

UNIVERSIDADE DE BRASÍLIA  
INSTITUTO DE BIOLOGIA  
DEPARTAMENTO DE BIOLOGIA CELULAR  
PÓS-GRADUAÇÃO EM BIOLOGIA MOLECULAR

**CARACTERIZAÇÃO MOLECULAR DOS HPVs DE  
ALTO RISCO DOS GENÓTIPOS -53, -56 E -66  
INFECTANDO MULHERES NO DISTRITO FEDERAL  
E ENTORNO.**

Patrícia Soares Wyant

Brasília  
2007

UNIVERSIDADE DE BRASÍLIA  
INSTITUTO DE BIOLOGIA  
DEPARTAMENTO DE BIOLOGIA CELULAR  
PÓS-GRADUAÇÃO EM BIOLOGIA MOLECULAR

**DISSERTAÇÃO DE MESTRADO**  
**CARACTERIZAÇÃO MOLECULAR DOS HPVs DE ALTO**  
**RISCO DOS GENÓTIPOS -53, -56 E -66 INFECTANDO**  
**MULHERES NO DISTRITO FEDERAL E ENTORNO.**

Patrícia Soares Wyant

Orientador: Prof. Dr. Marcelo de Macedo Brígido

Co-orientadora: Profa. Dra. Cláudia Renata Fernandes Martins

Colaboradoras:

Daniela Marreco Cerqueira – Doutoranda da Universidade de Brasília

Daniella Sousa Moraes – Mestranda da Universidade de Brasília

Tainá Raiol Alencar – Doutoranda da Universidade de Brasília

Brasília

2007

UNIVERSIDADE DE BRASÍLIA  
INSTITUTO DE BIOLOGIA  
DEPARTAMENTO DE BIOLOGIA CELULAR  
PÓS-GRADUAÇÃO EM BIOLOGIA MOLECULAR

**DISSERTAÇÃO DE MESTRADO**  
**CARACTERIZAÇÃO MOLECULAR DOS HPVS DE ALTO**  
**RISCO DOS GENÓTIPOS -53, -56 E -66 INFECTANDO**  
**MULHERES NO DISTRITO FEDERAL E ENTORNO.**

Patrícia Soares Wyant

Orientador: Prof. Dr. Marcelo de Macedo Brígido

Co-orientadora: Profa. Dra. Cláudia Renata Fernandes Martins

Banca examinadora:

Profa. Dra. Andréa Queiroz Maranhão (Suplente) – Universidade de Brasília

Prof. Dr. José Paulo Gagliardi Leite – Fundação Oswaldo Cruz

Prof. Dr. Marcelo de Macedo Brígido – Universidade de Brasília

Prof. Dr. Renato de Oliveira Resende – Universidade de Brasília

Brasília  
2007

"A ignorância gera confiança com mais freqüência do que o conhecimento: são aqueles que sabem pouco, e não aqueles que sabem muito, que tão positivamente afirmam que esse ou aquele problema jamais será resolvido pela ciência."

*Charles Darwin*

À minha mãe, que me formou como sou e por isso é parte de cada conquista minha.

Ao Claudiner, que incansavelmente me estimula a seguir em frente apesar dos obstáculos.

À Claudia Renata, pessoa e profissional batalhadora com quem tanto aprendi.

Às vítimas do HPV, que anonimamente tornaram possível este trabalho.

## AGRADECIMENTOS

À minha mãe **Fernanda Soares**, meu irmão **Rafael Wyant**, meu pai **Roland Wyant** e à **Kelle Virgínia** pelo carinho, compreensão, apoio e incentivo constantes;

Aos tios **Marilene Ghilardi** e **Antônio Ghilardi**, e primas **Andrea Ghilardi** e **Alessandra Ghilardi** pelo importante incentivo, amizade e respeito;

Ao **Claudiner**, meu amigo e namorado pela força, amor, carinho e dedicação.

À **Profª Cláudia Renata Fernandes Martins**, pela amizade, confiança e por ter me influenciado de forma positiva na minha formação profissional e pessoal;

Ao **Profº Marcelo de Macedo Brígido**, por ter me permitido ser sua orientanda e pelas valiosas recomendações durante a realização deste trabalho;

Ao **Profº José Paulo Gagliardi Leite**, à **Profª Sônia Bão** e à **Profª Marisa Ferreira** pela disponibilidade e apoio essenciais ao desenvolvimento desse trabalho;

Aos amigos **Tainá Alencar**, **Eduardo Ramalho**, **Ana Paula Oliveira**, **Daniela Cerqueira**, **Daniella Sousa**, **Natália Milanezi**, **Nazle Vêras**, **Margareti Medeiros**, **Márcio Rojas**, **David Neves**, **Geni Noceti**, **Érico Dianese**, **Regina Amorim** e **Verônica Veras**, do grupo de virologia humana da UnB, pela amizade e colaboração na execução deste trabalho;

À **Profª Marisa Ferreira** e aos **técnicos Arenildo**, **Arlindo**, **César**, **Francisca**, **Kamila** e **Marivaldo** do laboratório de Fitopatologia, pela importante e freqüente colaboração;

À **Profª Lídia Moraes**, à **Profª Cynthia Kyaw** e à **Profª Andréia Maranhão**, pelas importantes considerações apresentadas durante o exame de qualificação deste projeto;

Ao **Profº José Paulo Gagliardi Leite** e ao **Profº Renato de Oliveira Resende** e à **Profª Andréa Queiroz Maranhão**, membros da banca examinadora, pelas valiosas considerações prestadas durante a defesa dessa dissertação;

Ao **Sócrates**, **Fabício**, **Marcos**, **Luciana** e **Mário**, pela colaboração no seqüenciamento automático realizado no Laboratório de Biologia Molecular da Universidade de Brasília e na plataforma de seqüenciamento no Cenargen – Embrapa;

Aos queridos **amigos de curso (Biologia 01/2001)**, em especial **Marcelo Salviano**, **Fernando Fonseca**, **Antônio Gomes**, pelo constante companheirismo, aprendizado e carinho;

À **Ana**, funcionária da Secretaria do Departamento de Biologia Celular, por me atender sempre com cordialidade e eficiência;

Ao **Programa de Pós-Graduação em Biologia Molecular** da Universidade de Brasília por me permitir a realização desse Mestrado;

À **CAPES**, **FINATEC** e **Programa Nacional de AIDS/DST** pelo apoio financeiro.

## ÍNDICE

Resumo .....	viii
Abstract.....	ix
Índice de figuras .....	x
Índice de tabelas .....	xi
Lista de Siglas e de Abreviaturas .....	xii
1 Introdução .....	13
2 Revisão Bibliográfica .....	15
2.1 O Vírus .....	15
2.2 Classificação dos Papilomavírus Humanos .....	15
2.3 Organização Genômica .....	17
2.4 A Região Reguladora (LCR) .....	18
2.5 Fatores de Transcrição Envolvidos no Ciclo de Infecção do HPV .....	19
2.6 As Proteínas Virais e o Ciclo de Infecção Viral .....	21
2.6.1 A entrada do vírus na célula .....	21
2.6.2 A replicação viral .....	22
2.6.3 A transformação celular .....	23
2.6.4 A montagem viral e liberação das partículas .....	26
2.7 A Doença .....	27
2.7.1 Co-fatores no desenvolvimento do câncer cervical .....	30
2.7.2 Prevenção.....	30
2.7.3 Vacinas contra o câncer cervical.....	31
2.8 Os Genótipos -53, -56 e -66 do HPV .....	32
2.9 Estudos de Variantes .....	33
2.10 Justificativa e Objetivos .....	41
3 Metodologia .....	43
3.1 Amostras Usadas no Estudo .....	43
3.2 Amplificação do DNA Viral.....	44
3.3 Seqüenciamento .....	45
3.4 Análise das Seqüências.....	46
3.5 Predição de Sítios de Ligação para Fatores de Transcrição na LCR .....	46
3.6 Análise Filogenética .....	47
4 RESULTADOS.....	48

4.1	Amplificação do DNA Viral.....	48
4.2	Predição de Sítios de Ligação para Fatores de Transcrição na LCR .....	48
4.3	Análise das Sequências.....	53
4.3.1	HPV-53.....	53
4.3.2	HPV-56.....	54
4.3.3	HPV-66.....	56
4.4	Análise Filogenética .....	57
4.4.1	HPV-53.....	57
4.4.2	HPV-56.....	59
4.4.3	HPV-66.....	61
5	Discussão .....	63
6	Conclusões .....	68
7	Perspectivas .....	69
8	Bibliografia .....	70
9	Anexo I .....	82
	HPV-53 E6.....	82
	HPV-53 LCR.....	83
	HPV-53 L1 .....	84
	HPV-56 E6.....	85
	HPV-56 LCR.....	86
	HPV-56 L1 .....	87
	HPV-66 E6.....	88
	HPV-66 LCR.....	89
	HPV-66 L1 .....	90
10	Anexo II .....	92



## RESUMO

O câncer cervical é a segunda causa de morte na população feminina e está associado à infecção pelos papilomavírus humanos (HPVs). Os diferentes genótipos do HPV têm potencial oncogênico distinto e, da mesma forma, variantes intragenótipo também podem apresentar oncogenicidade diferenciada. Vários trabalhos foram realizados no sentido de caracterizar os variantes dos HPVs-16 e -18, porém, existe pouca informação sobre os variantes dos outros genótipos de alto risco. O objetivo desse trabalho foi ampliar os conhecimentos acerca da variabilidade genética dos HPVs -53, -56 e -66. Para isso foram amplificadas e seqüenciadas as regiões genômicas LCR, E6 e L1 de seis amostras do HPV-53, cinco do HPV-56 e seis do HPV-66. As seqüências geradas foram analisadas quanto à variação de nucleotídeos em relação às seqüências de referência dos HPVs-53, -56 e -66. Além disso, foi feita a análise filogenética das amostras do Distrito Federal (DF) juntamente com amostras de outras partes do mundo a partir de seqüências depositadas no GenBank. A análise das seqüências de E6 e LCR das amostras do HPV-53 revelou a presença de 6 variantes com relação à seqüência de referência. Na análise do gene L1 foram detectados 3 variantes. Dentre as amostras do HPV-56, na análise de E6 e de LCR foram detectados 3 variantes e no gene L1 um variante foi detectado em 3 das 5 amostras. A análise das amostras do HPV-66 revelou a presença de um variante na análise do gene E6, nenhum variante na LCR e 2 variantes na análise do gene L1. Foram observadas substituições nucleotídicas não silenciosas, levando à troca de aminoácidos, o que pode indicar uma diversidade biológica e patogênica entre os variantes de cada genótipo do HPV. Além disso, as variações detectadas em potenciais sítios de ligação para fatores transcricionais, celulares ou viral, podem influenciar a expressão dos oncogenes virais. Na análise filogenética não foi observado agrupamento étnico ou geográfico que pudesse ser correlacionado com os variantes detectados, como observado para HPV-16 e -18. A análise filogenética dos HPVs-53 e -56 revelou uma divisão dicotômica somente encontrada na análise de subtipos. Filogeneticamente, o HPV-66 foi o genótipo que apresentou maior conservação, com amostras de várias regiões se agrupando no mesmo ramo onde se localizou a seqüência de referência. Mais estudos devem ser conduzidos a fim de esclarecer o impacto das variações encontradas neste trabalho, visto que podem interferir na infectividade e patogenicidade viral.

## ABSTRACT

Cervical cancer is the second major cause of death in the female population and is associated with human papillomavirus (HPV) infection. Different HPV genotypes have different oncogenic potential, and different variants intragenotype can also present different oncogenic potential. Several works have been done with the aim of describing HPV-16 and -18 variants. However, there is little information about variants of the other high-risk genotypes. The aim of this study was to improve the knowledge about genetic variability of HPVs-53, -56, and -66. We amplified and sequenced the LCR, E6 and L1 genomic regions of six HPV-53, five HPV-56 and six HPV-66 isolates. The sequences were compared to the reference sequences of HPVs-53, -56 and -66 to evaluate the presence of nucleotide substitutions. We also performed a phylogenetic analysis with the Federal District (DF) isolates and those of other parts of the world, using sequences from the GenBank. The analysis of the HPV-53 E6 and LCR isolates showed the presence of six variants in comparison with the reference sequence. In the analysis of the L1 gene we detected 3 variants. The analysis of HPV-56 E6 and LCR isolates showed the presence of 3 variants. On the L1 gene, 1 variant was detected. The analysis of HPV-66 isolates revealed the presence of one variant in E6 and 2 variants in L1, no variants were found in LCR. There were observed non-synonymous nucleotide substitutions, leading to amino acid exchanges, which may indicate a biological and pathogenic diversity between the HPVs genotypes. Furthermore, the substitutions detected on the putative transcription factors binding sites may have some implication on viral oncogenes expression. The phylogenetic analysis of HPVs-53 and -56 did not show ethno-geographical clustering that could be correlated with the detected variants, as observed for HPV-16 and -18. The HPVs-53, -56 and -66 phylogenetic analysis showed a dichotomic division only described to subtypes analysis. Phylogenetically, the HPV-66 was the most conserved genotype, with isolates from many regions clustering in the same branch, with the reference sequence. More studies must be conducted to clarify the impact of the variations found in this study, once they can interfere with the viral infectivity and pathogenicity.

## ÍNDICE DE FIGURAS

Figura 1. Localização do colo uterino e representação do câncer cervical .....	13
Figura 2. Micrografia eletrônica de papilomavírus .....	15
Figura 3. Árvore filogenética dos papilomavírus.....	17
Figura 4. Representação esquemática do genoma do HPV .....	18
Figura 5. Representação esquemática da LCR do HPV-16.....	19
Figura 6. Associação de E7 à proteína do retinoblastoma.....	24
Figura 7. Degradação de p53 mediada por E6.....	25
Figura 8. Ciclo de infecção pelo HPV.....	27
Figura 9. História natural da infecção pelo HPV .....	28
Figura 10. Progressão das lesões com o tempo.....	29
Figura 11. Árvore filogenética do HPV-16 .....	34
Figura 12. Diversidade intragenótipo do HPV-53 .....	36
Figura 13. Diversidade intragenótipo do HPV-56 .....	37
Figura 14. Diversidade intragenótipo do HPV-66 .....	38
Figura 15. Árvore filogenética do HPV-53 .....	40
Figura 16. Análise da qualidade dos fragmentos de DNA amplificados por PCR .....	48
Figura 17. Mapeamento dos potenciais sítios de ligação na LCR do HPV-53.....	49
Figura 18. Mapeamento dos potenciais sítios de ligação na LCR do HPV-56.....	50
Figura 19. Mapeamento dos potenciais sítios de ligação na LCR do HPV-66.....	51
Figura 20. Representação esquemática da LCR dos HPV-16, -53, -56, -66 e -6.....	52
Figura 21. Análise filogenética do HPV-53.....	58
Figura 22. Análise filogenética do HPV-56.....	60
Figura 23. Análise filogenética do HPV-66.....	62

## ÍNDICE DE TABELAS

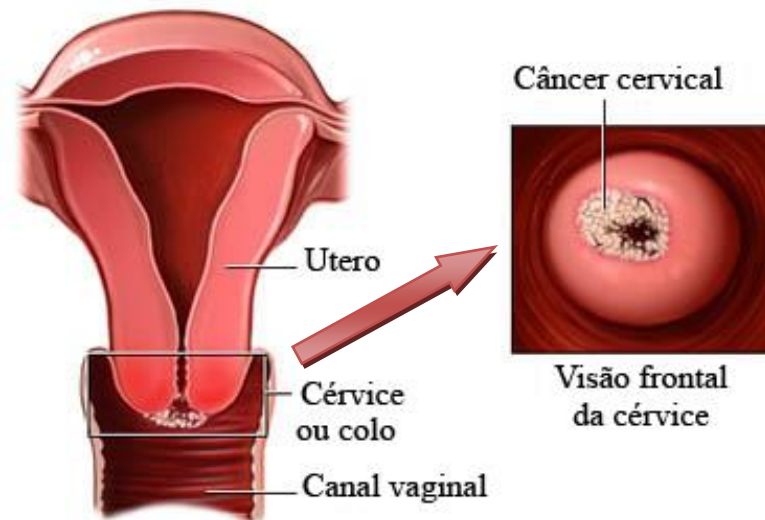
Tabela 1. Dados das amostras caracterizadas neste estudo .....	43
Tabela 2. Iniciadores usados na amplificação de E6, LCR e L1 .....	44
Tabela 3. Variabilidade genética de E6 das amostras do HPV-53.....	53
Tabela 4. Variabilidade genética da LCR das amostras do HPV-53 .....	54
Tabela 5. Variabilidade genética de L1 das amostras do HPV-53.....	54
Tabela 6. Variabilidade genética de E6 das amostras do HPV-56.....	55
Tabela 7. Variabilidade genética da LCR das amostras do HPV-56 .....	55
Tabela 8. Variabilidade genética de L1 das amostras do HPV-56.....	56
Tabela 9. Variabilidade genética de E6 das amostras do HPV-66.....	56
Tabela 10. Variabilidade genética de L1 das amostras do HPV-66.....	57

