rôle suppresseur plutôt qu'activateur pour la duplication *in vivo*. La présence d'un site de liaison au facteur *c-myb* au sein de la duplication expliquerait possiblement ce rôle suppresseur de la duplication. La suppression de la transcription virale par *c-myb* a déjà été démontrée chez le virus SV40 [109] et chez le virus de l'hépatite B [22].

Pour le VPH-31, nous avons tenté d'établir un lien entre certains variants et l'origine ethnique de la femme infectée. Pour le type 16, la construction d'arbres phylogénétiques a clairement démontré la présence de cinq lignées pour ce type (2 lignées africaines, 1 européenne, 1 asiatique et 1 asiatique-américaine), chaque lignée étant associée à une région du monde [68]. Le même travail a été accompli pour le VPH-18 avec des résultats similaires [114]. La présence de telles lignées au sein de l'arbre phylogénétique découle du fait qu'il y a peu de variations observées chez les variants provenant d'une région particulière, alors qu'un haut taux de variation est observé entre les régions. Pour le VPH-31, nous disposions d'isolats provenant de femmes de différentes origines ethniques et vivant au Canada. Les femmes caucasiennes étaient plus fréquemment infectées par des variants prototypiques, alors que les femmes d'origine africaine étaient infectées préférentiellement par des variants non-prototypiques. Plus précisément, le variant 31-LCR-4 était détecté plus souvent chez les femmes d'origine africaine. Il est important de mentionner que le fait qu'un variant soit associé à une branche africaine ne signifie pas qu'il ne peut infecter des individus d'un autre groupe ethnique. C'est pourquoi il aurait été intéressant d'avoir connaissance de l'origine ethnique du ou des partenaires sexuels des femmes de notre cohorte. Cela aurait peut-être permis d'établir une association plus forte entre les variants et l'origine ethnique des femmes infectées par le VPH-31.

L'analyse du polymorphisme des cadres de lecture E6 et E7 a démontré que ceux-ci étaient moins variables que le LCR chez les VPH-31 et 33, à l'instar des types 16, 58 et 52 [45, 14]. Par contre, ce détail n'était pas observé chez le VPH-35, ce qui apporte un autre argument pour affirmer que différents types de VPH peuvent adopter un polymorphisme différent. Dans aucun cas, nous n'avons observé de délétion ou d'insertion dans E6 et E7. Il est vrai que de telles mutations dans des régions codantes ne sont pas favorisées, étant donné le risque de perte de fonction.

Les domaines fonctionnels des protéines E6 et E7 ont été décrits pour le VPH-16, mais non pour les VPH-31, 33 et 35 [38, 129]. Donc, pour évaluer l'impact des mutations significatives retrouvées dans les variants des VPH-31, 33 et 35, nous avons dû aligner les séquences du VPH-16 avec celles des types à l'étude. Évidemment, un tel procédé est insuffisant pour réellement déterminer l'impact d'une mutation sur la fonction d'une protéine s'il n'est pas complété d'une étude *in vitro*. Ces dernières analyses débordent le cadre de notre étude, dont le but est de trouver des mutations qui pourraient avoir un lien avec l'établissement de la persistance. D'autre part, il était difficile d'évaluer l'impact potentiel de chaque mutation individuelle sur la persistance étant donné le trop petit nombre de participantes infectées par les types 31/33/35 dont nous disposions.

Pour le VPH-31, nous avons trouvé 13 mutations dans le cadre de lecture E6, dont 6 étaient significatives. La mutation significative 475 (A pour G; K123R) survenait dans un site potentiel de dégradation de p53. Aucune autre mutation ne semblait toucher un domaine fonctionnel de la protéine E6. D'ailleurs, aucune de ces mutations n'était associée à la persistance. Pour le VPH-33 par contre, pour lequel nous avons 10 variations dans E6 dont 6 entraînaient des mutations significatives de la protéine. La mutation 521 (A138V) se trouvait dans un motif en «doigt de zinc» de la protéine, mais nous ignorons si cette variation a un impact réel sur le comportement biologique de la protéine. En revanche, la mutation 355 (G à T; V83L) apportait un grand intérêt. Celle-ci entraîne le remplacement d'une valine par une leucine en position 83 de la protéine E6.

Une telle mutation a déjà été décrite chez le VPH-16 (variant 350G; L83V) et était associée à la persistance et le développement de LSILs [184]. De plus, cette mutation se retrouvait chez le VPH-52. Cependant, cette association semblait dépendre du bassin de population sur laquelle l'étude portait [184, 183]. Il nous a été impossible d'établir une association entre la mutation 355 de la protéine E6 du VPH-33 et la persistance, mais il est intéressant de constater qu'une variation aussi similaire puisse se retrouver chez trois types différents, bien que très près phylogénétiquement.

La mutation 355, de même que les mutations 214, 315, 325, 387, 521 et 534 ont déjà été décrites dans une étude ultérieure faite sur des femmes japonaises [179]. Ce rapport décrivait 14 mutations en comparaison aux 10 décrites dans notre étude. De plus, 3 de nos mutations n'étaient pas retrouvées dans cette étude, contre 7 que nous n'avions pas. De telles différences pourraient s'expliquer par le petit échantillonnage dont nous disposions, mais l'étude japonaise ne portait elle-même que sur 20 isolats VPH-33+ provenant de femmes cancéreuses. Il est donc possible que ces différences soient plutôt tributaires des 2 populations à l'étude et, en se basant sur les études phylogénétiques des VPH-16 et 18 [68, 114], que l'échantillonnage des variants du VPH-33 se soit fait au sein de deux branches phylogénétiques différentes. Le plus grand nombre de mutations retrouvé chez les variants japonais s'expliquerait alors par le fait que ceux-ci auraient évolué à partir de variants européens, suivant le mouvement des populations humaines.

Pour le cadre de lecture E6 du VPH-35, seulement 2 variations génétiques sur 10 résultaient en un changement d'acide aminé. De même que pour le type 33, aucune de ces mutations ne touchait les régions impliquées dans la liaison et la dégradation de p53 [20].

La protéine E7 est constituée des trois domaines CR1, CR2 et CR3, le domaine CR2 étant le plus important pour la transformation du fait de son implication dans la liaison à pRb [53]. Pour le VPH-31, des mutations significatives ont été identifiées dans les domaines CR2 et CR3. La mutation 626 (C pour T; H23Y) était située dans un domaine potentiel de liaison avec le pRb. Pour le VPH-33, aucune mutation significative

n'était retrouvée dans le domaine CR2. Cependant, deux mutations significatives (586 : K5R et 606 : V12I) étaient situées dans le domaine CR1 et deux autres dans le domaine CR3. Des observations indiquent que le domaine CR1 aurait une importance dans la transformation cellulaire, bien que la mécanique du phénomène n'ait pas encore été élucidée [53]. Quant au domaine CR3, il contiendrait un site de fixation supplémentaire à pRb qui serait suffisant pour empêcher la réclusion de E2F par ce régulateur du cycle cellulaire [124]. Pour le VPH-35, la mutation significative 628 (Y23H) affectait le domaine de liaison à pRb, les deux autres mutations significatives retrouvées pour ce type se trouvant dans la région CR3.