## LISTA DE SIGLAS E DE ABREVIATURAS

- ANVISA – Agência Nacional de Vigilância Sanitária;
- AP-1 – Proteína ativadora -1 (*Activator Protein-1*);
- BLAST – Ferramenta básica de alinhamento local (*Basic Local Alignment Search Tool*);
- C/EBP – Proteína de ligação à seqüência CCAAT (*CCAAT enhancer binding protein*);
- CDK – quinase dependente de ciclina (*Cyclin Dependent Kinase*);
- cMyc – *Cellular myelocytomatosis oncogene*;
- DF – Distrito Federal;
- E6AP – Proteína de associação a E6 (*E6 Associated Protein*);
- EGR-R – receptor de fator de crescimento epidermal (*epidermal growth factor receptor*);
- GRE – Elemento de resposta a glicocorticóide (*Glucocorticoid Response Element*);
- HPV – Papilomavírus humano (*Human Papillomavirus*);
- HSIL – Lesão escamosa epitelial de alto grau (*High-grade squamous intraepithelial lesion*);
- hTERT – subunidade catalítica da enzima telomerase (*human Telomerase Reverse Transcriptase*);
- INCA – Instituto Nacional de Câncer;
- ITS – Infecção de Transmissão Sexual;
- LCR – Longa região de controle (*Long Control Region*);
- LSIL – Lesão escamosa epitelial de baixo grau (*Low-grade squamous intraepithelial lesion*);
- NF1 – Fator nuclear - 1 (*Nuclear Factor 1*);
- NIC – Neoplasia Intraepitelial Cervical;
- Oct-1 – Fator de ligação a octâmero (*Octamer binding factor -1*);
- ORF – Fase de leitura aberta (*Open Reading Frame*);
- Pap – Papanicolau;
- pRb – Proteína do Retinoblastoma (*Retinoblastoma protein*);
- PRE – Elemento de resposta a progesterona (*Progesterone Response Element*);
- PV – Papilomavírus (*Papillomavirus*);
- RFLP - Análise de polimorfismo dos fragmentos de restrição (*Restriction Fragment Length Polymorphism*);
- Sp1 – Fator promotor seletivo -1 (*Selective promoter factor 1*);
- TEF-1 – Fator potencializador transcricional (*Transcriptional enhancer factor-1*);
- VLP – Partícula semelhante ao vírus (*Virus Like Particle*);
- YY1 – yin-yang factor -1;

## 1 INTRODUÇÃO

O câncer cervical (Figura 1), também denominado câncer de colo do útero, é a segunda causa de morte na população feminina e está associado à infecção pelos papilomavírus humanos (HPVs). O DNA do HPV é detectado em mais de 99% dos casos de câncer de colo do útero (Scheurer *et al*, 2005; Yu *et al*, 2005). A infecção por HPV é considerada um passo intermediário necessário para o desenvolvimento do câncer cervical (Franco e Harper, 2005).



**Figura 1. Localização do colo uterino e representação do câncer cervical.** O colo do útero é o segmento do útero que fica em contato com o canal vaginal. Fonte: Nucleus Medical Art.

A infecção genital por HPV é a doença de transmissão sexual mais comumente diagnosticada em populações jovens sexualmente ativas (Koutsky *et al*, 2002; Scheurer *et al*, 2005). Estima-se que 75% da população em idade reprodutiva esteja infectada com HPV genital e que a maioria das mulheres sexualmente ativas irá adquirir infecção por HPV ao longo da vida (Villa *et al*, 2005; Scheurer *et al*, 2005).

O câncer de colo do útero é o segundo tipo de câncer mais comum entre mulheres no mundo, com cerca de 471.000 novos casos e 230.000 mortes por ano. Oitenta por cento desses casos ocorrem em países em desenvolvimento e, em algumas regiões, esse é o câncer mais comum entre as mulheres (Harper *et al*, 2004; Ministério da Saúde, 2006).

No Brasil, estimou-se em 19.260 o número de novos casos de câncer de colo do útero no ano de 2006. Excluindo-se os tumores de pele não melanoma, o câncer de colo do útero foi o mais incidente na região Norte com 22 casos para cada 100.000 mulheres (22/100.000), o segundo mais incidente nas regiões Sul (28/100.000), Centro-Oeste (21/100.000) e Nordeste (17/100.000), e o terceiro mais incidente na região Sudeste (20/100.000) (Ministério da Saúde, 2006).

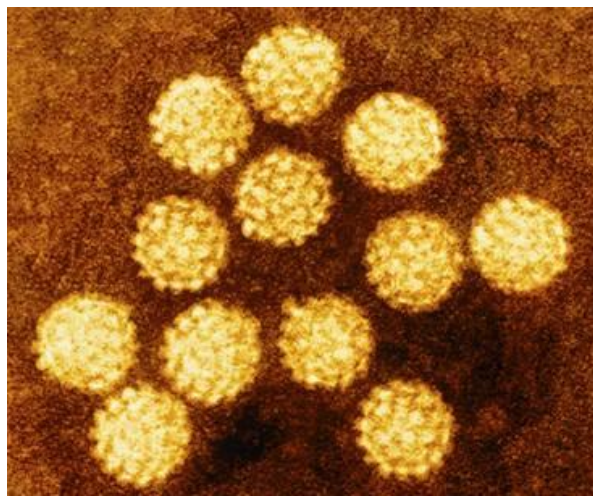
Na região Centro-Oeste, a estimativa foi de 1.430 novos casos para 2006, sendo 220 no Distrito Federal (Ministério da Saúde, 2006). Esses números podem estar subestimados, visto que, tanto o câncer de colo do útero, quanto a infecção por HPV, não são doenças de notificação compulsória (Ministério da Saúde, 2003).

## 2 REVISÃO BIBLIOGRÁFICA

### 2.1 OS PAPILOMAVÍRUS

Os papilomavírus (PVs) são um grupo diverso de vírus encontrados em mais de 20 espécies de mamíferos, bem como em aves e répteis. Devido à importância médica, os papilomavirus humanos (HPVs) têm sido os mais estudados (Doorbar, 2005).

Os HPVs são vírus não envelopados de simetria icosaédrica e medem aproximadamente 55 nm de diâmetro (Fauquet *et al*, 2005) (Figura 2).



**Figura 2.** Micrografia eletrônica de papilomavírus Fonte: Stannard (1995).

Os HPVs pertencem à família *Papillomaviridae* e aos gêneros *Alphapapillomavirus*, HPVs que preferencialmente infectam a mucosa oral ou anogenital em humanos e primatas; *Betapapillomavirus*, *Gammapapillomavirus*, *Nupapillomavirus* e *Mupapillomavirus*, HPVs que infectam preferencialmente a pele de humanos (Fauquet *et al*, 2005).

### 2.2 CLASSIFICAÇÃO DOS PAPILOMAVÍRUS HUMANOS

A classificação dos HPVs baseia-se na divergência genômica de L1, assim, a família *Papillomaviridae* engloba 16 gêneros, que apresentam entre si uma diferença maior que 40% na seqüência nucleotídica do gene L1, enquanto as espécies apresentam diferença entre 30% e 40% em L1. As espécies são denominadas pelo genótipo mais estudado entre os genótipos que as compõem, por exemplo, os HPVs-16, -31, -33, -35, -52, -58, -67 pertencem à espécie *Human Papillomavirus 16*. Além da classificação em família, gênero e espécie os papilomavírus (PVs) são classificados em genótipos, subtipos e variantes. Um genótipo de PV difere de outro em pelo menos 10% na seqüência de nucleotídeos do gene L1. O termo



subtipo é empregado para identificar genomas de PV com seqüências de nucleotídeos de L1 que diferem entre 2% a 10% daquela do genótipo mais próximo. Os variantes de genótipos de PV diferem em menos de 2% na seqüência de nucleotídeos de L1, e em 5% na LCR (Ho *et al*, 1993; de Villiers *et al*, 2004).

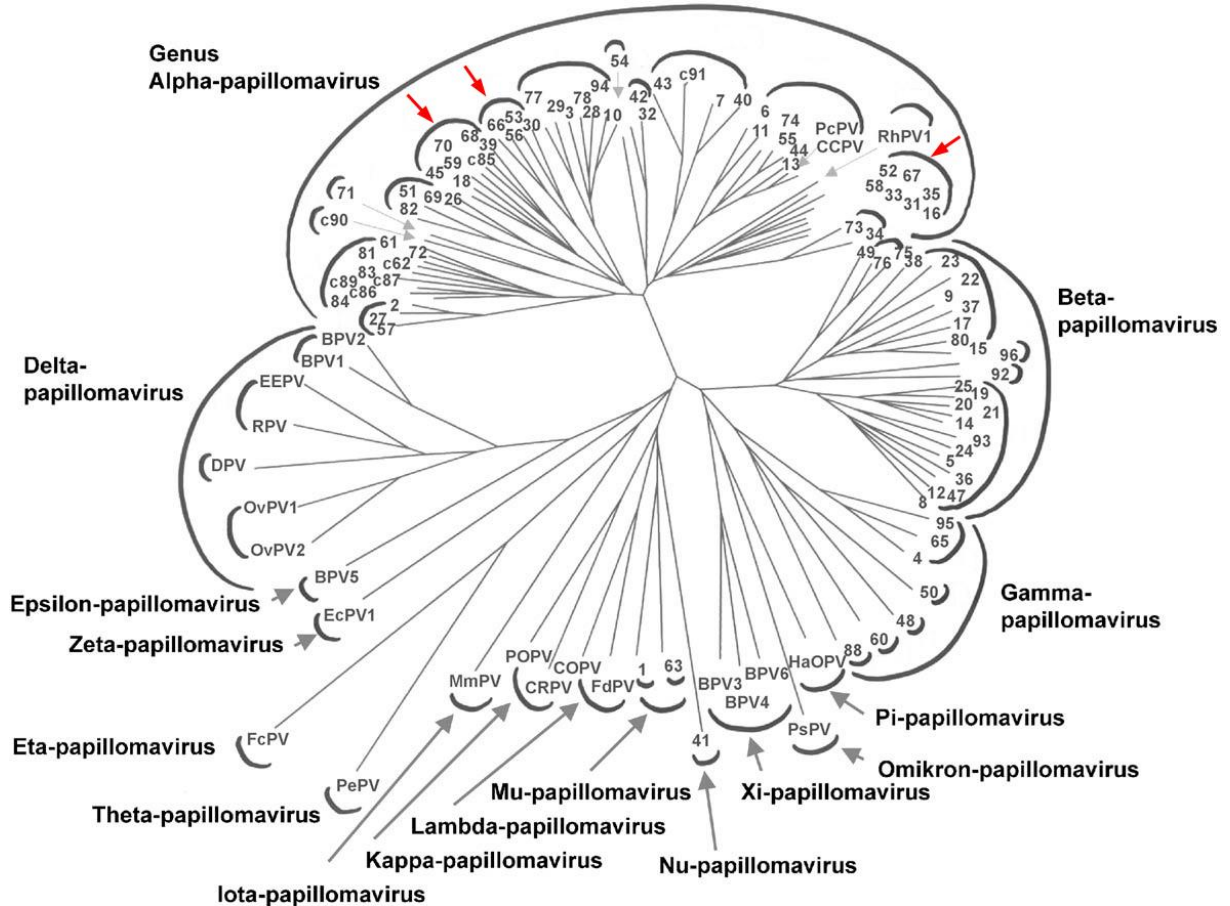
Até o presente já foram descritos mais de 100 genótipos do HPV (de Villiers *et al*, 2004). Os genótipos são nomeados pela sigla HPV seguida de um número que é dado seqüencialmente, à medida que diferentes genótipos são descobertos (Bernard *et al*, 1994).

Os genótipos dos HPVs são classificados em cutâneos e mucosotróficos, dependendo do tropismo pelo tecido. Os cutâneos são aqueles que infectam a pele e os mucosotróficos são os que infectam as mucosas urogenitais, anais e oro-respiratórias (Sterling e Tyring, 2001; Scheurer *et al*, 2005). Os HPVs mucosotróficos podem ainda ser definidos como de baixo ou alto risco, de acordo com o potencial de progressão para neoplasias<sup>1</sup>. Os de alto risco possuem potencial oncogênico elevado, sendo considerados os principais causadores de neoplasias cervicais em todo o mundo (zur Hausen, 1991; Scheurer *et al*, 2005). Infecções por genótipos oncogênicos do HPV representam 50% a 75% de todas as infecções do HPV (Villa *et al*, 2005; Scheurer *et al*, 2005).

Dezoito genótipos são de particular interesse por estarem freqüentemente associados ao desenvolvimento de câncer anogenital e são, portanto, classificados como HPVs de alto risco (Muñoz *et al*, 2003). Esses genótipos do HPV pertencem ao gênero *Alphapapillomavirus* e às espécies *Human Papillomavirus 16*, *Human Papillomavirus 18* e *Human Papillomavirus 53*, separadas em três ramos filogenéticos distintos (Fauquet *et al*, 2005) (Figura 3).

---

<sup>1</sup> Neoplasia: Crescimento celular descontrolado que ocorre na ausência de demanda fisiológica.



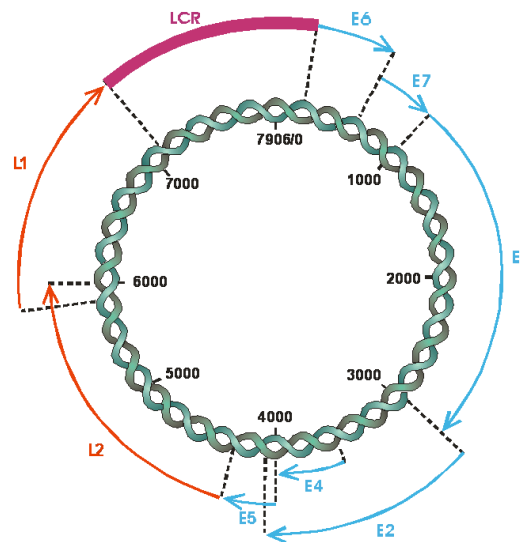
**Figura 3. Árvore filogenética dos papilomavírus com base na análise da sequência do gene L1.** Os números nas pontas dos ramos identificam os genótipos do HPV, c-números indicam candidatos a genótipos do HPV. Todas as outras abreviações se referem a genótipos de papilomavírus animais. Os semicírculos externos identificam os gêneros e os semicírculos internos as espécies. Setas vermelhas indicam as espécies contendo HPVs de alto-risco. Fonte: Modificado de de Villiers *et al* (2004).

### 2.3 ORGANIZAÇÃO GENÔMICA

O genoma dos HPVs é constituído de um DNA de dupla fita circular de aproximadamente 8.000 pares de base associada às histonas celulares H2a, H2b, H3 e H4, que forma um complexo semelhante à cromatina. Possui 8 regiões abertas de leitura (*Open Reading Frames* - ORFs) e uma região não codificadora LCR (*Long Control Region*), com elementos regulatórios e sítios de ligação para fatores de transcrição celulares e viral (Burd, 2003; Scheurer *et al*, 2005).

Apenas uma das duas fitas de DNA é transcrita. O genoma pode ser dividido em três partes: uma região precoce (*E-early*) com cerca de 4kb, a qual codifica proteínas não estruturais E1, E2, E4, E5, E6 e E7; uma região tardia (*L-late*) com aproximadamente 3kb, que codifica as duas proteínas do capsídeo L1 e L2; e uma região de cerca de 1kb não codificadora LCR, que contém elementos de regulação da replicação viral e expressão gênica (Münger *et al*, 2004).

A representação esquemática do genoma do HPV é apresentada na Figura 4.



**Figura 4. Representação esquemática do genoma do HPV.** Observam-se na figura, em azul, a região precoce, responsável pela síntese das proteínas não estruturais E1, E2, E4, E5, E6 e E7, em vermelho, a região tardia, que codifica as proteínas estruturais L1 e L2, e, em lilás, a região reguladora ou LCR (*long control region*). As setas, ao final de cada gene, indicam a direção única em que ocorre a transcrição dos genes do HPV. Na parte interna da figura estão indicadas as posições nucleotídicas no genoma viral a cada 1000pb. Figura modificada de Burk *et al* (1999).