Un des créneaux majeur de notre étude était la persistance en relation avec le polymorphisme des VPH-31/33/35. L'étude de la persistance chez les papillomavirus est d'une grande importance, du fait que la persistance est un déterminant majeur du développement du cancer du col de l'utérus. Effectivement, le cancer prenant plusieurs décennies pour se développer, le virus doit se maintenir suffisamment longtemps dans l'organisme pour causer des lésions malignes. Chez le VPH-16, le polymorphisme génétique était associé à la persistance virale [37]. De plus, des variants du VPH-16 étaient plus fortement associés à l'apparition de lésions de haut grade [178] et au développement du cancer [64]. L'intervalle de temps médian pour la perte de l'infection d'un VPH oncogène étant de 8.9 mois [41], notre définition de la persistance impliquait la détection d'ADN viral sur une période de plus de 9 mois. Les variants étaient regroupés en deux classes : les variants «prototype-like» (avec au plus 1 variation les différenciant du prototype) et les variants «non-prototype» (avec plus d'une mutation).

Pour le VPH-31, nous n'avons pu établir aucun lien significatif entre le polymorphisme de celui-ci et la persistance, et ce à tous les niveaux (LCR, E6 et E7). Les variants non-prototypes n'étaient pas associés avec la persistance comme cela avait été observé pour le VPH-52. De plus, aucun variant et aucune mutation particulière ne semblait montrer de relation avec la persistance. De toute manière, la recherche d'une association entre un variant individuel et la persistance était fortement minée par le trop

petit nombre d'isolats disponibles par variants. Par contre, nous avons pu faire ressortir certains éléments intéressants, bien qu'encore une fois nous n'ayons pas établi de relation significative. Ainsi, parmi les femmes infectées par le variant 31-LCR-4, qui semble être d'origine africaine, aucune n'avait d'infection transitoire contre 6 persistantes. Ce n'est qu'une observation mais néanmoins, il a déjà été démontré chez le VPH-16 que l'infection aux variants non-européens présentait un facteur de risque plus élevé pour l'infection persistante [170] et le développement du cancer [64]. D'autre part, nous avons remarqué une association, bien que non-significative, entre l'infection persistante et un résultat d'examen cytologique anormal. Le développement de lésions de haut grade a déjà été associé à la persistance virale chez le VPH-16 [67].

De par le petit nombre de participantes infectées par les VPH-33 et 35, les résultats obtenus pour ces deux types furent combinés pour évaluer l'impact du polymorphisme génétique sur la persistance. De plus, pour le VPH-35, aucune participante n'était infectée par des variants prototypes ou «prototype-like» au niveau du LCR et du cadre de lecture E6. Nos analyses ont donc porté sur le cadre de lecture E7. Ainsi, nous avons établi que pour ces types, l'infection par un variant non-prototype était significativement associée à la persistance avec un risque relatif de 2 (95% IC, 1.35-2.97). C'est la première fois qu'il est rapporté que le polymorphisme du cadre de lecture E7 est lié à la persistance. Encore une fois, il fut impossible d'établir une relation entre des variants ou des mutations individuelles et la persistance, à une exception près. En effet, tel que décrit précédemment, l'absence de la duplication de 78 paires de base dans le LCR du VPH-33 était associée à un risque relatif pour la persistance de 1.8 (95% IC, 1.19-1.72,  $p=0.03$ ), ce qui suggère un rôle suppresseur de la duplication sur la transcription virale. Cette suppression pourrait se faire via la liaison du facteur de transcription *c-myb* sur des sites potentiels de fixation retrouvés dans la duplication.



## CONCLUSION

Cette étude présente plusieurs intérêts. D'abord, bien qu'il y ait eu quelques publications sur le VPH à haut risque de type 31, aucune n'a jamais porté sur son polymorphisme. Pour le VPH-33, quelques études ont touché très succinctement aux variations au niveau de la protéine de la capsid L1 et du cadre de lecture E6. Enfin, le polymorphisme du VPH-35 n'a été abordé que dans le cadre d'une étude portant sur le polymorphisme de 12 types de VPH au niveau de L1. Depuis que la recherche sur les VPHs a pris son essor, seuls les VPH-16 et 18 ont fait l'objet d'études exhaustives. Il est vrai que ce sont généralement eux qui sont le plus souvent retrouvés dans les cas de cancer du col utérin. Cependant, l'étude d'autres types que ceux-ci peut permettre d'observer des variations au sein de la grande famille des papillomavirus au niveau du polymorphisme et du comportement biologique. Notre étude souligne d'ailleurs plusieurs variations dans le polymorphisme de types qui sont pourtant très apparentés phylogénétiquement. D'autre part, nos observations apportent plusieurs nouveautés en rapport avec le polymorphisme et la persistance.

D'abord, contrairement à ce qu'il fut rapporté pour d'autres types, le polymorphisme du VPH-31 n'était pas associé à la persistance virale. Le statut VIH des participantes ne semblait pas non plus avoir d'influence sur l'infection à ce type. Notre rapport est aussi le premier à rapporter une association entre la variation génétique du cadre de lecture E7 et la persistance. Cette association était observée chez les types 33 et 35. Enfin, bien que la présence d'une duplication de 78 paires de base dans le LCR du VPH-33 était déjà connue, l'existence de variants n'ayant pas cette duplication constitue un fait nouveau et extrêmement intéressant qui mérite qu'on s'y attarde. Ainsi, il pourrait y avoir un grand intérêt à étudier l'influence de la présence ou absence de cette duplication sur le développement de lésions de haut grade ou sur le développement du cancer du col utérin.