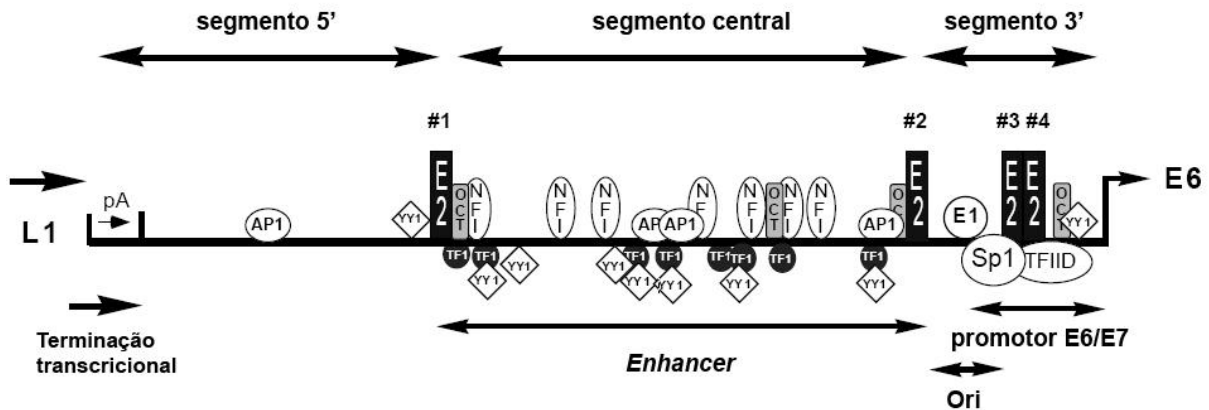
#### 2.4 A REGIÃO REGULADORA (LCR)

A LCR é um segmento não codificador de aproximadamente 850pb que se localiza entre o final do gene L1 e o início do gene E6 (Figura 4). É a região que controla a replicação viral e apresenta vários sítios de ligação para fatores transcricionais celulares e viral (O'Connor *et al*, 1995).

A LCR pode ser dividida em 3 partes: a região 5', o segmento central e a região 3', que está localizada diretamente *upstream* dos genes E6 e E7. No segmento central da LCR está presente o *enhancer* epitélio-específico, que é mais ativo em células epiteliais e é dependente de fatores celulares de transcrição. Sítios para vários fatores celulares incluindo AP-1, NF1, Oct-1, TEF1, TEF2, YY1, e para hormônios esteróides são identificados nessa região (Figura 5) (Kissel'ov, 2000). Uma explicação plausível para a especificidade epitelial da transcrição dos HPVs genitais seria a existência de um ou mais fatores de transcrição, ou co-fatores necessários para a transcrição, produzidos somente em células epiteliais e em nenhum outro tipo celular (Cripe *et al*, 1987; Gloss *et al*, 1987). A região 5' da LCR contém sinais para terminação e poliadenilação de transcritos virais tardios e um sítio de ligação à matriz nuclear (Kissel'ov, 2000).

Na LCR localiza-se o promotor responsável pela transcrição dos genes precoces e tardios, incluindo os oncogenes E6 e E7 (O'Connor *et al*, 1995). A transcrição do RNA começa a

partir do promotor na região 3' da LCR, diretamente *upstream* do gene E6. O promotor de E6/E7 do HPV-16 contém seis elementos conservados: os sítios 3 e 4 de ligação a E2, um sítio de ligação para o fator de transcrição promotor-seletivo 1 (*Selective promoter factor 1*-Sp1) sobreposto ao sítio 3 de E2 e o sítio 4 de ligação a E2 sobreposto ao “TATAbox” (Figura 5). A ocupação alternada desses sítios pelo respectivo fator de reconhecimento promove diferentes níveis de atividade do promotor de E6/E7 (Gloss *et al*, 1990; Tan *et al*, 1992).



**Figura 5. Representação esquemática da LCR do HPV-16.** Quatro sítios de ligação para E2 estão presentes na LCR, sendo que dois deles dividem a LCR em segmentos funcionais distintos, os quais são denominados segmento 5', central e 3'. O segmento 5' contém o sinal de terminação, denotado “pA”, o segmento central é o *enhancer* epitélio específico, o qual contém a maioria dos sítios de ligação para fatores de transcrição, e o segmento 3' contém a origem de replicação e o promotor E6/E7. Fonte: Modificado de O'Connor *et al* (1995).

## 2.5 FATORES DE TRANSCRIÇÃO ENVOLVIDOS NO CICLO DE INFECÇÃO DO HPV

**Proteína ativadora 1 (*activator protein-1* - AP-1):** É um fator de transcrição que se liga ao sítio de ligação como heterodímeros formados de proteínas derivadas das famílias de genes *jun* e *fos* (Meyers *et al*, 1994). Offord *et al* (1990) sugeriram que proteínas que reconhecem sítios de ligação na LCR do HPV para AP-1 estariam presentes apenas em queratinócitos. Dessa forma AP-1 poderia contribuir para a ativação epitélio-específica do *enhancer*, tendo um papel crítico durante a expressão de genes precoces do HPV, em particular na expressão das oncoproteínas E6 e E7 (O'Connor *et al*, 1995).

**Fator nuclear 1 (*Nuclear factor 1* - NF1):** NF1 explica parte da especificidade epitelial do *enhancer* do HPV. Esse gene é expresso a partir de um dentre quatro genes de uma mesma família, dependendo do tipo celular. Em células epiteliais, NF1 é composto principalmente de subunidades derivadas do gene NF1-C, cujos produtos ativam a transcrição. Em células não epiteliais, entretanto, NF1 é composto principalmente de subunidades derivadas do gene NF1-X, cujos produtos não ativam a transcrição (Apt *et al*, 1993).

**Fator de ligação em octâmero -1 (*Octamer binding factor -1, Oct-1*):** Quando presente em células epiteliais em concentrações fisiológicas, Oct-1 se liga a NF-1 estabilizando a ligação deste ao seu sítio de reconhecimento. Esse evento provoca um aumento dos níveis de transcrição do genoma viral (O'Connor e Bernard, 1995).

**Elementos de Resposta à Progesterona (*Progesterone Response Element, PRE*) e a glicocorticóides (*Glucocorticoid Response Element, GRE*):** Sítios de ligação de PRE e GRE estão presentes em *enhancers* de alguns HPVs genitais. PRE e GRE são fatores transcricionais que estimulam a transcrição viral, o que resulta na transformação aumentada de células em cultura (Gloss *et al*, 1987; Chan *et al*, 1989). Essas observações *in vitro* fornecem uma explicação molecular para a evidência epidemiológica de que mulheres que são expostas a altos níveis de progesterona por longos períodos apresentam um maior risco de desenvolverem câncer cervical (Schiffman e Brinton, 1995).

**Fator YY1 (*yin-yang factor -1*):** YY1 é um fator de transcrição multifuncional que ativa ou reprime a transcrição de muitos genes celulares e virais. Essa versatilidade funcional pode ser atribuída aos seus múltiplos domínios transcricionais. Foi demonstrado que YY1 está envolvido na transcrição dos HPVs-16 e -18. Em HPV-16 atua como um repressor transcricional (Galvin e Shi, 1997). A proteína YY1 pode reprimir a transcrição suprimindo a atividade de AP-1 (O'Connor *et al*, 1996) ou competindo com SP1 pela ligação a seus respectivos sítios, que estão sobrepostos na LCR (Dong e Pfister, 1999). No caso do HPV-18, inicialmente, YY1 foi descrito como repressor do promotor. Análises posteriores revelaram que, no contexto da LCR completa do HPV-18, YY1 atua como um ativador ao formar um complexo com C/EBP $\beta$ , e se ligar à região localizada *upstream* do sítio de ligação de YY1 próximo ao promotor. Os resultados sugerem que o complexo C/EBP $\beta$ -YY1 é um regulador positivo do HPV18, tendo um papel crítico na regulação da atividade da LCR (Bauknecht e Shi, 1998).

**C/EBP $\beta$  (*CCAAT enhancer binding protein - beta*):** É um membro de uma família de fatores transcricionais envolvidos na diferenciação celular (C/EBP). Em HPV-18, um aumento nos níveis de C/EBP $\beta$  resulta em forte repressão da atividade da LCR, por desfazer a ligação do complexo Proteína de Ligação ao TATAbox (*TATAbox Binding Protein - TBP*) – TATAbox (Bauknecht e Shi, 1998).

**Sp1 (*Transcription factor specificity protein 1*):** Foi sugerido que a função primária de Sp1 é a de facilitar a ligação dos componentes ao complexo de pré-iniciação da transcrição. Alternativamente, foi proposto que Sp1 pode facilitar a ligação da TBP ao TATAbox ao promotor, desenrolando ou flexionando a fita de DNA (Rose *et al*, 1998).

**cMyc (*cellular myelocytomatosis oncogene*):** Os genes Myc compreendem uma pequena família de genes e estão envolvidos no surgimento de diferentes tipos de tumor em humanos. c-Myc é capaz de desregular a proliferação celular e inibir a diferenciação celular e sua expressão desregulada leva ao crescimento celular, reduz a adesão celular e promove metástase<sup>2</sup> e instabilidade genômica. Muitos cânceres, inclusive o câncer cervical, apresentam níveis elevados de c-Myc. Assim, essa proteína desempenha um papel importante na transformação de células epiteliais (Subramanyam e Krishna, 2006).

## 2.6 AS PROTEÍNAS VIRAIS E O CICLO DE INFECÇÃO VIRAL

O ciclo de infecção dos papilomavírus é dependente da diferenciação celular dos queratinócitos. Estudos sugerem que, para que a lesão seja mantida, o vírus precisa infectar uma célula tronco epitelial. Para que HPV's do gênero *Alphapapillomavirus*, que infectam mucosas, atinjam células da camada basal do epitélio é necessária uma quebra do epitélio estratificado, como microlesões ou abrasão, que ocorrem na pele ou mucosa por contato direto, como resultado, por exemplo, de atividade sexual ou durante o parto (Figura 8) (Schneider, 1994; Doorbar, 2005).

### 2.6.1 A entrada do vírus na célula

Ainda é controversa a natureza do receptor responsável pela entrada do vírus na célula. Evander *et al* (1997) sugeriram que a entrada da partícula do papilomavírus na célula é mediada pela ligação à integrina  $\alpha 6$ , enquanto estudos recentes sugerem que proteoglicanas heparan-sulfato são os receptores celulares responsáveis pelo reconhecimento e ligação inicial dos vírions às células por meio de interação com o C-terminal de L1, a proteína do capsídeo (Joyce *et al*, 1999; Shafti-Keramat *et al*, 2003). Porém, Patterson *et al* (2005) demonstraram que a infecção de queratinócitos por HPV-31b não é dependente de heparan-sulfato, sugerindo que pode haver diferença de receptores celulares na ligação e internalização dos diferentes tipos do HPV. Assim, o processo pelo qual os HPV's entram na célula parece ser diferente entre os genótipos. Sabe-se que, para os HPV's-16 e -58, a internalização ocorre por endocitose mediada por clatrin, enquanto que o HPV-31 parece ser internalizado por meio da membrana caveolar<sup>3</sup> (Bousarghin *et al*, 2003).

---

<sup>2</sup> Migração de células de câncer para outros tecidos ou órgãos à distância do tumor primário. Que ocorre por meio do sistema circulatório (sanguíneo e linfático).

<sup>3</sup> Caveolae: Pequenas invaginações na membrana plasmática onde há acúmulo de colesterol e concentração de proteínas de membrana. Uma das funções dessa estrutura é a internalização de certos vírus e bactérias patogênicas (Anderson, 1998).

Após a internalização, as partículas do HPV se desmontam em endossomos tardios e/ou em lisossomos e ocorre a transferência do DNA viral para o núcleo. Essa transferência parece ser facilitada pela proteína do capsídeo viral L2 (Day *et al*, 2004).

Inicialmente, o genoma permanece nas células da camada basal, como epissomo estável, sem integração ao genoma da célula hospedeira e ocorre síntese das proteínas precoces E1, E2, E4 e E5 (Figura 8) (Doorbar, 2006).

### **2.6.2 A replicação viral**

O padrão de expressão dos genes virais nas células basais não é bem definido, mas acredita-se que as proteínas virais E1 e E2 são mais expressas que as outras proteínas precoces. A proteína precoce E1 tem importante papel na manutenção do DNA viral na forma epissomal (Wilson *et al*, 2002) e, durante a infecção natural, é expressa em níveis muito baixos e requer a presença de E2 para se ligar de forma eficiente ao seu sítio de ligação. A formação de um complexo E1- E2 na origem de replicação induz uma distorção no DNA viral, que facilita o recrutamento de moléculas adicionais de E1 e o eventual desligamento de E2 (Sarafi e McBride, 1995). Além disso, E2 facilita a correta segregação do genoma viral durante a divisão celular, uma vez que o genoma viral se replica junto com o DNA da célula durante a fase S e os genomas replicados são particionados igualmente durante a divisão celular (You *et al*, 2004; McPhillips *et al*, 2005).

Em adição a esse papel na replicação e segregação do genoma, E2 atua como fator de transcrição e pode regular o promotor viral precoce e controlar a expressão dos oncogenes virais E6 e E7. A proteína E2 existe como um homodímero e se liga especificamente à sequência nucleotídica palindrômica ACCG(N<sub>4</sub>)GCCT (McBride *et al*, 1991). Existem 4 sítios de ligação para E2 na LCR do HPV. A capacidade de E2 de reprimir ou ativar a expressão gênica viral e a regulação da expressão de genes precoces de acordo com sua abundância pode ser resultado de diferenças na afinidade de E2 pelos seus sítios de ligação (Hines *et al*, 1998). Em baixos níveis, E2 atua como um ativador transcricional. No HPV-16, supõe-se que os sítios de ligação 1 e 2 são os primeiros a serem ocupados, quando E2 está presente em baixos níveis, levando à ativação do promotor. Com o aumento nos níveis de E2, a ocupação dos seus demais sítios de ligação impede a ligação de fatores de transcrição como Sp1 e TBP, que são necessários para a ativação do promotor (Dong *et al*, 1994).

Embora E1 e E2 sejam primordiais na amplificação do genoma, as proteínas virais E4 e E5 também contribuem nesse processo (Genther *et al*, 2003; Wilson *et al*, 2005). Genomas mutantes em E5 dos HPVs-16 e -31 exibem níveis mais baixos de amplificação do genoma

em relação ao tipo selvagem, e isso pode ser devido à habilidade de E5 em modular a sinalização celular. E5 é uma proteína transmembrana que reside predominantemente no retículo endoplasmático, mas que pode se associar com a ATPase de prótons vacuolar inibindo diretamente a bomba de prótons e, dessa forma, retardar o processo de acidificação endossomal. Suspeita-se que isso afete a reciclagem dos receptores de fator de crescimento epidermal (EGF-R) na superfície celular, inibindo a degradação do EGF-R internalizado, o que aumenta a sinalização mediada por receptores EGF-R e a manutenção de um ambiente favorável para a replicação nas camadas superiores do epitélio, inibindo a diferenciação dos queratinócitos (Straight *et al*, 1993; Disbrow *et al*, 2005).

Em contraste, o papel de E4 na amplificação do genoma ainda não foi completamente estabelecido. E4 se acumula na célula no momento da amplificação do genoma viral e a sua perda tem mostrado prejudicar a replicação em sistemas experimentais (Nakahara *et al*, 2005; Wilson *et al*, 2005).

### **2.6.3 A transformação celular**

Em lesões cervicais causadas por HPV, o aumento da proliferação das células epiteliais suprabasais é atribuído à expressão dos oncogenes virais E6 e E7. Durante a infecção natural, a atividade desses genes estimula a divisão de um pequeno número de células infectadas, aumentando o número de células que subsequenteiramente irão produzir partículas infecciosas. A habilidade de E6 e E7 em levar as células até a fase S é também necessária, juntamente com E1 e E2, para a replicação de epissomos virais acima da camada basal (Doorbar, 2006).

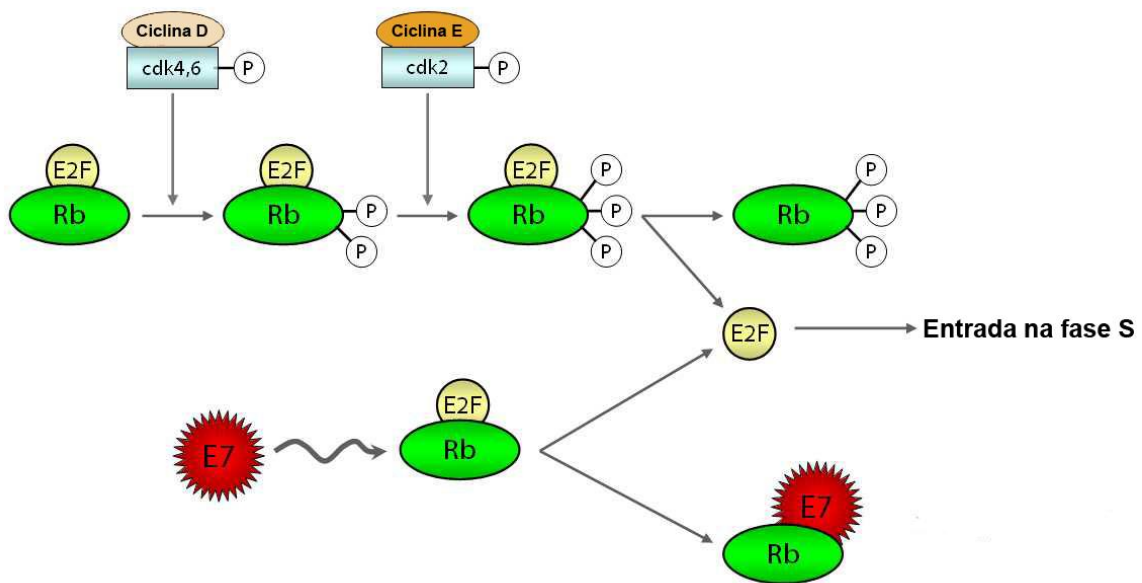
As células suprabasais normalmente saem do ciclo celular e começam o processo de diferenciação terminal, produzindo uma barreira de proteção que é normalmente provida pelas camadas superficiais da mucosa. Em queratinócitos infectados por HPV de alto risco, entretanto, a parada da progressão do ciclo celular é perdida e a diferenciação terminal não ocorre (Sherman *et al*, 1997).