## RÉFÉRENCES

### Journaux scientifiques

1. Ahdieh L, Klein RS, Burk R, et al: Prevalence, incidence, and type-specific persistence of human papillomavirus in human immunodeficiency virus (HIV)-positive and HIV-negative women. *J Inf Dis* 2001; 184: 682-690.
2. Aho J, Hankins C, Tremblay C, et al: Molecular analysis of human papillomavirus type 52 infection in HIV-infected women. *Journal of Infectious Diseases*; submitted 2002
3. Arany I, Grattendick KG, Tyring SK: Interleukin-10 induces transcription of the early promoter of human papillomavirus type 16 (HPV16) through the 5'-segment of the upstream regulatory region (URR). *Antiviral Res* 2002; 55: 331-339.
4. Asin SN, Wildt-Perinic D, Mason SI, Howell AL, Wira CR, Fanger MW: Human immunodeficiency virus type 1 infection of human uterine epithelial cells: viral shedding and cell contact-mediated infectivity. *J Infect Dis* 2003; 187: 1522-1533.
5. Barbosa MS, Vass WC, Lowy DR, Schiller JT: In vitro biological activities of the E6 and E7 genes vary among human papillomaviruses of different oncogenic potential. *J Virol* 1991; 65: 292-298.
6. Beckman G, Birgander R, Sjalander A, et al: Is p53 polymorphism maintained by natural selection? *Hum Hered* 1994; 44: 266-270.

7. Bible JM, Mant C, Best JM, et al: Cervical lesions are associated with human papillomavirus type 16 intratypic variants that have high transcriptional activity and increased usage of common mammalian codons. *J Gen Virol* 2000; 81 Pt 6: 1517-1527.
8. Blanton RA, Perez-Reyes N, Merrick DT, McDougall JK: Epithelial cells immortalized by human papillomaviruses have premalignant characteristics in organotypic culture. *Am J Pathol* 1991; 138: 673-685.
9. Bosch FX, Castellsague X, Munoz N, et al: Male sexual behavior and human papillomavirus DNA: key risk factors for cervical cancer in Spain. *J Natl Cancer Inst* 1996; 88: 1060-1067.
10. Bosch FX, Lorincz A, Munoz N, Meijer CJLM, Shah KV: The causal relation between human papillomavirus and cervical cancer. *J Clin Pathol* 2002; 55: 244-265.
11. Bosch FX, Manos MM, Muñoz N, et al: Prevalence of human papillomavirus in cervical cancer: A worldwide perspective. *J Natl Cancer Inst* 1995; 87: 796-802.
12. Brinton LA, Reeves WC, Brenes MM, et al: The male factor in the etiology of cervical cancer among sexually monogamous women. *Int J Cancer* 1989; 44: 199-203.
13. Brisson J, Bairati I, Morin C, et al: Determinants of persistent detection of Human papillomavirus DNA in the uterine cervix. *J Inf Dis* 1996; 173: 794-799.

14. Chan PK, Lam CW, Cheung TH, et al: Association of human papillomavirus type 58 variant with the risk of cervical cancer. *J Natl Cancer Inst* 2002; 94: 1249-1253.
15. Chan S-Y, Ho L, Ong C-K, et al: Molecular variants of human papillomavirus type 16 from four continents suggest ancient pandemic spread of the virus and its coevolution with humankind. *J Virol* 1992; 66: 2057-2066.
16. Chan SY, Bernard HU, Ong CK, Chan SP, Hoffman B, Delius H: Phylogenetic analysis of 48 papillomavirus types and 28 subtypes and variants: a showcase for the molecular evolution of DNA viruses. *J Virol* 1992; 66: 5714-5725.
17. Chen EY, Howley PM, Levinson AD, Seeburg PH: The primary structure and genetic organization of the bovine papillomavirus type 1 genome. *Nature* 1982; 299: 529-534.
18. Chow LT, Reilly SS, Broker TR, Taichman LB: Identification and mapping of human papillomavirus type 1 RNA transcripts recovered from plantar warts and infected epithelial cell cultures. *J Virol* 1987; 61: 1913-1918.
19. Cole ST, Streek RE: Genome organization and nucleotide sequence of human papillomavirus type 33, which is associated with cervical cancer. *J Virol* 1986; 58: 991-995.
20. Crook T, Tidy JA, Vousden KH: Degradation of p53 can be targeted by HPV E6 sequences distinct from those required for p53 binding and trans-activation. *Cell* 1991; 67: 547-556.
21. Cuthill S, Maiman M, Fruchter RG, Lopatinsky I, Cheng CC: Complications after treatment of cervical intraepithelial neoplasia in women infected

- with the human immunodeficiency virus. *J Reprod Med JID* 1995; 40: 823-828.
22. David-Cordonnier MH, Hamdane M, D'Halluin JC: c-Myb protein binds to the EP element of the HBV enhancer and regulates transcription in synergy with NF-M. *Biochim Biophys Acta* 1999; 1446: 82-92.
  23. DeGregori J, Leone G, Ohtani K, Miron A, Nevins JR: E2F-1 accumulation bypasses a G1 arrest resulting from the inhibition of G1 cyclin-dependent kinase activity. *Genes Dev* 1995; 9: 2873-2887.
  24. Delmas MC, Larsen C, van Benthem B, et al: Cervical squamous intraepithelial lesions in HIV-infected women: prevalence, incidence and regression. European Study Group on Natural History of HIV Infection in Women. *AIDS* 2000; 14: 1775-1784.
  25. Demeret C, Yaniv M, Thierry F: The E2 transcriptional repressor can compensate for Sp1 activation of the human papillomavirus type 18 early promoter. *J Virol* 1994; 68: 7075-7082.
  26. Dhar R, McClements WL, Enquist LW, Vande WG: Nucleotide sequences of integrated Moloney sarcoma provirus long terminal repeats and their host and viral junctions. *Proc Natl Acad Sci U S A* 1980; 77: 3937-3941.
  27. Dolei A, Curreli S, Marongiu P, et al: Human immunodeficiency virus infection in vitro activates naturally integrated human papillomavirus type 18 and induces synthesis of the L1 capsid protein. *J Gen Virol* 1999; 80: 2937-2944.

28. Dong XP, Pfister H: Overlapping YY1- and aberrant SP1-binding sites proximal to the early promoter of human papillomavirus type 16. *J Gen Virol* 1999; 80 ( Pt 8): 2097-2101.
29. Dong XP, Stubenrauch F, Beyer-Finkler E, Pfister H: Prevalence of deletions of YY1-binding sites in episomal HPV 16 DNA from cervical cancers. *Int J Cancer* 1994; 58: 803-808.
30. Doorbar J, Ely S, Sterling J, McLean C, Crawford L: Specific interaction between HPV-16 E1-E4 and cytokeratins results in collapse of the epithelial cell intermediate filament network. *Nature* 1991; 352: 824-827.
31. Durst M, Gissman L, Thernberg H, zur Hausen H: A papillomavirus DNA from a cervical carcinoma and its prevalence in cancer biopsy samples from different geographic regions. *Proc Natl Acad Sci USA* 1983; 80: 3812-3815.
32. Durst M, Glitz D, Schneider A, zur HH: Human papillomavirus type 16 (HPV 16) gene expression and DNA replication in cervical neoplasia: analysis by in situ hybridization. *Virology* 1992; 189: 132-140.
33. Dyson N, Howley PM, Munger K, Harlow E: The human papillomavirus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product. *Science* 1989; 243: 934-947.
34. Ellerbrock TV, Chiasson MA, Bush TJ, et al: Incidence of cervical squamous intraepithelial lesions in HIV-infected women. *JAMA* 2000; 283: 1031-1037.

35. Evander M, Edlund K, Gustafsson A, et al: Human papillomavirus infection is transient in young women: population-based cohort study. *J Inf Dis* 1995; 171: 1026-1030.
36. Evander M, Frazer IH, Payne E, Qi YM, Hengst K, McMillan NA: Identification of the alpha6 integrin as a candidate receptor for papillomaviruses. *J Virol* 1997; 71: 2449-2456.
37. Ferenczy A, Franco EL: Persistent human papillomavirus infection and cervical neoplasia. *Lancet Oncology* 2002; 3: 11-16.
38. Foster SA, Demers GW, Etscheid BG, Galloway DA: The ability of human papillomavirus E6 proteins to target p53 for degradation in vivo correlates with their ability to abrogate actinomycin D-induced growth arrest. *J Virol* 1994; 68: 5698-5705.
39. Franco EL, Duarte-Franco E, Ferenczy A: Cervical cancer: epidemiology, prevention and the role of human papillomavirus infection. *CMAJ: Canadian Medical Association Journal* 2001; 164: 1017-1025.
40. Franco EL, Rohan TE, Villa LL: Epidemiologic evidence and human papillomavirus infection as a necessary cause of cervical cancer. *J Natl Cancer Inst* 1999; 91: 506-511.
41. Franco EL, Villa LL, Sobrinho JP, et al: Epidemiology of acquisition and clearance of cervical Human Papillomavirus infection in women from a high-risk area for cervical cancer. *J Inf Dis* 1999; 180: 1415-1423.
42. Fruchter RG, Maiman M, Sedlis A, Bartley L, Camilien L, Arrastia CD: Multiple recurrences of cervical intraepithelial neoplasia in women with the human immunodeficiency virus. *Obstet Gynecol* 1996; 87: 338-344.

43. Fujinaga Y, Okazawa K, Nishikawa A, et al: Sequence variation of human papillomavirus type 16 E7 in preinvasive and invasive cervical neoplasias. *Virus Genes* 1994; 9: 85-92.
44. Geng Y, Eaton EN, Picon M, et al: Regulation of cyclin E transcription by E2Fs and retinoblastoma protein. *Oncogene* 1996; 12: 1173-1180.
45. Giannoudis A, Herrington CS: Human papillomavirus variants and squamous neoplasia of the cervix. *J Pathol* 2001; 193: 295-302.
46. Giannoudis A, van Duin M, Snijders PJF, Herrington CS: Variation in the E2-binding domain of HPV 16 is associated with high-grade squamous intraepithelial lesions of the cervix. *British J Cancer* 2001; 84: 1058-1063.
47. Giri I, Yaniv M: Structural and mutational analysis of E2 trans-activating proteins of papillomaviruses reveals three distinct functional domains. *EMBO J* 1988; 7: 2823-2829.
48. Gissmann L, Hausen HZ: Human papilloma virus DNA: physical mapping and genetic heterogeneity. *Proc Natl Acad Sci U S A* 1976; 73: 1310-1313.
49. Gissmann L, Pfister H, zur HH: Human papilloma viruses (HPV): characterization of four different isolates. *Virology* 1977; 76: 569-580.
50. Grassmann K, Kratzer F, Petry KU, Iftner T: Functional characterization of naturally occurring mutants of human papillomavirus type 16 with large deletions in the non-coding region [letter]. *Int J Cancer* 1996; 68: 265-269.



51. Greider CW: Telomere length regulation. *Annu Rev Biochem* 1996; 65: 337-365.
52. Gruss P, Dhar R, Khoury G: Simian virus 40 tandem repeated sequences as an element of the early promoter. *Proc Natl Acad Sci U S A* 1981; 78: 943-947.
53. Gulliver GA, Herber RL, Liem A, Lambert PF: Both conserved region 1 (CR1) and CR2 of the human papillomavirus type 16 E7 oncogene are required for induction of epidermal hyperplasia and tumor formation in transgenic mice. *J Virol* 1997; 71: 5905-5914.
54. Hagensee ME, Olson NH, Baker TS, Galloway DA: Three-dimensional structure of vaccinia virus-produced human papillomavirus type 1 capsids. *J Virol* 1994; 68: 4503-4505.
55. Hagensee ME, Yaegashi N, Galloway DA: 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. *J Virol* 1993; 67: 315-322.
56. Hanahan D, Weinberg RA: The hallmarks of cancer. *Cell* 2000; 100: 57-70.
57. Harris RW, Brinton LA, Cowdell RH, et al: Characteristics of women with dysplasia or carcinoma in situ of the cervix uteri. *Br J Cancer* 1980; 42: 359-369.
58. Heck DV, Yee CL, Howley PM, Munger K: Efficiency of binding the retinoblastoma protein correlates with the transforming capacity of the E7 oncoproteins of the human papillomaviruses. *Proc Natl Acad Sci U S A* 1992; 89: 4442-4446.

59. Heinzl PA, Chan SY, Ho L, et al: Variation of human papillomavirus type 6 (HPV-6) and HPV-11 genomes sampled throughout the world. *J Clin Microbiol* 1995; 33: 1746-1754.
60. Helland A, Borresen AL, Kaern J, Ronningen KS, Thorsby E: HLA antigens and cervical carcinoma. *Nature* 1992; 356: 23-Helland, A.
61. Hemminki K, Li X, Mutanen P: Age-incidence relationships and time trends in cervical cancer in Sweden. *Eur J Epidemiol* 2001; 17: 323-328.
62. Herrero R, Potischman N, Brinton LA, et al: A case-control study of nutrient status and invasive cervical cancer. I. Dietary indicators. *Am J Epidemiol* 1991; 134: 1335-1346.
63. Herrington CS: Human papillomaviruses and cervical neoplasia. Interaction of HPV with other factors. *Clin Pathol* 1995; 48: 1-6.
64. Hildesheim A, Schiffman M, Bromley C, et al: Human papillomavirus type 16 variants and risk of cervical cancer. *J Natl Cancer Inst* 2001; 93: 315-318.
65. Hildesheim A, Schiffman MH, Gravitt PE, et al: Persistence of type-specific human papillomavirus infection among cytologically normal women. *J Inf Dis* 1994; 169: 235-240.
66. Ho GY, Bierman R, Beardsley L, Chang CJ, Burk RD: Natural history of cervicovaginal papillomavirus infection in young women. *N Engl J Med* 1998; 338: 423-428.