Em lesões de alto grau e cânceres, o DNA do HPV está quase sempre integrado ao genoma do hospedeiro. Isso causa a perda da ORF de E1 e o rompimento da ORF de E2. No entanto, as ORFs de E6 e E7 são preservadas e ativas transcricionalmente em linhagens celulares de câncer cervical e neoplasias associadas ao HPV (Durst *et al*, 1992). Isso resulta no término da amplificação do DNA viral e na ativação constante do promotor precoce do HPV (Stoler *et al*, 1992).

O mecanismo básico pelo qual os papilomavírus estimulam a progressão do ciclo celular é bem conhecido. A proteína do retinoblastoma (pRb) é um regulador negativo do ciclo celular

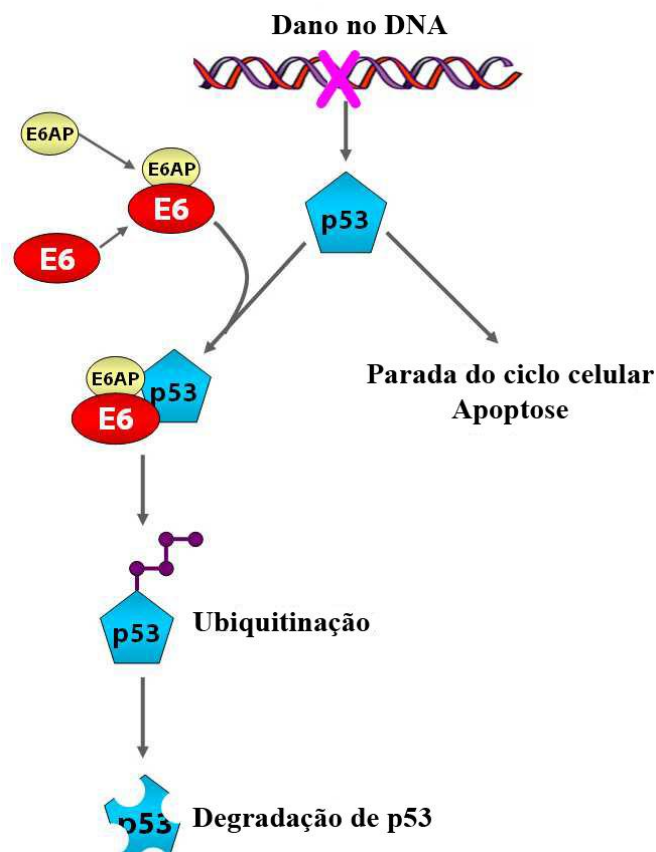


na transição da fase G1 para S e seu estado de fosforilação é regulado durante no ciclo celular. No estado hipofosforilado, pRb se liga ao fator transcricional E2F durante a fase G1. Com a fosforilação de pRb por quinases dependentes de ciclinas (CDK), em resposta a sinais de proliferação celular, o complexo pRb/E2F se dissocia e E2F é liberado para agir como um ativador transcricional de genes envolvidos na progressão do ciclo celular. A associação de E7 à pRb impede a ligação desta a E2F, promovendo então a progressão contínua para a fase S do ciclo celular (Figura 6) (Hebner e Laimins, 2006).



**Figura 6. Associação de E7 a proteína do retinoblastoma.** A fosforilação sequencial da proteína do retinoblastoma (Rb) por complexos ciclina/CDK inibem a atividade repressora de Rb. E7 se liga a Rb em sua forma hipofosforilada. Essa ligação desfaz o complexo entre Rb e o fator de transcrição celular E2F, resultando na liberação deste fator, o que permite que a célula entre na fase S do ciclo celular. Fonte: Modificado de Jo e Kim (2005).

A função da proteína viral E6 complementa a de E7 e, nos HPV de alto risco, as duas proteínas são expressas simultaneamente a partir de um único pré-mRNA bicistrônico (Tang *et al*, 2006). A reentrada no ciclo celular mediada por E7 nas células das camadas superiores (que deveriam sair do ciclo celular e entrar no processo de diferenciação) deveria levar essas células ao processo de apoptose. Esse processo é impedido pela proteína E6 dos HPV de alto risco, que se liga à proteína supressora de tumor p53 por meio da proteína de associação a E6 (E6AP), um polipeptídeo celular que tem atividade de ubiquitina ligase, formando um complexo trimérico. Esse complexo atua marcando especificamente p53 para a degradação, por meio da ubiquitinação (Huibregtse *et al*, 1991) (Figura 7).



**Figura 7. Degradação de p53 mediada por E6.** Danos no DNA induzem a ativação de p53 levando tanto à parada do ciclo celular quanto à apoptose. E6 se liga a E6-AP e o complexo formado se liga a p53. E6-AP ubiquitina p53, que é rapidamente degradada pelo proteassomo. Fonte: Modificado de Jo e Kim (2005).

Embora as proteínas E6 de todos os HPVs genitais tenham provavelmente alguma propriedade de ligação a p53, somente as dos tipos de alto risco se ligam com alta afinidade e provocam a degradação da mesma (Lechner e Laimins, 1994). O papel anti-apoptótico de E6 é um fator chave no desenvolvimento de cânceres cervicais, pois compromete a efetividade de reparo celular do DNA e permite o acúmulo de mutações secundárias (Doorbar, 2006).

Outra importante função da proteína E6 dos HPVs de alto risco é a ativação da telomerase em células infectadas. Em células normais, vários ciclos de replicação do DNA resultam no encurtamento dos telômeros, eventualmente produzindo instabilidade cromossomal e senescência celular. Foi demonstrado que as proteínas E6 dos HPVs de alto risco ativam a subunidade catalítica da enzima telomerase (*human Telomerase Reverse Transcriptase - hTERT*) aumentando o tamanho dos telômeros em células epiteliais. Além de estender a vida de células epiteliais para produção de progênie viral, esse efeito também influencia a carcinogênese induzida por HPV, pois a super-expressão de hTERT somada à expressão de E7 pode ser suficiente para imortalizar queratinócitos humanos primários (Hebner e Laimins, 2006).

#### 2.6.4 A montagem viral e liberação das partículas

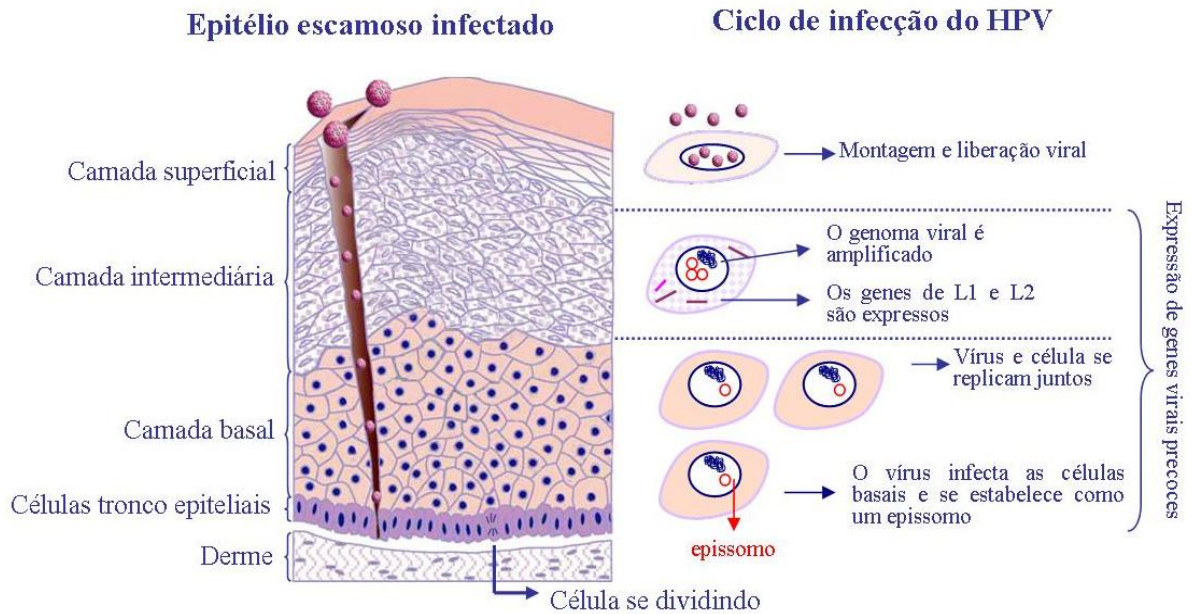
O estágio final no ciclo replicativo dos HPVs requer que os genomas replicados sejam empacotados em partículas infecciosas. Já nas camadas superiores do epitélio, as proteínas do capsídeo (L1 e L2) se acumulam (Figura 8), com a expressão de L2 precedendo a expressão de L1 (Florin *et al*, 2002). Após a síntese no citoplasma, L1 e L2 são direcionadas para o núcleo, onde ocorre a montagem das novas partículas virais (Turek e Smith, 1996).

A montagem dos vírions infecciosos nas camadas superficiais do epitélio pode requerer E2 em adição às proteínas do capsídeo L1 e L2. Zaho *et al* (2000) sugeriram que E2 pode aumentar a eficiência da encapsidação do genoma durante a infecção natural.

Embora as partículas de HPV possam se montar na ausência de L2, sua presença contribui para o empacotamento eficiente e aumenta a infectividade do vírus. Para o HPV-31, foi demonstrado que a perda de L2 resulta em uma redução de 10 vezes na eficiência do empacotamento e uma redução de 100 vezes na infectividade viral, quando comparado ao HPV-31 selvagem (Holmgren, 2005).

A proteína L1 é o principal elemento estrutural do capsídeo viral, o qual é formado por 360 cópias de L1 organizadas em 72 capsômeros (cada capsômero é formado por 5 moléculas de L1) arranjados na superfície do icosaedro. Estudos mostram que existem apenas 12 cópias da proteína L2 por vírion (Trus *et al*, 1997; Modis *et al*, 2002).

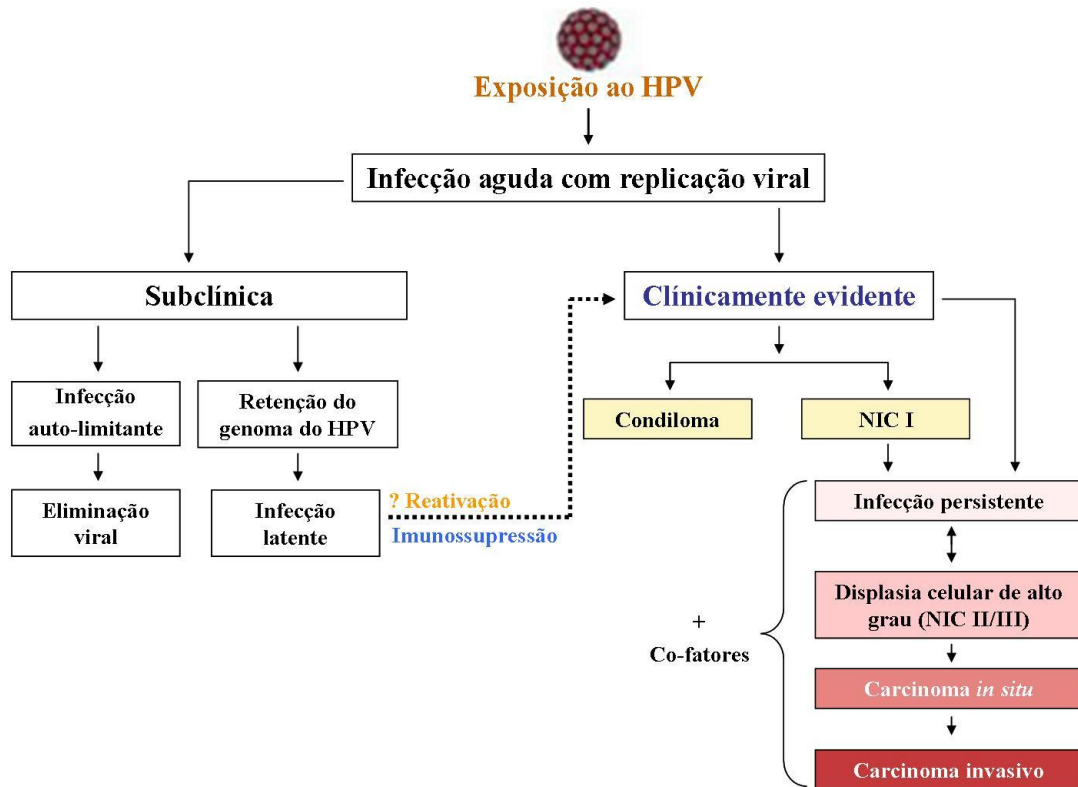
A liberação das partículas virais requer um escape eficiente da célula, processo que parece ser facilitado pela proteína E4, a qual rompe a rede de queratina, afetando a integridade celular (Bryan e Brown, 2000; Wang *et al*, 2004). Os vírions são liberados durante a descamação celular, na superfície da lesão (Turek e Smith, 1996).



**Figura 8. Ciclo de infecção pelo HPV.** Micro-traumas no epitélio expõem as células basais e facilitam a infecção pelo HPV. Após a entrada na célula, o vírus é desnudado e o genoma viral alocado no núcleo celular. O DNA viral permanece como um epissomo e se divide junto, mas de forma independente, do genoma da célula hospedeira. Após a divisão, as células migram da camada basal passando pela intermediária e atingindo a superficial. Porém, durante esse processo as células infectadas continuam se multiplicando, o que resulta na amplificação do genoma viral em milhares de cópias por célula. Na camada intermediária do epitélio, L1 e L2 são expressas. A montagem dos vírions e o empacotamento do DNA celular ocorrem na camada superficial e as partículas infecciosas são liberadas com a esfoliação das células. Fonte: Adaptado de Doorbar (2005) e Merk (2006).

## 2.7 A DOENÇA

Muitas mulheres adquirem infecções por HPV na adolescência por transmissão sexual. Em aproximadamente 80% dessas mulheres a infecção é transitória, sem anormalidades epiteliais. Nas 20% restantes, a infecção por HPV leva ao desenvolvimento de lesões e infecção persistente, um pré-requisito para transformação maligna (Steenbergen *et al*, 2005) (Figura 9).



**Figura 9. História natural da infecção pelo HPV.** A maior parte das infecções por HPV é subclínica (sem sintomas). A resposta imune pode erradicar a infecção ou suprimir a níveis não detectáveis. Algumas infecções são suprimidas, mas o genoma do HPV é mantido em estado latente (infecção subclínica com um pequeno número de células mantendo poucas cópias de DNA). Ainda não está claro como ocorre a latência em hospedeiros imunocompetentes e os mecanismos que levam à reativação do HPV a um estado detectável. Fonte: Modificado de Merk (2006).