67. Ho GYF, Burk RD, Klein S, et al: Persistent genital human papillomavirus infection as a risk factor for persistent cervical dysplasia. *J Natl Cancer Inst* 1995; 87: 1365-1371.
68. Ho L, Chan SY, Burk RD, 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* 1993; 67: 6413-6423.
69. Ho WZ, Harouse JM, Rando RF, Gonczol E, Srinivasan A, Plotkin SA: Reciprocal enhancement of gene expression and viral replication between human cytomegalovirus and human immunodeficiency virus type 1. *J Gen Virol* 1990; 71 ( Pt 1): 97-103.
70. Hoppe-Seyler F, Butz K, zur HH: Repression of the human papillomavirus type 18 enhancer by the cellular transcription factor Oct-1. *J Virol* 1991; 65: 5613-5618.
71. Hu X, Pang T, Guo Z, et al: HPV16 gene variations in invasive cervical squamous cell carcinoma and cancer in situ from Russian patients. *British J Cancer* 2001; 84: 791-795.
72. Jamieson DJ, Duerr A, Burk R, et al: Characterization of genital human papillomavirus infection in women who have or who are at risk of having HIV infection. *Am.J.Obstet.Gynecol.* 2002; 186: 21-27.
73. Jenson AB, Rosenthal JD, Olson C, Pass F, Lancaster WD, Shah K: Immunologic relatedness of papillomaviruses from different species. *J Natl Cancer Inst* 1980; 64: 495-500.

74. Jeon S, Allen-Hoffmann BL, Lambert PF: Integration of human papillomavirus type 16 into the human genome correlates with a selective growth advantage of cells. *J Virol* 1995; 69: 2989-2997.
75. Kabsch K, Alonso A: The human papillomavirus type 16 E5 protein impairs TRAIL- and FasL-mediated apoptosis in HaCaT cells by different mechanisms. *J Virol* 2002; 76: 12162-12172.
76. Kammer C, Tommasino M, Syrjanen S, et al: Variants of the long control region and the E6 oncogene in European human papillomavirus type 16 isolates: implications for cervical disease. *Br J Cancer* 2002; 86: 269-273.
77. Kammer C, Warthorst U, Torrez-Martinez N, Wheeler CM, Pfister H: Sequence analysis of the long control region of human papillomavirus type 16 variants and functional consequences for P97 promoter activity. *J Gen Virol* 2000; 81: 1975-1981.(Abstract)
78. Kanaya T, Kyo S, Laimins LA: 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* 1997; 237: 159-169.
79. Kim NW, Piatyszek MA, Prowse KR, et al: Specific association of human telomerase activity with immortal cells and cancer. *Science* 1994; 266: 2011-2015.
80. Kirnbauer R, Taub J, Greenstone H, et al: Efficient self-assembly of human papillomavirus type 16 L1 and L1-L2 into virus-like particles. *J Virol* 1993; 67: 6929-6936.

81. Kiviat NB, Koutsky LA: Specific human papillomavirus types as the causal agents of most cervical intraepithelial neoplasia: implications for current views and treatment. *J Natl Cancer Inst* 1993; 85: 934-935.
82. Kjellberg L, Wang Z, Wiklund F, et al: Sexual behaviour and papillomavirus exposure in cervical intraepithelial neoplasia: a population-based case-control study. *J Gen Virol* 1999; 80: 391-398.
83. Klein RS, Ho GYF, Vermund SH, Fleming I, Burk RD: Risk factors for squamous intraepithelial lesions on pap smear in women at risk for human immunodeficiency virus infection. *Journal of Infectious Disease* 1994; 170: 404-409.
84. Klingelutz AJ, Foster SA, McDougall JK: Telomerase activation by the E6 gene product of human papillomavirus type 16. *Nature* 1996; 380: 79-82.
85. Koutsky LA, Ault KA, Wheeler CM, et al: A controlled trial of a human papillomavirus type 16 vaccine. *N Engl J Med* 2002; 347: 1645-1651.
86. Kurman RJ, Malkasian GDJ, Sedlis A, Solomon D: From Papanicolaou to Bethesda: the rationale for a new cervical cytologic classification. *Obstet Gynecol* 1991; 77: 779-782.
87. Kurvinen K, Yliskoski M, Saarikoski S, Syrjanen K, Syrjanen S: Variants of the long control region of human papillomavirus type 16. *Eur J Cancer* 2000; 36: 1402-1410.
88. 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. *J Virol* 1993; 67: 1058-1066.

89. La Vecchia C, Franceschi S, Decarli A: Papanicolaou smears and risk of cervical neoplasia. *Lancet* 1984; 2: 1335-1338.
90. Lacey JV, Brinton LA, Abbas FM, et al: Oral contraceptives as risk factors for cervical adenocarcinomas and squamous cell carcinomas. *Cancer Epidemiol Biom Prev* 1999; 8: 1079-1085.
91. Laimins LA: The biology of human papillomaviruses: from warts to cancer. *Infect Agents Dis* 1993; 2: 74-86.
92. Lambert PF, Hubbert NL, Howley PM, Schiller JT: Genetic assignment of multiple E2 gene products in bovine papillomavirus-transformed cells. *J Virol* 1989; 63: 3151-3154.
93. Lee D, Lee B, Kim J, Kim DW, Choe J: cAMP response element-binding protein-binding protein binds to human papillomavirus E2 protein and activates E2-dependent transcription. *J Biol Chem* 2000; 275: 7045-7051.
94. Lehr E, Brown DR: Infection with the oncogenic human papillomavirus type 59 alters protein components of the cornified cell envelope. *Virology* 2003; 309: 53-60.
95. Maiman M, Fruchter RG, Serur E, Remy JC, Feuer G, Boyce J: Human immunodeficiency virus infection and cervical neoplasia. *Gyn Oncol* 1990; 38: 377-382.
96. Malim MH, Hauber J, Le SY, Maizel JV, Cullen BR: The HIV-1 rev transactivator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature* 1989; 338: 254-257.