O câncer cervical progride a partir de lesões pré-malignas não invasivas, denominadas neoplasias intraepiteliais cervicais (NICs) ou lesões intraepiteliais escamosas<sup>4</sup> (squamous intraepithelial lesions – SILs). Essas lesões pré-malignas são classificadas histologicamente com base nas atipias<sup>5</sup> das células epiteliais: NIC I corresponde a uma displasia<sup>6</sup> branda; NIC II à displasia moderada e NIC III à displasia severa e carcinoma<sup>7</sup> *in situ*. NIC I é também classificada como lesão intraepitelial escamosa de baixo grau (*low-grade* SIL – LSIL) e NIC II/III como lesão intraepitelial escamosa de alto grau (*high-grade* SIL– HSIL) (Steenbergen *et al*, 2005) . A infecção, tanto por HPVs oncogênicos como não oncogênicos, pode causar LSIL na cérvix uterina, enquanto que a maioria das lesões cervicais classificadas como HSIL,

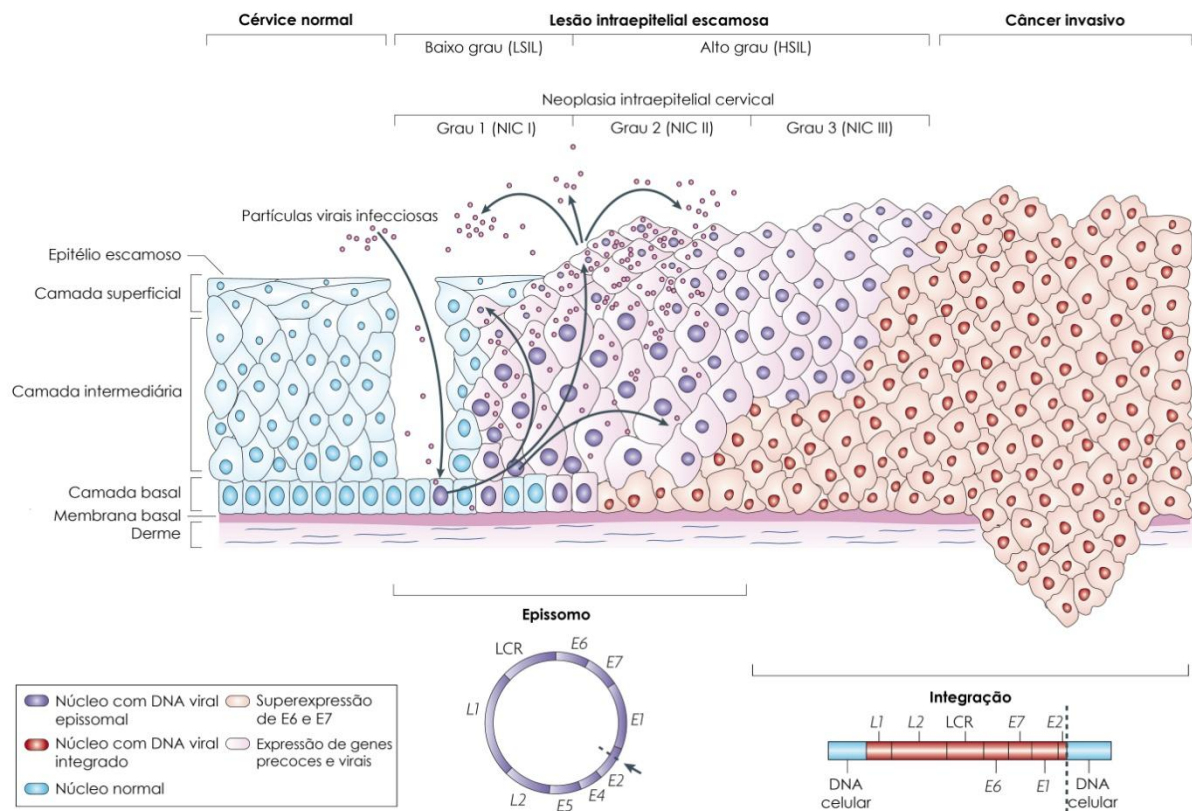
<sup>4</sup> Epitélio escamoso: Epitélio composto por várias camadas celulares.

<sup>5</sup> Atipia: Desvio da morfologia normal.

<sup>6</sup> Displasia: Crescimento celular anômalo que resulta em células que diferem de tamanho, forma ou arranjo em relação às outras células do mesmo tipo de tecido.

<sup>7</sup> Carcinoma: Tumor maligno formado de células epiteliais que infiltram as estruturas vizinhas podendo ocasionar metástases; Sinônimo de câncer.

carcinoma *in situ* ou câncer invasivo é positiva para genótipos oncogênicos do HPV (Baseman e Koutsky, 2005) (Figura 10).



**Figura 10. Progressão das lesões com o tempo.** As células da camada basal do epitélio cervical se localizam sobre a membrana basal, que é sustentada pela derme. Os papilomavírus humanos (HPVs) acessam as células basais por meio de micro abrasões no epitélio cervical. Após a infecção, os genes precoces E1, E2, E4, E5, E6 e E7 são expressos e o DNA viral se replica como DNA epissomal (núcleo roxo). Nas camadas superiores do epitélio (camadas intermediária e superficial) o genoma viral é replicado em maior escala, e os genes E4, L1 e L2 são expressos. L1 e L2 promovem a encapsidação dos genomas virais no núcleo. Os vírions podem então iniciar uma nova infecção. Nas lesões epiteliais de baixo grau há a replicação viral produtiva. Uma parte das infecções por HPV de alto risco progredem para a neoplasia intraepitelial de alto grau (NIC III). A progressão de lesões não tratadas ao câncer invasivo está associada com a integração do genoma do HPV ao cromossomo do hospedeiro (núcleo vermelho), com a perda ou rompimento de E2, e subsequente superexpressão dos oncogenes E6 e E7. Fonte: Woodman *et al* (2007).

Uma teoria comum até recentemente era que mulheres que desenvolviam câncer cervical sempre progrediam por estágios distintos de lesões epiteliais, a partir de baixo grau para moderado e, posteriormente, para de alto grau. Estudos de história natural, entretanto, têm questionado a noção de estágios cervicais pré-cancerosos progressivos contínuos e levado os pesquisadores a concluir que lesões cervicais de baixo e alto-grau são processos distintos de infecção por HPV. LSIL parece ser uma manifestação transitória da infecção viral produtiva, onde o epitélio infectado por HPV passa pela diferenciação e maturação e exhibe apenas pequenas anormalidades celulares. HSIL, o verdadeiro precursor do câncer cervical, ocorre quando a infecção pelo HPV em células imaturas em replicação impede a maturação e diferenciação epitelial, levando à replicação celular continuada e ao acúmulo de



anormalidades genéticas, que podem levar ao surgimento de células cancerosas. LSIL pode se estabelecer antes, ao mesmo tempo, ou na ausência de HSIL (Baseman e Koutsky, 2005).

### 2.7.1 Co-fatores no desenvolvimento do câncer cervical

Embora muitas mulheres sejam infectadas por HPV, a maioria não progride para o câncer cervical. Portanto, outros fatores provavelmente estão envolvidos no desenvolvimento da doença. Três principais grupos de co-fatores são (Muñoz *et al*, 2006):

- Co-fatores ambientais ou exógenos, incluindo contraceptivos hormonais, tabagismo, número de partos e co-infecções com outros agentes sexualmente transmissíveis;
- Co-fatores virais, como infecção por genótipos específicos, co-infecção com outros genótipos do HPV ou variantes, carga viral e integração viral;
- Co-fatores do hospedeiro, incluindo hormônios endógenos, fatores genéticos e outros fatores relacionados à resposta imune.

Acredita-se que o mecanismo pelo qual o número elevado de partos aumenta o risco para o desenvolvimento de carcinomas cervicais seja a manutenção da zona de transformação<sup>8</sup> na exocérvice<sup>9</sup> por muitos anos, aumentando a exposição ao HPV. No entanto, fatores hormonais também podem estar envolvidos. Os efeitos do tabagismo no desenvolvimento do carcinoma incluem redução de resposta imune na cérvice e dano genético direto causado pela carcinogênese relacionada ao tabaco. Os contraceptivos hormonais podem agir como co-fatores no desenvolvimento do câncer cervical por meio do aumento da expressão gênica do HPV na cérvice via elementos de resposta a hormônios (estrógeno e/ou progesterona) presentes no genoma viral (Muñoz *et al*, 2006).

### 2.7.2 Prevenção

A estratégia mais utilizada para prevenção do desenvolvimento do câncer de colo do útero é o exame de Papanicolau (Pap). O Pap baseia-se na coloração de células epiteliais obtidas da cérvice, na expectativa de que anormalidades nucleares detectáveis (discariose<sup>10</sup>) sejam representativas de lesões subjacentes definidas histologicamente. Esse exame reduz a incidência do câncer de colo do útero ao permitir a detecção e a excisão das neoplasias

---

<sup>8</sup> Zona de transformação: Área localizada entre o epitélio escamoso e o epitélio colunar.

<sup>9</sup> Exocérvice: Porção do colo uterino que se projeta para a vagina, onde situa-se o acesso à cavidade uterina.

<sup>10</sup> Discariose: Perda da relação do núcleo com o citoplasma.

intraepiteliais cervicais – NICs ou do adenocarcinoma<sup>11</sup> *in situ* – AIS antes do desenvolvimento do câncer invasivo. Apesar de o diagnóstico ser subjetivo, o Pap é considerado o procedimento de maior sucesso no controle do câncer cervical (Cuschieri e Cubie, 2005).

### 2.7.3 Vacinas contra o câncer cervical

Nos últimos anos as atenções se voltaram para a possibilidade de prevenir lesões pré-cancerosas e o câncer cervical por meio de vacinação contra o HPV (Harper *et al*, 2004; Mandic e Vujkov, 2004; Reddy *et al*, 2004; Franco e Harper, 2005; Villa *et al*, 2005; Lowy e Schiller, 2006). Zhou *et al* (1991) sintetizaram partículas semelhantes a vírus (*virus like particles* – VLPs) expressando apenas dois genes do HPV-16 (L1 e L2) em células eucarióticas (células de levedura). Outros cientistas refinaram a montagem das VLPs e demonstraram que apenas o gene L1 era necessário para a produção das VLPs do HPV que, quando injetadas em voluntários humanos, poderiam estimular altos títulos de anticorpos neutralizantes (Brown *et al*, 2001; Evans *et al*, 2001; Harro *et al*, 2001).

Mais de 35 genótipos do HPV infectam o trato genital. Destes, os genótipos do HPV mais freqüentemente associados ao desenvolvimento do câncer cervical são os HPVs-16 e -18, que juntos são responsáveis por 70% dos casos de câncer de colo do útero e neoplasias intraepiteliais de alto grau. Os HPVs-6 e -11 causam 90% das verrugas anogenitais. Assim, vacinas profiláticas que têm esses genótipos como alvos foram desenvolvidas (Harper *et al*, 2004; Villa *et al*, 2005).

Duas vacinas estão em desenvolvimento, uma fabricada pelo laboratório Merck Sharp e Dhome (internacionalmente conhecida como Gardasil®), e a outra a desenvolvida pela empresa Glaxo SmithKline (internacionalmente conhecida como Cervarix®). A Gardasil® é uma vacina profilática quadrivalente que apresenta VLPs dos HPVs-16, -18, -6 e -11 e a Cervarix® é uma vacina profilática bivalente, contendo VLPs dos HPVs-16 e -18. Essas vacinas induzem a produção de anticorpos neutralizantes que previnem a infecção pelos genótipos do HPV representados na vacina. As vacinas mostraram uma eficiência de 100% na proteção contra o desenvolvimento de lesões cervicais pré-cancerosas e verrugas genitais (Gardasil®) causados pelos genótipos do HPV presentes na vacina, com pouco ou nenhum efeito colateral. A expectativa é de que os efeitos de proteção das vacinas durem pelo menos quatro anos e meio a partir do início da vacinação (Villa *et al*, 2005; Harper *et al*, 2006).

---

<sup>11</sup> Adenocarcinoma: Neoplasia maligna em células epiteliais glandulares (epitélio formado por uma única camada de células colunares produtoras de muco).



A vacina Gardasil® já está disponível no mercado americano e, aprovada pela Agência Nacional de Vigilância Sanitária (ANVISA), começou a ser comercializada no Brasil no início de 2007. Ela deve ser administrada em 3 doses em um período de seis meses, sendo as segunda e terceira doses com dois e seis meses de intervalo em relação à primeira dose, respectivamente. A recomendação do Centro de Controle e Prevenção de Doenças (*Centers for Disease Control and Prevention – CDC*) e da ANVISA para uma melhor eficácia é de que a vacina seja aplicada em mulheres jovens, de 9 a 26 anos, que nunca tiveram contato com HPV. Porém, como dificilmente uma mulher em idade sexualmente ativa que já entrou em contato com o HPV está infectada com todos os tipos presentes na vacina, a mesma pode ser válida mesmo em idades superiores a 26 anos. A vacina Cervarix® está em fase de análise pelas autoridades regulatórias mundiais, ainda não tendo sido aprovada em nenhum país.

Em adição às vacinas preventivas, estão sendo realizadas pesquisas para o desenvolvimento de vacinas terapêuticas contra HPV. Em geral, essas vacinas têm como alvo as principais oncoproteínas E6 e E7, que são importantes na indução e manutenção da transformação celular, sendo expressas nas células dos carcinomas cervicais. Dessa forma, espera-se que a resposta imune contra essas duas oncoproteínas possa erradicar os tumores já estabelecidos (Roden *et al*, 2004).

## **2.8 OS GENÓTIPOS -53, -56 E -66 DO PAPILOMAVÍRUS HUMANO**

Os HPVs dos genótipos -53, -56 e -66 pertencem à espécie *Human Papillomavirus 53*, gênero *Alphapapillomavirus* (Fauquet *et al*, 2005).

O HPV-53 foi isolado pela primeira vez a partir do esfregaço cervical de uma mulher grávida sem anormalidade clínica ou citológica (Gallahan *et al*, 1989). Subseqüentemente, este genótipo tem sido detectado em esfregaços com citologia normal e em displasias, sendo considerado um genótipo de HPV de alto risco. A prevalência do HPV-53 observada em estudos epidemiológicos no mundo varia de 0,5 % a 25% (Brown *et al*, 1999; Gravitt *et al*, 2000; Castellsagué *et al*, 2001; Brown *et al*, 2002; Chan *et al*, 2002; Clifford *et al*, 2005; Stevens *et al*, 2006) e no Brasil de 3,0% a 24,4% (Franco *et al*, 1999, Gonçalves *et al*, 1999; Levi *et al*, 2002; Cerqueira *et al*, 2007). No Distrito Federal a prevalência foi de 6,2% em mulheres HIV-1 soropositivas (Cerqueira *et al*, 2007) e de 6,0% em mulheres HIV-1 soronegativas (Câmara *et al*, 2003). Em um estudo conduzido na Austrália, o HPV-53 foi, juntamente com o HPV-16, o genótipo mais comum entre as mulheres estudadas (Brestovac *et al*, 2005).

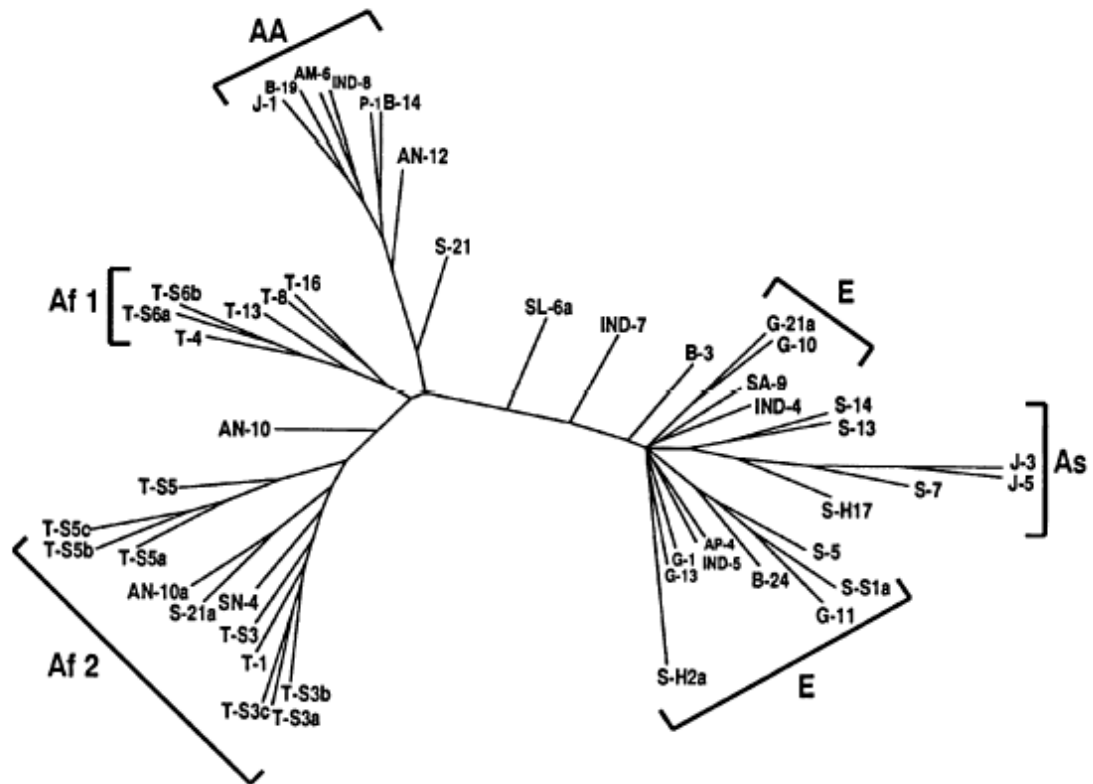
O HPV-56 foi detectado pela primeira vez em uma amostra de neoplasia intraepitelial cervical (NIC) (Lörincz *et al*, 1989) e tem sido encontrado em cânceres cervicais e lesões precursoras de todos os graus. É considerado um genótipo de alto risco e a prevalência observada nos estudos varia entre 0,1% e 18% (Gravitt *et al*, 2000; Castellsagué *et al*, 2001; Brown *et al*, 2002; Chan *et al*, 2002; Luque *et al*, 2004; Clifford *et al*, 2005; Lee *et al*, 2005; Stevens *et al*, 2006). No Brasil a prevalência varia de 1% a 11% (Franco *et al*, 1999; Levi *et al*, 2002; Cerqueira *et al*, 2007). No Distrito Federal, a prevalência foi de 6,2% em mulheres HIV-1 soropositivas (Cerqueira *et al*, 2007) e não há dados sobre a prevalência do HPV-56 em mulheres HIV-1 soronegativas. Em estudo realizado nos Estados Unidos com mulheres HIV-1 soropositivas, o HPV-56 foi o mais prevalente entre os genótipos estudados tendo sido detectado em 18% das amostras (Luque *et al*, 2004).