97. Matlashewski G: The cell biology of human papillomavirus transformed cells. *Anticancer Res* 1989; 9: 1447-1456.
98. 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. *EMBO Journal* 1994; 13: 1460-1466.
99. Mayrand MH, Coutlee F, Hankins C, et al: Detection of human papillomavirus type 16 DNA in consecutive genital samples does not always represent persistent infection as determined by molecular variant analysis. *J Clin Microbiol* 2000; 38: 3388-3393.
100. McBride AA, Romanczuk H, Howley PM: The papillomavirus E2 regulatory proteins. *J Biol Chem* 1991; 266: 18411-18414.
101. Minkoff H, Ahdieh L, Massad LS, et al: The effect of highly active antiretroviral therapy on cervical cytologic changes associated with oncogenic HPV among HIV-infected women. *AIDS* 2001; 15: 2157-2164.
102. Minkoff H, Feldman J, DeHovitz J, Landesman S, Burk R: A longitudinal study of human papillomavirus carriage in human immunodeficiency virus-infected and human immunodeficiency virus-uninfected women. *American Journal of Obstetrics & Gynecology* 1998; 178: 982-986.
103. Mittal R, Pater A, Pater MM: Multiple human papillomavirus type 16 glucocorticoid response elements functional for transformation, transient expression, and DNA-protein interactions. *J Virol* 1993; 67: 5656-5659.
104. Munger K, Phelps WC, Bubb V, Howley PM, Schlegel R: The E6 and E7 genes of the human papillomavirus type 16 together are necessary and sufficient

- for transformation of primary human keratinocytes. *J Virol* 1989; 63: 4417-4421.
105. Munger K, Werness BA, Dyson N, Phelps WC, Harlow E, Howley PM: Complex formation of human papillomavirus E7 proteins with the retinoblastoma tumor suppressor gene product. *EMBO J* 1989; 8: 4099-4105.
  106. Munoz N: Human papillomavirus and cancer: the epidemiological evidence. *Journal of Clinical Virology* 2000; 19: 1-5.
  107. Nakagawa M, Stites DP, Farhat S, et al: T-cell proliferative response to human papillomavirus type 16 peptides: relationship to cervical intraepithelial neoplasia. *Clin Diag Lab Immunol* 1996; 3: 205-210.
  108. Nakagawa M, Stites DP, Patel S, et al: Persistence of human papillomavirus type 16 infection is associated with lack of cytotoxic T lymphocyte response to the E6 antigens. *J Infect Dis* 2000; 182: 595-598.
  109. Nakagoshi H, Nagase T, Kanei-Ishii C, Ueno Y, Ishii S: Binding of the c-myc proto-oncogene product to the simian virus 40 enhancer stimulates transcription. *J Biol Chem* 1990; 265: 3479-3483.
  110. Nakahara T, Nishimura A, Tanaka M, Ueno T, Ishimoto A, Sakai H: Modulation of the cell division cycle by human papillomavirus type 18 E4. *J Virol* 2002; 76: 10914-10920.
  111. Nasiell K, Roger V, Nasiell M: Behavior of mild cervical dysplasia during long-term follow-up. *Obstet Gynecol* 1986; 67: 665-669.



112. Nelson LM, Rose RC, Moroianu J: The L1 major capsid protein of human papillomavirus type 11 interacts with Kap beta2 and Kap beta3 nuclear import receptors. *Virology* 2003; 306: 162-169.
113. Nindl I: HPV variants and risk of cervical cancer. *Papillomavirus Report* 2002; 13: 1-6.
114. Ong CK, Chan SY, Campo MS, et al: Evolution of human papillomavirus type 18: an ancient phylogenetic root in Africa and intratype diversity reflect coevolution with human ethnic groups. *J Virol* 1993; 67: 6424-6431.
115. Palefsky JM: Human papillomavirus infection and anogenital neoplasia in human immunodeficiency virus-positive men and women. *J Natl Cancer Inst* 1998; Monographs.: 15-20.
116. Palella FJJ, Delaney KM, Moorman AC, et al: Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. *N Engl J Med* 1998; 338: 853-860.
117. Park JS, Hwang ES, Lee CJ, et al: Mutational and functional analysis of HPV-16 URR derived from Korean cervical neoplasia. *Gyn Oncol* 1999; 74: 23-29.
118. Park P, Copeland W, Yang L, Wang T, Botchan MR, Mohr IJ: The cellular DNA polymerase alpha-primase is required for papillomavirus DNA replication and associates with the viral E1 helicase. *Proc Natl Acad Sci U S A* 1994; 91: 8700-8704.
119. Park TW, Fujiwara H, Wright TC: Molecular biology of cervical cancer and its precursors. *Cancer* 1995; 76: 1902-1913.

120. Parkin DM, Bray F, Ferlay J, Pisani P: Estimating the world cancer burden: Globocan 2000. *Int J Cancer* 2001; 94: 153-156.
121. Parkin DM, Muir CS: Cancer Incidence in Five Continents. Comparability and quality of data. *IARC Sci Publ* 1992; 45-173.
122. Parkin DM, Pisani P, Ferlay J: Estimates of the worldwide incidence of 25 major cancers in 1990. *Int J Cancer* 1999; 80: 827-841.
123. Pastrana DV, Vass WC, Lowy DR, Schiller JT: NHPV16 VLP vaccine induces human antibodies that neutralize divergent variants of HPV16. *Virology* 2001; 279: 361-369.
124. Patrick DR, Oliff A, Heimbrook DC: Identification of a novel retinoblastoma gene product binding site on human papillomavirus type 16 E7 protein. *J Biol Chem* 1994; 269: 6842-6850.
125. Pecoraro G, Morgan LD, defendi V: Evolution of in vitro transformation and tumorigenesis of HPV16 and HPV18 immortalized primary cervical epithelial cells. *Am J Pathol* 1991; 138: 1-7.
126. Peitsaro P, Johansson B, Syrjanen S: Integrated human papillomavirus type 16 is frequently found in cervical cancer precursors as demonstrated by a novel quantitative real-time PCR technique. *J Clin Microbiol* 2002; 40: 886-891.
127. Peng YC, Breiding DE, Sverdrup F, Richard J, Androphy EJ: AMF-1/Gps2 binds p300 and enhances its interaction with papillomavirus E2 proteins. *J Virol* 2000; 74: 5872-5879.

128. Penn I: Cancer in the immunosuppressed organ recipient. *Transplant Proc* 1991; 23: 1771-1772.
129. Phelps WC, Munger K, Yee CL, Barnes JA, Howley PM: Structure-function analysis of the human papillomavirus type 16 E7 oncoprotein. *J Virol* 1992; 66: 2418-2427.
130. Pim D, Collins M, Banks L: Human papillomavirus type 16 E5 gene stimulates the transforming activity of the epidermal growth factor receptor. *Oncogene* 1992; 7: 27-32.
131. Rheinwald JG, Green H: Epidermal growth factor and the multiplication of cultured human epidermal keratinocytes. *Nature* 1977; 265: 421-424.
132. Romanczuk H, Howley PM: Disruption of either the E1 or the E2 regulatory gene of human papillomavirus type 16 increases viral immortalization capacity. *Proc Natl Acad Sci U S A* 1992; 89: 3159-3163.
133. Romanczuk H, Villa LL, Schlegel R, Howley PM: The viral transcriptional regulatory region upstream of the E6 and E7 is a major determinant of the differential immortalization activities of human papillomavirus types 16 and 18. *J Virol* 1991; 65: 2739-2744.
134. Rose BR, Thompson CH, Zhang J, et al: Sequence variation in the upstream regulatory region of HPV 18 isolates from cervical cancers. *Gyn Oncol* 1997; 66: 282-289.
135. Rotkin ID: A comparison review of the epidemiological studies in cervical cancer related to current searches for transmissible agents. *Cancer Res* 1973; 33: 1353-1367.