O HPV-66 foi detectado pela primeira vez em uma biópsia de carcinoma invasivo e é também considerado um HPV de alto risco, tendo sido encontrado principalmente em lesões de alto grau, como as neoplasias intraepiteliais cervicais II e III (NIC II/III). O HPV-66 apresenta prevalência de 0,2% a 15% (Gravitt *et al*, 2000; Richardson *et al*, 2000; Castellsagué *et al*, 2001; Brown *et al*, 2002; Chan *et al*, 2002; Clifford *et al*, 2005; Lee *et al*, 2005). No Brasil, a prevalência variou de 1% a 11% (Franco *et al*, 1999; Levi *et al*, 2002; Cerqueira *et al*, 2007). No Distrito Federal, a prevalência foi de 6,2% em mulheres HIV-1 soropositivas (Cerqueira *et al*, 2007) e de 3,0% em mulheres HIV-1 soronegativas (Câmara *et al*, 2003).

A grande variação observada na prevalência desses vírus pode ser explicada pelas diferenças nos métodos de diagnóstico, na população estudada (pacientes com ou sem lesão preditiva de infecção por HPV; HIV-1 soropositiva ou soronegativa) e no tamanho amostral.

## **2.9 ESTUDOS DE VARIANTES**

Estudos com variantes do HPV-16 e HPV-18 têm mostrado uma co-evolução com os três maiores ramos filogenéticos humanos (africanos, caucasianos e asiáticos) (Ho *et al*, 1993; Ong *et al*, 1993). A Figura 11 mostra a árvore filogenética do HPV-16, destacando os cinco clados nos quais os variantes foram agrupados: Europeu (E), Asiático (As), Asiático-Americano (AA), Africano 1 (Af-1) e Africano 2 (Af-2) (Ho *et al*, 1993).



**Figura 11. Árvore filogenética do HPV-16.** Os variantes foram encontrados a partir da análise da LCR de 301 amostras provenientes de 25 grupos étnicos distintos. Fonte: Ho *et al* (1993).

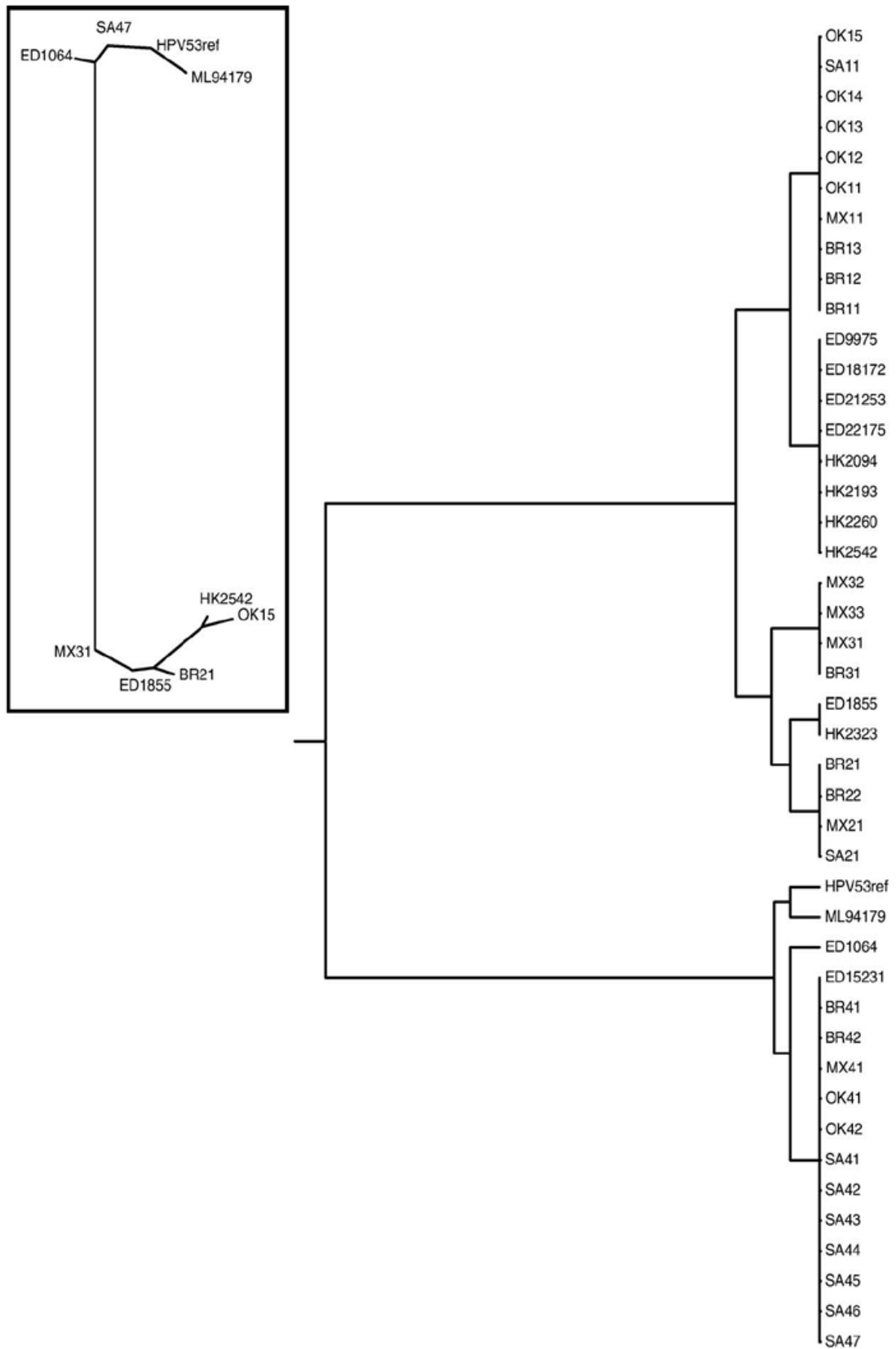
Há indicações de que variantes do mesmo genótipo do HPV diferem biologicamente e etiológicamente. Os diferentes genótipos do HPV têm potencial oncogênico distinto e, da mesma forma, os variantes intragenótipo também podem apresentar oncogenicidade diferenciada. No caso do HPV-16, vários estudos sugerem a associação entre variantes não-Europeus e o risco de incidência e prevalência de lesões de alto grau. Os variantes Europeus estariam associados a um menor risco de desenvolvimento dessas lesões (Xi *et al*, 1997; Veress *et al*, 1999; Kämmer, 2000; Villa *et al*, 2000; Berumen *et al*, 2001). Variantes não-Europeus, tanto do HPV-16 quanto do HPV-18, têm sido associados também com um maior risco de infecções persistentes (Villa *et al*, 2000).

Entretanto, a informação sobre os variantes de outros genótipos do HPV de alto risco que não os HPVs-16 e -18 é limitada, sendo necessárias mais pesquisas a fim de ampliar o conhecimento acerca da variabilidade genética, permitindo um melhor entendimento sobre esses variantes (Scheurer *et al*, 2005).

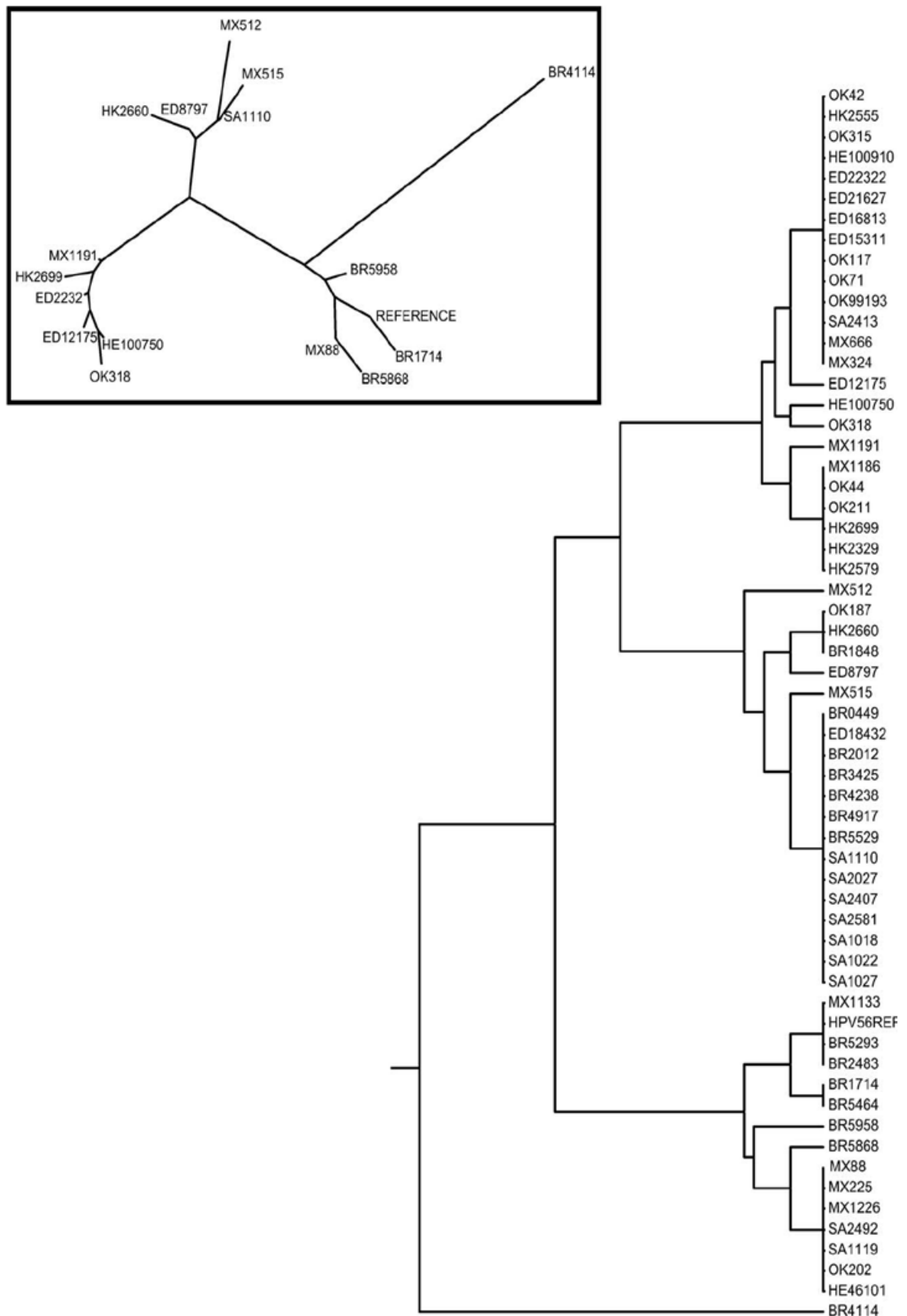
Recentemente, foi realizado um estudo multicêntrico sobre a diversidade genética dos HPVs-53, -56 e -66 por Prado *et al* (2005). O objetivo deste estudo foi investigar se os modelos de diversidade intragenótipo observados para os HPVs-16 e -18 poderiam ser

aplicados para todos os HPVs genitais associados ao câncer. Nesse estudo foram encontradas semelhanças com padrões observados na análise de variantes dos HPVs-16 e -18, como alguns aglomerados geográficos de variantes para HPV-53 e -56, mas também foram encontradas discrepâncias, como a ramificação dicotômica das árvores HPV-53 e -56 e a falta de associação geográfica para os variantes do HPV-66.

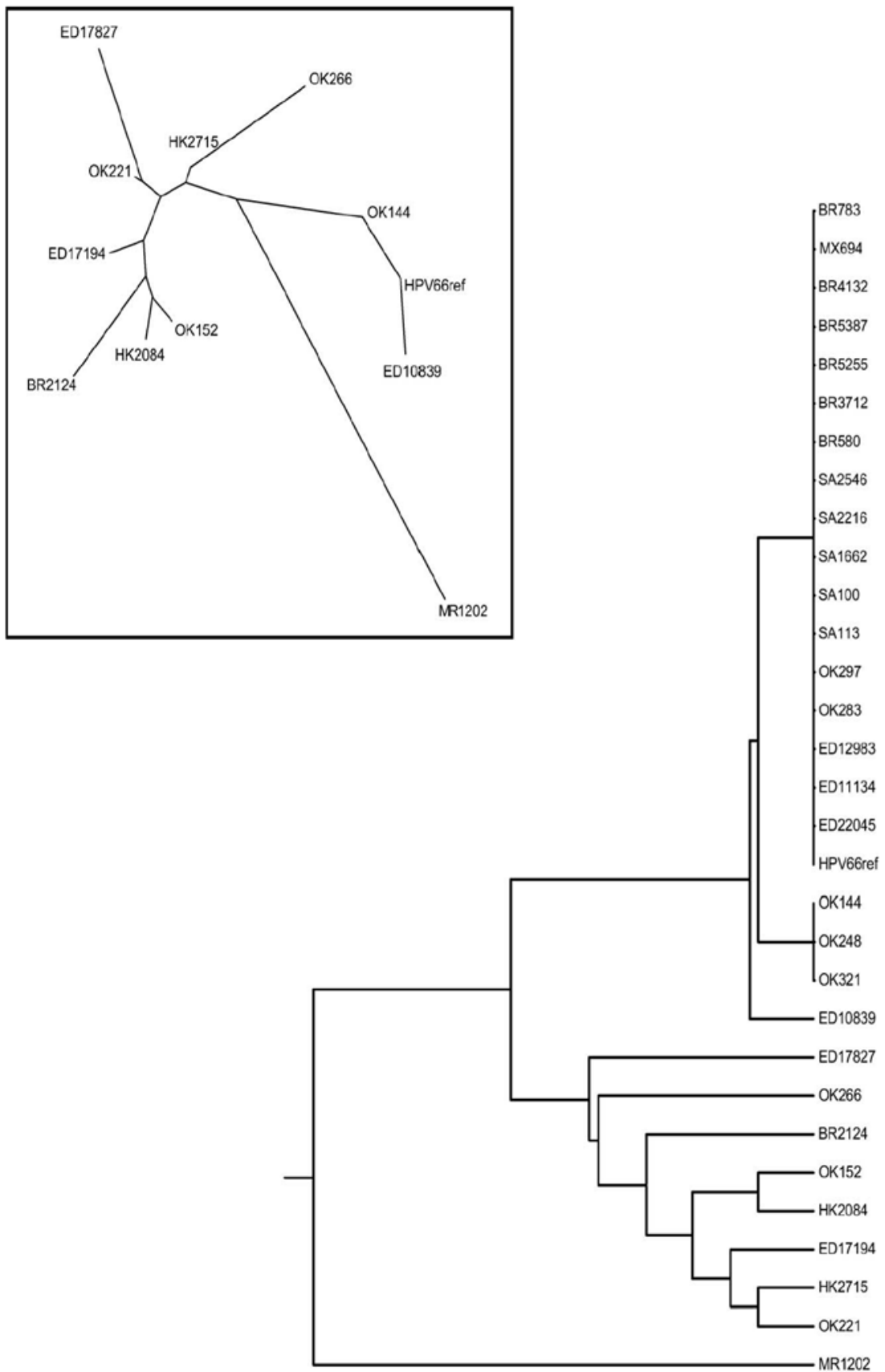
No estudo anteriormente mencionado, a análise filogenética das 44 amostras (nove variantes) do HPV-53 revelou uma ramificação dicotômica, nunca antes descrita para variantes do HPV e um pequeno aglomerado geográfico para amostras da África (Figura 12). Para o HPV-56, foram encontrados 17 variantes genômicos e também foi observada uma divisão dicotômica e um pequeno aglomerado geográfico para amostras da África do Sul, Hong Kong e Europa (Escócia – ED) (Figura 13). Para o HPV-66, foram analisadas 31 amostras e a análise das seqüências da LCR revelou a presença de 12 variantes monofiléticos. Dezoito das 31 amostras apresentaram a seqüência de referência e não foi encontrado nenhum agrupamento geográfico (Figura 14).



**Figura 12. Diversidade intragenótipo do HPV-53.** As árvores mostram a relação entre os nove variantes obtidos a partir da análise da LCR de 44 amostras. BR: Brasil; HK: Hong Kong; MX: México; ED: Escócia; SA: África do Sul; OK: Estados Unidos; ML: Mali. Fonte: Prado *et al* (2005).



**Figura 13. Diversidade intragenótipo do HPV-56.** As árvores filogenéticas representam a relação entre variantes do HPV-56 baseadas na homologia da LCR. BR: Brasil; HK: Hong Kong; MX: México; ED: Escócia; SA: África do Sul; OK: Estados Unidos e HE: Alemanha. Fonte: Prado *et al* (2005).