136. Rowson KE, Mahy BW: Human papova (wart) virus. *Bacteriol Rev* 1967; 31: 110-131.
137. Ruutu M, Peitsaro P, Johansson B, Syrjanen S: Transcriptional profiling of a human papillomavirus 33-positive squamous epithelial cell line which acquired a selective growth advantage after viral integration. *Int J Cancer* 2002; 100: 318-326.
138. Saccucci P, Mastrone M, Are P, Pisani G, Provenza C: Rapidly progressive squamous cell carcinoma of the cervix in a patient with acquired immunodeficiency syndrome: case report. *Eur J Gynaecol Oncol* 1996; 17: 306-308.
139. Scheffner M, Werness BA, Huilbregtse JM, Levine AJ, Howley PM: The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell* 1990; 63: 1129-1136.
140. Schlegel R, Phelps WC, Zhang YL, Barbosa M: Quantitative keratinocyte assay detects two biological activities of human papillomavirus DNA and identifies viral types associated with cervical carcinoma. *EMBO J* 1988; 7: 3181-3187.
141. Schwarz E, Freese UK, Gissmann L, et al: Structure and transcription of human papillomavirus sequences in cervical carcinoma cells. *Nature* 1985; 314: 111-114.
142. Seif I, Khoury G, Dhar R: The genome of human papovavirus BKV. *Cell* 1979; 18: 963-977.

143. Seo YS, Muller F, Lusk M, Hurwitz J: Bovine papilloma virus (BPV)-encoded E1 protein contains multiple activities required for BPV DNA replication. *Proc Natl Acad Sci U S A* 1993; 90: 702-706.
144. Shah KV, Solomon L, Daniel R, Cohn S, Vlahov D: Comparison of PCR and Hybrid capture methods for detection of human papillomavirus in injection drug-using women at high risk of human immunodeficiency virus infection. *J Clin Microbiol* 1997; 35: 517-519.
145. Sibbet G, Romero-Graillet C, Meneguzzi G, Saveria CM: alpha6 integrin is not the obligatory cell receptor for bovine papillomavirus type 4. *J Gen Virol* 2000; 81 Pt 6: 1629
146. Smith JS, Munoz N, Herrero R, et al: Evidence for Chlamydia trachomatis as a human papillomavirus cofactor in the etiology of invasive cervical cancer in Brazil and the Philippines. *J Infect Dis* 2002; 185: 324-331.
147. Southern SA, McDicken IW, Herrington CS: Evidence for keratinocyte immortalization in high-grade squamous intraepithelial lesions of the cervix infected with high-risk human papillomaviruses. *Lab Invest* 2000; 80: 539-544.
148. Spruck CH, Won KA, Reed SI: Deregulated cyclin E induces chromosome instability. *Nature* 1999; 401: 297-300.
149. Stewart AC, Eriksson AM, Manos MM, et al: Intratype variation in 12 human papillomavirus types: a worldwide perspective. *J Virol* 1996; 70: 3127-3136.

150. Stoler MH, Broker TR: In situ hybridization detection of human papillomavirus DNAs and messenger RNAs in genital condylomas and a cervical carcinoma. *Hum Pathol* 1986; 17: 1250-1258.
151. Stoppler MC, Ching K, Stoppler H, Clancy K, Schlegel R, Icenogle J: 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* 1996; 70: 6987-6993.
152. Storey A, Thomas M, Kalita A, et al: Role of a p53 polymorphism in the development of human papillomavirus-associated cancer. *Nature* 1998; 393: 229-234.
153. Straight SW, Herman B, McCance DJ: The E5 oncoprotein of human papillomavirus type 16 inhibits the acidification of endosomes in human keratinocytes. *J Virol* 1995; 69: 3185-3192.
154. Straight SW, Hinkle PM, Jewers RJ, McCance DJ: The E5 oncoprotein of human papillomavirus type 16 transforms fibroblasts and effects the downregulation of the epidermal growth factor receptor in keratinocytes. *J Virol* 1993; 67: 4521-4532.
155. Strickler HD, Palefsky JM, Shah KV, et al: Human papillomavirus type 16 and immune status in human immunodeficiency virus-seropositive women. *J Natl Cancer Inst* 2003; 95: 1062-1071.
156. Su W, Jackson S, Tjian R, Echols H: DNA looping between sites for transcriptional activation: self-association of DNA-bound Sp1. *Genes Dev* 1991; 5: 820-826.

157. Sun XW, Kuhn L, Ellerbrock TV, Chiasson MA, Bush TJ, Wright TC, Jr.: Human papillomavirus infection in women infected with the human immunodeficiency virus. *New Engl J Med* 1997; 337: 1343-1349.
158. Tan SH, Leong LE, Walker PA, Bernard HU: The human papillomavirus type 16 E2 transcription factor binds with low cooperativity to two flanking sites and represses the E6 promoter through displacement of Sp1 and TFIID. *J Virol* 1994; 68: 6411-6420.
159. Tan W, Schwartz S: The Rev protein of human immunodeficiency virus type 1 counteracts the effect of an AU-rich negative element in the human papillomavirus type 1 late 3' untranslated region. *J Virol* 1995; 69: 2932-2945.
160. Thierry F, Yaniv M: The BPV1-E2 trans-acting protein can be either an activator or a repressor of the HPV18 regulatory region. *EMBO J* 1987; 6: 3391-3397.
161. Thorner LK, Lim DA, Botchan MR: DNA-binding domain of bovine papillomavirus type 1 E1 helicase: structural and functional aspects. *J Virol* 1993; 67: 6000-6014.
162. Turek LP: The structure, function, and regulation of papillomaviral genes in infection and cervical cancer. *Adv Virus Res* 1994; 44: 305-356.
163. Ustav M, Ustav E, Szymanski P, Stenlund A: Identification of the origin of replication of bovine papillomavirus and characterization of the viral origin recognition factor E1. *EMBO J* 1991; 10: 4321-4329.
164. van Duin M, Snidjers PJF, Vossen MTM, et al: Analysis of human papillomavirus type 16 E6 variants in relation to p53 codon 72