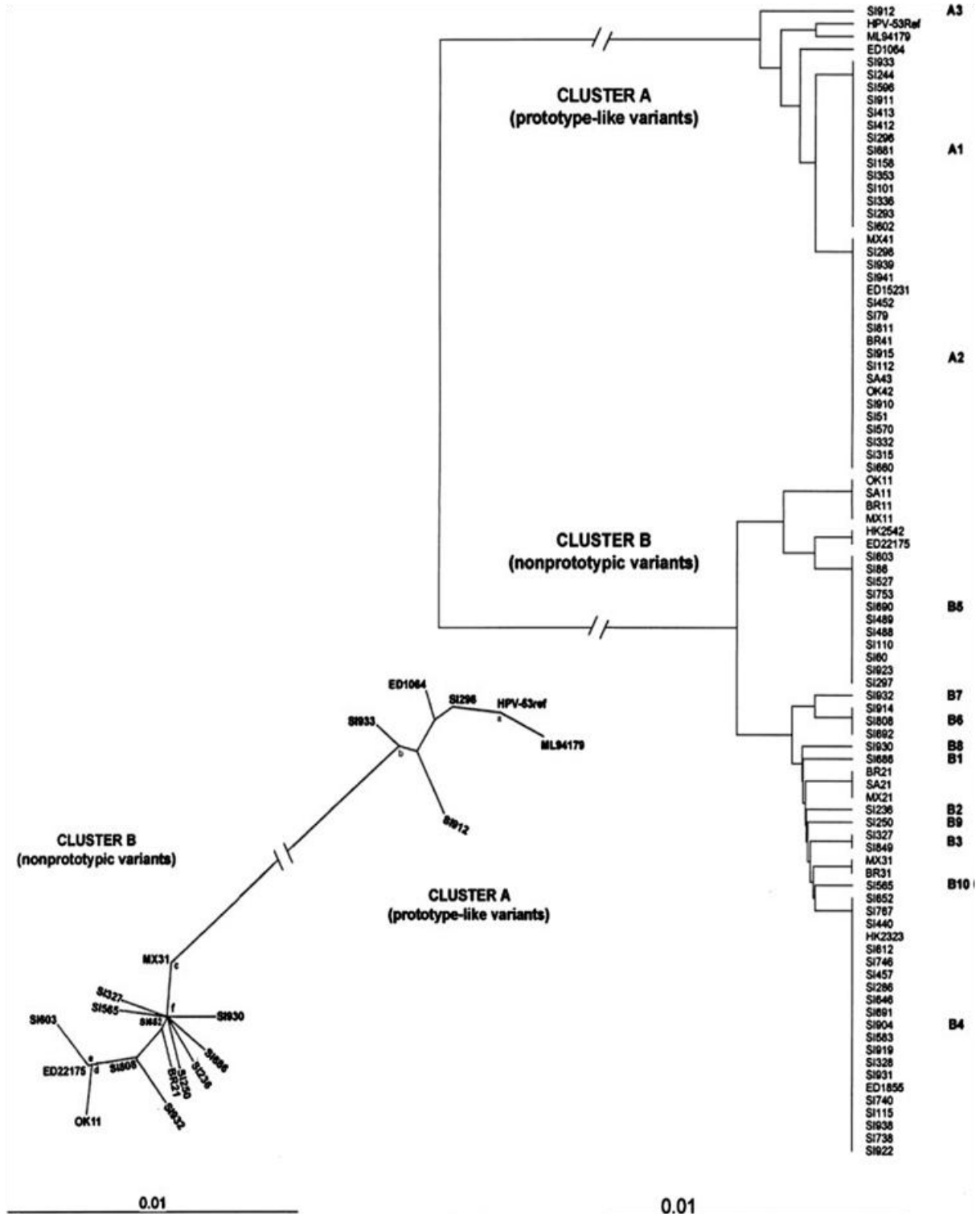


**Figura 14. Diversidade intragenótipo do HPV-66.** As árvores filogenéticas representam a relação entre variantes do HPV-66 baseadas na homologia da LCR. BR: Brasil; HK: Hong Kong; MX: México; ED: Escócia; SA: África do Sul; OK: Estados Unidos e MR: Marrocos. Fonte: Prado *et al* (2005).

Kocjan *et al* (2007) realizaram um trabalho que teve como objetivo investigar a diversidade genômica do HPV-53 em mulheres européias. Para isso foram analisadas três diferentes regiões genômicas: LCR, E6 e E7 de 70 amostras do HPV-53. Neste estudo foram encontrados 13 variantes genômicos para LCR e E6. Os variantes genômicos do HPV-53 encontrados não tiveram associação significativa a nenhum grau de displasia.

No estudo anteriormente mencionado, foram feitas análises filogenéticas para a LCR das 70 amostras analisadas no trabalho juntamente com 20 amostras de sete diferentes regiões geográficas (Brasil, Hong Kong, Mali, México, Escócia, África do Sul e Estados Unidos) descritas por Prado *et al* (2005). A análise confirmou a ramificação dicotômica somente descrita para subtipos do HPV verificada por Prado *et al* (2005) (Figura 15).





**Figura 15.** Árvore filogenética do HPV-53 baseada no algoritmo UPGMA. Variantes da Eslovênia (SI), Brasil (BR), Escócia (ED), Hong Kong (HK), Mali (ML), México (MX), África do Sul (SA) e Estados Unidos (OK). A árvore baseada no algoritmo neighbor-joining foi incluída para aumentar o impacto visual da filogenia em estrela como oposta à evolução por ramificação dicotômica. Os valores de *bootstrap* (%) acima de 50% são mostrados: a (72,7), b (100,0), c (100,0), d (50,5) e e (52,3) e f (74,0). O ramo que leva aos *clusters* A e B é mostrado por uma linha interrompida, uma vez que a distância entre eles poderia exceder a escala da figura. As escalas mostradas medem a distância genética em substituições por nucleotídeo. Fonte: Kocjan *et al* (2007).

## 2.10 JUSTIFICATIVA E OBJETIVOS

A infecção por HPVs de alto risco é considerada um fator necessário para o desenvolvimento do câncer de colo do útero (Franco e Harper, 2005), que é o segundo tipo de câncer mais comum em mulheres em todo o mundo. A cada ano, 490.000 mulheres são diagnosticadas com câncer de colo do útero e 290.000 morrem devido a essa doença (American Cancer Society, 2005). O impacto do câncer de colo do útero é maior em relação aos outros tipos de câncer por afetar mulheres de 30 a 60 anos, quando estão no pico da vida familiar e da produtividade (Mohan e Ind, 2004). A infecção por HPV é a infecção de transmissão sexual (ITS) de etiologia viral mais comum no mundo todo. Estima-se que 75% da população em idade reprodutiva esteja infectada com HPV genital. Desse total, 50% a 75% correspondem à infecção por genótipos oncogênicos (Villa *et al*, 2005; Scheurer *et al*, 2005).

Dentre os mais de 100 genótipos do HPV já descritos, dezoito são de particular interesse por estarem frequentemente associados ao desenvolvimento do câncer cervical: HPVs-16, -18, -30, -31, -33, -35, -39, -45, -52, -53, -56, -58, -59, -66, -67, -68, -70 e -85 (Muñoz *et al*, 2003). Os diferentes genótipos do HPV têm potencial oncogênico distinto e, da mesma forma, os variantes intragenótipo podem também apresentar oncogenicidade diferenciada (Villa *et al*, 2000).

Diferenças na seqüência de nucleotídeos que são encontradas entre variantes de um genótipo podem refletir em variações nos aminoácidos codificados, levando à diferença no potencial de infecção entre essas variantes (Bernard, 2002). Variações na seqüência de nucleotídeos da LCR podem resultar na perda da capacidade de ligação de um fator transcricional específico (Dong *et al*, 1994), na mudança do seu sítio de reconhecimento, aumentando ou diminuindo a afinidade do fator ou ainda na alteração de sua capacidade em interagir com outras proteínas (Schmidt *et al*, 2001). A mudança de um único nucleotídeo nesta região pode influenciar a replicação e a transcrição viral, por meio de sua influência na formação de complexos de proteínas regulatórias do DNA (Kämmer *et al*, 2002).

Vários trabalhos foram realizados no sentido de caracterizar os variantes dos HPVs-16 e -18 (Ong *et al*, 1993; Berumen *et al*, 1997; Xi *et al*, 1997; Veress *et al*, 1999; Kämmer, 2000; Villa *et al*, 2000). Porém, existem poucas informações sobre os variantes dos outros genótipos de alto risco. Prado *et al* (2005) iniciaram a caracterização molecular dos HPVs -53, -56 e -66. Kocjan *et al* (2007) trabalharam também com caracterização molecular do HPV-53 em mulheres européias.

No Brasil existe uma carência de dados sobre a variabilidade genética dos HPVs de alto risco -53, -56 e -66. Em 2003, nosso grupo de pesquisa analisou a variabilidade do gene L1 de

duas amostras de HPV-53 e uma amostra de HPV-66 (Cerqueira *et al*, 2003). Não existem dados sobre a variabilidade do gene E6 e da LCR para esses genótipos e sobre a variabilidade genética do HPV-56 circulante no Brasil.

A análise das seqüências de L1, E6 e LCR, dos variantes do HPV-53, -56 e -66 encontrados no Distrito Federal pode contribuir para uma melhor compreensão da filogenia desses genótipos, subsidiar estudos sobre infectividade e patogenicidade desses genótipos e fornecer dados sobre a variabilidade genética do HPV, que podem ser considerados em futuros testes de vacinas contra esses vírus.

Nosso grupo de pesquisa, na Universidade de Brasília, vem trabalhando na caracterização dos variantes dos HPVs de alto risco -16, -18, -31, -33, -35, -52 e -58.

Com a finalidade de ampliar os conhecimentos acerca da variabilidade genética dos HPVs -53, -56 e -66 esse trabalho teve como objetivo geral:

- Gerar dados sobre a variabilidade genética da LCR e dos genes L1 e E6 de dos HPVs-53, -56 e 66 encontrados no Distrito Federal e entorno.

E objetivos específicos:

- Verificar possíveis correlações entre variantes e patogenicidade por meio da comparação de exames citopatológicos com o número e a posição das variações;
- Verificar quais alterações na seqüência de nucleotídeos da LCR podem estar associadas a potenciais sítios de ligação de fatores transcricionais celulares e viral;
- Comparar a variabilidade genética encontrada com a descrita na literatura;
- Realizar a análise filogenética dessas amostras e comparar os resultados obtidos com os dados de outros estudos já realizados relativos à filogenia;
- Descrever seqüências da LCR, L1 e E6 dos variantes do HPV-53, -56 e 66 encontrados no Distrito Federal.

### 3 METODOLOGIA

#### 3.1 AMOSTRAS USADAS NO ESTUDO

Neste trabalho foram usadas amostras provenientes de um estudo de prevalência do HPV no Distrito Federal e entorno desenvolvido por nosso grupo de trabalho (Cerqueira *et al*, 2007) coletadas no período de 2004 a 2007 com a aprovação do Comitê de Ética e Pesquisa da Secretaria do Estado de Saúde do DF. As amostras foram previamente extraídas pelo método descrito por Nindl *et al* (1999), no qual o material coletado é transferido para tubos estéreis, contendo 5 ml de solução salina tamponada (PBS). Os tubos são centrifugados para a sedimentação das células, que são ressuspensas em 1 ml de solução 10 mM Tris-HCl (pH 8.3), 50 mM KCl e aquecidas a 95 °C por 15 minutos. Além disso, as amostras foram genotipadas por análise de polimorfismo dos fragmentos de restrição (*Restriction Fragment Length Polymorphism* – RFLP) ou seqüenciamento automático como pertencendo aos genótipos -53, -56 e -66 do HPV.

Foram incluídas no estudo: seis amostras do HPV-53; cinco amostras do HPV-56 e seis amostras do HPV-66. Os dados das amostras estão sumarizados na Tabela 1.

**Tabela 1. Dados das amostras caracterizadas neste estudo.** Resultados de exames citopatológicos, idade da paciente na data da coleta, genótipo(s) do(s) HPV(s) presente(s) na amostra. Na citopatologia: “LSIL” (lesão intraepitelial escamosa de baixo grau) indica efeito citopático compatível com HPV, “HSIL” (lesão intraepitelial escamosa de alto grau) indica NIC II (displasia moderada) ou NIC III (displasia severa e carcinoma in situ), “Normal” indica ausência de alterações citopatológicas e “Alterações benignas” indica alterações celulares benignas incluindo: inflamação, reparação e metaplasia<sup>12</sup> escamosa.

Amostra	Citopatologia	Idade da paciente	Genótipo
DF-33	HSIL (NICII)	42	HPV-53
DF-128	LSIL	28	HPV-53 e HPV-58
DF-141	Alterações benignas	37	HPV-53, HPV-32 e HPV-62
DF-153	Alterações benignas	37	HPV-53
DF-180	Alterações benignas	50	HPV-53
DF-231	Normal	27	HPV-53
DF-08	Alterações benignas	48	HPV-56
DF-26	Alterações benignas	31	HPV-56
DF-80	Alterações benignas	24	HPV-56
DF-123	HSIL (NICIII)	32	HPV-56
DF-230	Normal	48	HPV-56
DF-24	HSIL (NICII)	29	HPV-66, HPV-33
DF-41	Alterações benignas	39	HPV-66
DF-107	Normal	29	HPV-66
DF-198	Alterações benignas	36	HPV-66
DF-224	LSIL	20	HPV-66, HPV-33 e HPV-61
DF-228	Alterações benignas	21	HPV-66

<sup>12</sup> Metaplasia: Conversão de um tipo de célula madura em outro tipo de célula.

### 3.2 AMPLIFICAÇÃO DO DNA VIRAL

Para a amplificação do gene E6 foram desenhados iniciadores específicos, para a amplificação da LCR foram usados iniciadores específicos descritos na literatura (Prado *et al*, 2005) e para a amplificação de um fragmento do gene L1, foi usado o par de iniciadores degenerados MY09/11 (Manos *et al*, 1989) (Tabela 2).

**Tabela 2. Iniciadores usados na amplificação da LCR, E6 e L1.** Iniciadores específicos para amplificação da LCR, e dos genes E6 e L1 dos HPVs-53, -56 e -66. Iniciadores da LCR de acordo com Prado *et al* (2005) e iniciadores consenso degenerados MY09/11 (Manos *et al*, 1989) para a amplificação de um fragmento do gene L1 de acordo com Manos *et al*, 1989, onde r = a/g, w = a/t, m = a/c, e y = c/t.

Iniciadores	Seqüência	Genótipo amplificado	Posição no genoma	Fragmento esperado
53E6F	5'-agtatctgtctagtaaaccagg-3'	HPV-53	7766 - 7787	693 pb
53E6R	5'-tatattgtggaagtgttggtac-3'	HPV-53	0581 - 0602	
53LCRF	5'-tttgcattgttgaataaatat-3'	HPV-53	7270 - 7291	
53LCRR	5'-agtaggattgttactttc-3'	HPV-53	0001 - 0018	602 pb
56E6F	5'-aggccaagtatctgtcagta-3'	HPV-56	7741 - 7760	714 pb
56E6R	5'-tctaatacaacgtcttcgagc-3'	HPV-56	0592- 0612	
56LCRF	5'-tgtgtcattattgtggctttgtttgt-3'	HPV-56	7323 - 7350	
56LCRR	5'-agctgccttttatatgtaccgttttc-3'	HPV-56	0056- 0080	603 pb
66E6F	5'-tcattttacaaaaccgtttacgg-3'	HPV-66	7772 - 7794	697 pb
66E6R	5'-attgtaggtcaattccgtttg-3'	HPV-66	0623- 0644	
66LCRF	5'-aacgatagttgtgtgtgtg-3'	HPV-66	7151 - 7170	
66LCRR	5'-acctgttttagcagacc-3'	HPV-66	7793- 7810	659 pb
MY09	5'-cgtccmarrggawactgatc-3'	Todos	Distinta para cada genótipo	450 pb
MY11	5'-gcmcagggwctataaayatgg-3'	Todos	Distinta para cada genótipo	

As reações da PCR (*Polimerase Chain Reaction*) foram realizadas em um volume final de 50µL, contendo 5-15µL do material genômico obtido, 1,5mM de MgCl<sub>2</sub> (Invitrogen), 250µM de dATP, dCTP, dTTP e dGTP (Invitrogen), 40pmoles de cada *primer* e 1U de Taq DNA polimerase (Invitrogen). Quarenta ciclos de amplificação foram conduzidos no termociclador MJ Research PTC-100 (primeiro ciclo: 95°C-5min, 55°C-1min, 72°C-2min; segundo ao 39º ciclo: 95°C-1min, 55°C-1min, 72°C-2min; 40º ciclo: 95°C-1min, 55°C-1min, 72°C-8min).

Todas as amostras foram amplificadas na presença de controles negativos. Os produtos da PCR obtidos foram analisados em gel de agarose 1% impregnado com brometo de etídio para verificação da qualidade do DNA.