- polymorphism genotypes in cervical carcinogenesis. *J Gen Virol* 2000; 81: 317-325.
165. Vaziri H, Schachter F, Uchida I, et al: Loss of telomeric DNA during aging of normal and trisomy 21 human lymphocytes. *Am J Hum Genet* 1993; 52: 661-667.
166. Veldman T, Liu X, Yuan H, Schlegel R: Human papillomavirus E6 and Myc proteins associate in vivo and bind to and cooperatively activate the telomerase reverse transcriptase promoter. *Proc Natl Acad Sci U S A* 2003; 100: 8211-8216.
167. Veress G, Szarka K, Dong XP, Gergely L, Pfister H: Functional significance of sequence variation in the E2 gene and the long control region of human papillomavirus type 16. *J Gen Virol* 1999; 80: 1035-1043.
168. Vermund SH, Kelly KF, Klein RS, et al: High risk of human papillomavirus infection and cervical squamous intraepithelial lesions among women with symptomatic human immunodeficiency virus infection. *Am J Obst Gyn* 1991; 165: 392-400.
169. Vernon SD, Hart CE, Reeves WC, Icenogle JP: The HIV-1 tat protein enhances E2-dependent human papillomavirus 16 transcription. *Virus Research* 1993; 27: 133-145.
170. Villa LL, Sichero L, Rahal P, et al: Molecular variants of human papillomavirus types 16 and 18 preferentially associated with cervical neoplasia. *J Gen Virol* 2000; 81 Pt 12: 2959-2968.



171. Walboomers JM, Jacobs MV, Manos MM, et al: Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol* 1999; 189: 12-19.
172. 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. *J Virol* 1996; 70: 4839-4844.
173. Watts KJ, Thompson CH, Cossart YE, Rose BR: Sequence variation and physical state of human papillomavirus type 16 cervical cancer isolates from Australia and New Caledonia. *Int J Cancer* 2002; 97: 868-874.
174. Weinberg RA: The retinoblastoma protein and cell cycle control. *Cell* 1995; 81: 323-330.
175. Wheeler CM, Yamada T, Hildesheim A, Jenison SA: Human papillomavirus type 16 sequence variants- identification by E6 and L1 lineage-specific hybridization. *J Clin Microbiol* 1997; 35: 11-19.
176. Winkelstein WJ: Smoking and cervical cancer--current status: a review. *Am J Epidemiol* 1990; 131: 945-957.
177. Woo MS, Sanchez I, Dynlacht BD: p130 and p107 use a conserved domain to inhibit cellular cyclin-dependent kinase activity. *Mol Cell Biol* 1997; 17: 3566-3579.
178. Xi LF, Koutsky LA, Galloway DA, et al: Genomic variation of human papillomavirus types 16 and risk for high grade cervical intraepithelial neoplasia. *J Natl Cancer Inst* 1997; 89: 796-802.

179. Xin CY, Matsumoto K, Yoshikawa H, et al: Analysis of E6 variants of human papillomavirus type 33, 52 and 58 in Japanese women with cervical intraepithelial neoplasia/cervical cancer in relation to their oncogenic potential. *Cancer Letters* 2001; 170: 19-24.
180. Yamada T, Manos MM, Peto J, et al: Human papillomavirus type 16 sequence variation in cervical cancers - a worldwide perspective. *J Virol* 1997; 71: 2463-2472.
181. Yamada T, Wheeler CM, Halpern AL, Stewart AC, Hildesheim A, Jenison SA: Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, L2, and L1 coding segments. *J Virol* 1995; 69: 7743-7753.
182. Zehbe I, Mytilineos J, Wikstrom I, Henriksen R, Edler L, Tommasino M: Association between human papillomavirus 16 E6 variants and human leukocyte antigen class I polymorphism in cervical cancer of Swedish women. *Hum Immunol* 2003; 64: 538-542.
183. Zehbe I, Tachezy R, Mytilineos J, et al: Human papillomavirus 16 E6 polymorphisms in cervical lesions from different European populations and their correlation with human leukocyte antigen class II haplotypes. *Int J Cancer* 2001; 94: 711-716.
184. Zehbe I, Wilander E, Delius H, Tommasino M: Human papillomavirus 16 E6 variants are more prevalent in invasive cervical carcinoma than the prototype. *Cancer Res* 1998; 58: 829-833.

185. Zerfass K, Levy LM, Cremonesi C, et al: Cell cycle-dependent disruption of E2F-p107 complexes by human papillomavirus type 16 E7. *J Gen Virol* 1995; 76 ( Pt 7): 1815-1820.
186. Zerfass K, Schulze A, Spitkovsky D, Friedman V, Henglein B, Jansen-Durr P: Sequential activation of cyclin E and cyclin A gene expression by human papillomavirus type 16 E7 through sequences necessary for transformation. *J Virol* 1995; 69: 6389-6399.
187. Zhou J, Sun XY, Louis K, Frazer IH: Interaction of human papillomavirus (HPV) type 16 capsid proteins with HPV DNA requires an intact L2 N-terminal sequence. *J Virol* 1994; 68: 619-625.
188. zur Hausen H: Papillomavirus infections--a major cause of human cancers. *Biochimica et Biophysica Acta* 1996; 1288: F55-F78
189. zur Hausen H: Yohei Ito Memorial Lecture: Papillomaviruses in human cancers. [Review] [40 refs]. *Leukemia* 1999; 13: 1-5.
190. zur Hausen H: Papillomaviruses causing cancer: evasion from host-cell control in early events in carcinogenesis. [Review] [123 refs]. *J Natl Cancer Inst* 2000; 92: 690-698.

### **Manuels de référence**

Fields BN, Knipe DM, Howley PM, Chanock RM, Melnick JL, Monath TP, Roizman B, Straus SE. *Virology, Third Edition*. Philadelphie : Lippincott-Raven, 1996.

Myers G, Bernard HU, Baker C, Halpern A, Delius H, Icenogle J. *Human Papillomaviruses*. Los Alamos : Los Alamos Laboratory, 1995.

### **Sites internet**

Centre François Baclesse : Centre de lutte contre le cancer. <http://www.baclesse.fr>

Northwestern University Medical School : Microbiology-Immunology.

<http://bugs.mimnet.northwestern.edu/mimweb/default.asp>

A closer look at papillomavirus variants.

Purdue University. Timothy S. Baker Laboratory. Determining the structure of macromolecules. <http://bilbo.bio.purdue.edu/~baker/>

Santé Canada. 2003. <http://www.hc-sc.gc.ca/>

UTBM : Graduate School of Biomedical Sciences. <http://gsbs.utmb.edu>

### **Autres**

Centers for Disease Control and Prevention. 1997

Papillomavirus Nomenclature Committee 1995 (14<sup>th</sup> International Papillomavirus Conference, Quebec City, Quebec Canada).