### 3.3 CLONAGEM

A reação de ligação dos produtos de PCR no plasmídeo pGEM®-T easy (Promega) foi realizada seguindo as instruções do fabricante.

Os plasmídeos foram utilizados para transformar bactérias *Escherichia coli* da linhagem DH5 $\alpha$ , pelo método de choque térmico. Inicialmente, 10  $\mu$ L da solução contendo os plasmídeos foram incubados com 300  $\mu$ L das células competentes no gelo por 40 minutos. Em seguida, as células foram submetidas a 42 °C por 2 minutos, em banho-maria, e novamente transferidas para o banho de gelo por 1 minuto. Foram adicionados 700  $\mu$ L de meio líquido 2xYT (1,6 mg/mL de triptona, 1,0 mg/mL de extrato de levedura e 85,6 mM de NaCl), que em seguida foi incubado a 37 °C por 1 hora. Após a incubação, os tubos foram centrifugados por 5 minutos a 3000 rpm (Jouan S.A. A14/V1), e 700  $\mu$ L foram removidos por aspiração. As bactérias foram espalhadas em placas de Petri contendo meio sólido 2xYT, com 16,7  $\mu$ mol de IPTG, 800  $\mu$ g de X-Gal e 4mg/mL de ampicilina.

As placas de Petri foram incubadas *overnight* em estufa a 37 °C. Os clones recombinantes foram selecionados e incubados *overnight* a 37°C sob agitação em meio líquido 2xYT, contendo 4mg/mL de ampicilina. O DNA plasmidial foi extraído utilizando-se o *Kit Sephaglas FlexiPrep*<sup>™</sup> (GE Healthcare) seguindo as instruções do fabricante.

O DNA plasmidial foi utilizado como molde para amplificação por PCR, utilizando os iniciadores universais M13. A reação foi realizada em um volume final de 50  $\mu$ L, contendo: 1  $\mu$ L do DNA plasmidial, 0,2 mM de dNTPs (Invitrogen), 0,4  $\mu$ M dos iniciadores M13F/M13R (Integrated DNA Technologies), 1,25 U da enzima Taq DNA Polimerase (Invitrogen), 1X de tampão da enzima (Invitrogen) e 1 mM de cloreto de magnésio (Invitrogen). As seguintes condições foram adotadas: desnaturação a 94°C por 2 minutos, seguidos de 35 ciclos de desnaturação a 94°C por 30 segundos, anelamento dos iniciadores a 55°C por 25 segundos e extensão a 72°C por 30 segundos.

Os clones amplificados foram submetidos à eletroforese em gel de agarose na concentração de 1%. Brometo de etídio foi utilizado para permitir a visualização dos fragmentos de DNA. Controles negativos foram incluídos durante os procedimentos de amplificação, para monitorar possíveis contaminações.

### 3.4 SEQÜENCIAMENTO

Os produtos da PCR que apresentaram o número de pares de base esperado foram purificados por precipitação com acetato de amônio 7,5M e etanol e seqüenciados

automaticamente pelo método dideoxi-terminal fluorescente nos seqüenciadores Megabace System 500 (Amersham Pharmacia Biotec), no laboratório de Biologia Molecular da Universidade de Brasília, e ABI PRISM modelo 3100, na plataforma de seqüenciamento do Cenargen - Embrapa.

### 3.5 ANÁLISE DAS SEQÜÊNCIAS

As seqüências foram analisadas quanto às divergências nucleotídicas em relação às seqüências de referência dos HPVs-53 (número de acesso: X74482), -56 (número de acesso: X74483) e -66 (número de acesso: U31794) e às seqüências desses genótipos já depositadas no GenBank por Cerqueira *et al* (2003), Prado *et al* (2005), Cerqueira *et al* (2007) e Kocjan *et al* (2007). As comparações foram feitas usando-se o programa BLAST (*Basic Local Alignment Search Tool*) (Altschul *et al*, 1997) seguidas da análise dos alinhamentos obtidos pelo programa CLUSTALW (Thompson *et al*, 1994). As substituições de nucleotídeos encontradas foram confirmadas por um segundo PCR e seqüenciamento.

### 3.6 PREDIÇÃO DE SÍTIOS DE LIGAÇÃO PARA FATORES DE TRANSCRIÇÃO NA LCR

Para a predição de sítios de ligação para fatores de transcrição foram usadas as seqüências de referência: HPV-6 (número de acesso no GenBank X00203); HPV-53 (número de acesso no GenBank: X74482); HPV-56 (X74483) e HPV-66 (U31794).

Os sítios de ligação para fatores de transcrição na LCR foram preditos pelo programa MATCH™, desenvolvido para localizar potenciais sítios de ligação para fatores de transcrição (Kel *et al*, 2003). MATCH™ usa a biblioteca de matrizes de nucleotídeos do TRANSFAC® 7.0 (Wingender *et al*, 1996) e procura por seqüências que combinem entre as matrizes de nucleotídeos que definem um sítio de ligação para determinado fator de transcrição e uma dada seqüência problema. O programa encontra-se disponível em: <http://www.gene-regulation.com/pub/programs.html#match>.

Os valores de *cut-off* e níveis de coincidência entre as seqüências dos sítios de ligação e as seqüências da LCR foram ajustados para minimizar, ao mesmo tempo, o número de falsos negativos e falsos positivos.

Para esta análise foram usadas apenas matrizes dos sítios de ligação para fatores de transcrição já estabelecidos e comprovados para o HPVs-16: AP-1, E2, GRE, NF-1, Oct-1, TATA, YY1, CEBP, Sp1 e c-Myc.

### 3.7 ANÁLISE FILOGENÉTICA

As seqüências da LCR obtidas foram, juntamente com as seqüências previamente descritas por Prado *et al* (2005) e Kocjan *et al* (2007), usadas na construção das árvores filogenéticas dos genótipos do HPV-53, -56 e -66. Para a análise filogenética do HPV-53 foram usados fragmentos de 508pb, para o HPV-56 foram usados fragmentos de 514pb e para o HPV-66 os fragmentos foram de 607pb. O alinhamento das seqüências foi realizado usando o programa CLUSTAL W (Thompson *et al*, 1994) e a construção de árvores filogenéticas pelo método de Máxima Verossimilhança usando o programa PHYLIP (Felsenstein, 1981). As árvores obtidas foram comparadas àquelas descritas por Prado *et al* (2005) e Kocjan *et al* (2007). Os valores de *bootstrap* foram calculados pelo método UPGMA com algoritmo Log Det e 1000 réplicas.

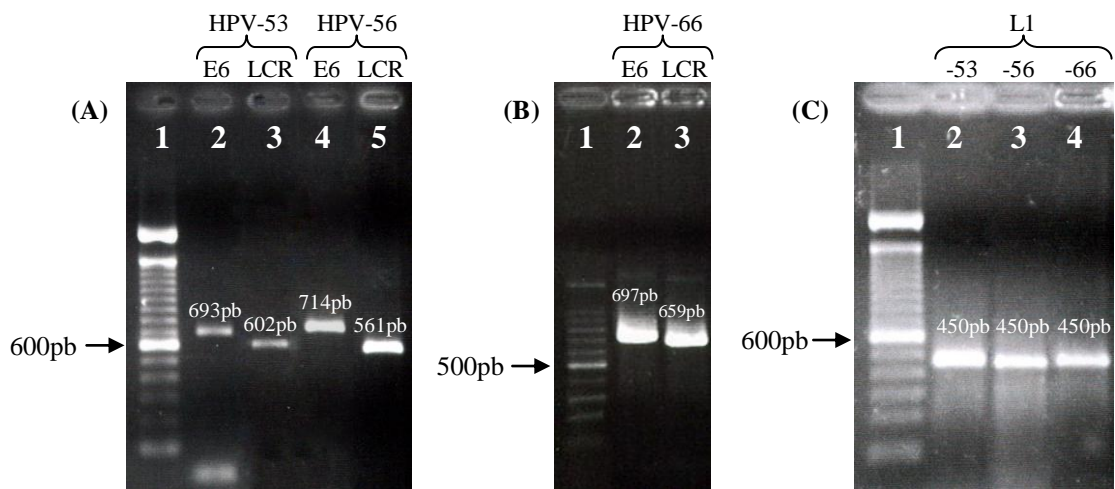


## 4 RESULTADOS

### 4.1 AMPLIFICAÇÃO DO DNA VIRAL

O material genético das seis amostras do HPV-53 e cinco amostras do HPV-56 foi amplificado para os genes E6 e L1 e para a LCR. As seis amostras do HPV-66 tiveram o material genético amplificado para os genes E6 e L1 e cinco delas para a LCR. Os produtos da PCR apresentaram o fragmento de número de pares de base esperado (Figura 16).

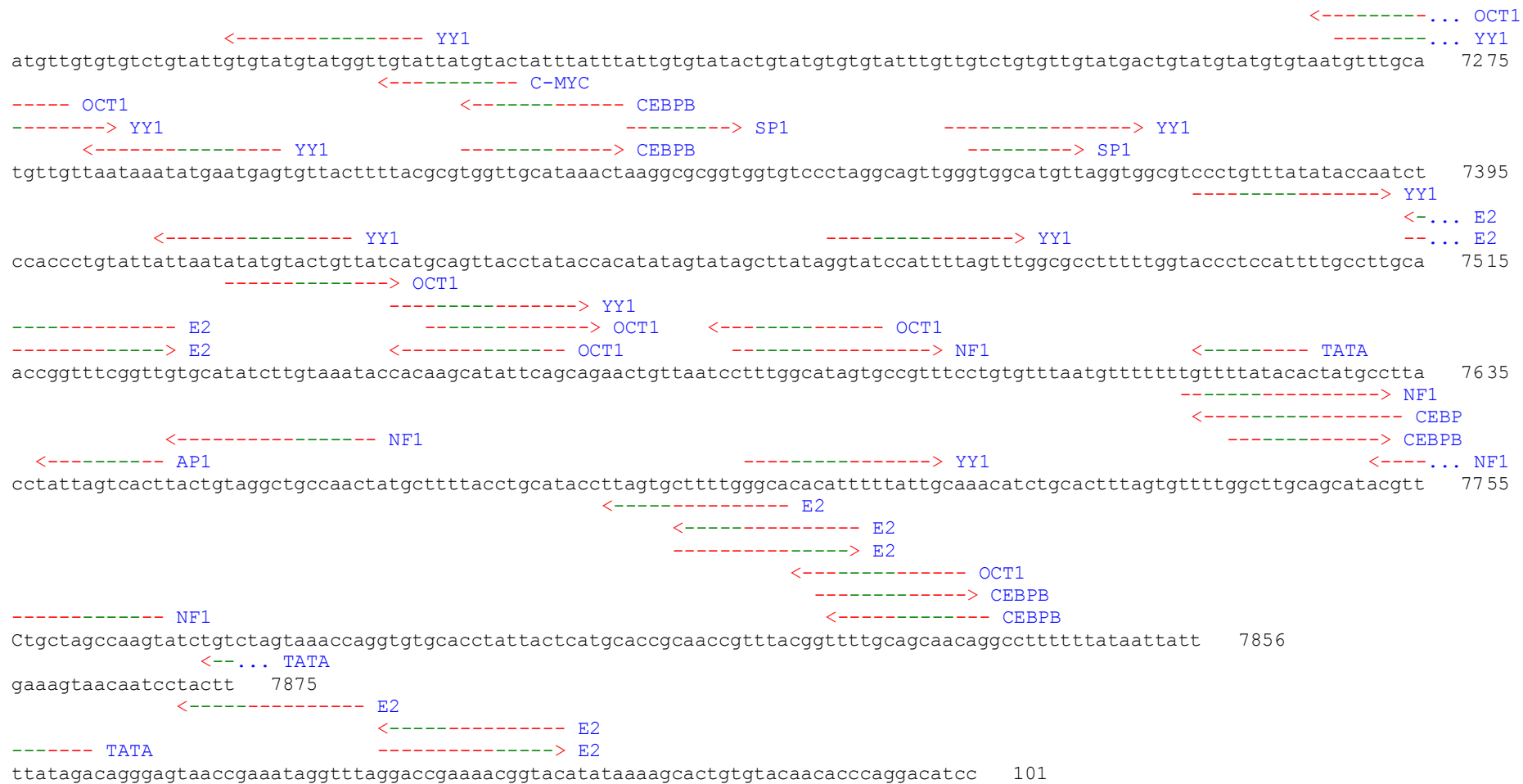
Os produtos da PCR foram seqüenciados automaticamente e as seqüências geradas estão no anexo I.



**Figura 16. Análise da qualidade dos fragmentos de DNA amplificados por PCR.** Géis de agarose 1% impregnados com brometo de etídio. (A) 1: marcador de 100pb (Invitrogen – banda forte: 600pb), 2: fragmento amplificado de amostra do HPV-53 para o gene E6 (aproximadamente 693pb), 3: fragmento amplificado de amostra do HPV-53 para a LCR (aproximadamente 602pb), 4: fragmento amplificado de amostra do HPV-56 para o gene E6 (aproximadamente 714pb), 5: fragmento amplificado de amostra do HPV-56 para a LCR (aproximadamente 561pb). (B) 1: marcador de 100pb (Ludwig Biotecnologia – banda forte: 500pb), 2: fragmento amplificado de amostra do HPV-66 para o gene E6 (aproximadamente 697pb), 3: fragmento amplificado de amostra do HPV-66 para a LCR (aproximadamente 659pb). (C) 1: marcador de 100pb (Invitrogen – banda forte: 600pb), 2 a 4: fragmentos amplificados de amostras dos HPVs-53, -56 e -66 para o gene L1 (aproximadamente 450pb).

### 4.2 PREDIÇÃO DE SÍTIOS DE LIGAÇÃO PARA FATORES DE TRANSCRIÇÃO NA LCR

O mapeamento dos potenciais sítios de ligação para cada genótipo foi gerado pelo programa MATCH™ usando as seqüências de referência dos HPVs-53 (número de acesso no CenBank: X74482), -56 (X74483) e -66 (U31794) (Figuras 17 a 19). No anexo II encontram-se o mapeamento no formato original fornecido pelo programa com os nomes das matrizes usadas na predição.



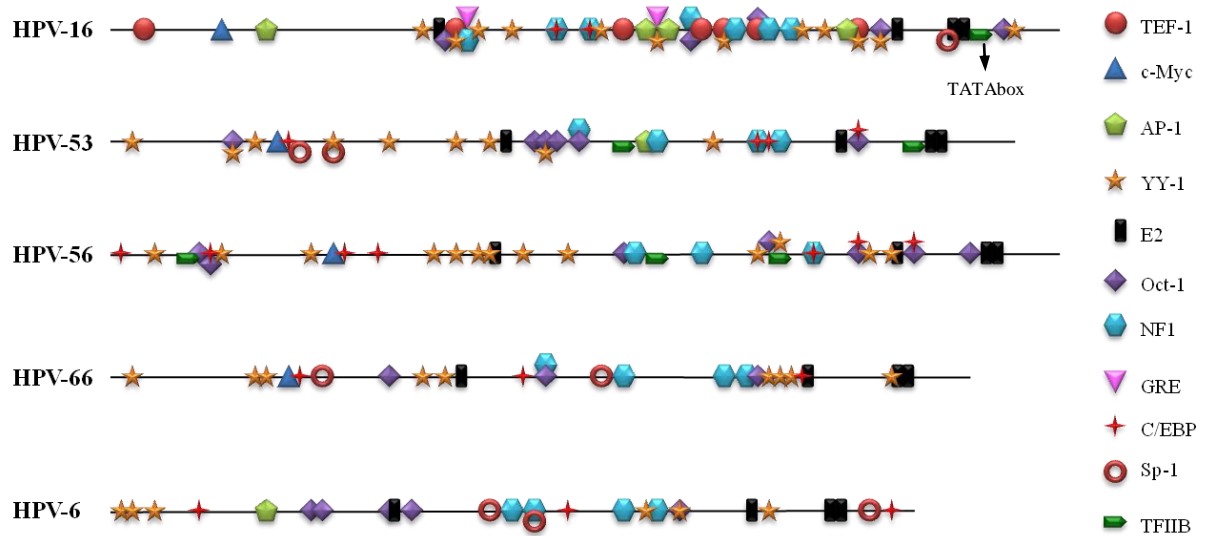
**Figura 17. Mapeamento dos potenciais sítios de ligação na LCR do HPV-53.** Em azul: o nome do fator de transcrição. Em vermelho: o sítio de ligação predito pelo programa MATCH™. Em verde o *core* do sítio de ligação. A direção da seta indica o sentido da fita de DNA na qual o fator se liga: <----- fita 3'-5'; -----> fita 5'-3'.



**Figura 18. Mapeamento dos potenciais sítios de ligação na LCR do HPV-56.** Em azul: o nome do fator de transcrição. Em vermelho: o sítio de ligação predito pelo programa MATCH™. Em verde o *core* do sítio de ligação. A direção da seta indica o sentido da fita de DNA na qual o fator se liga: <---- fita 3'- 5'; ----> fita 5'- 3'.



**Figura 19. Mapeamento dos potenciais sítios de ligação na LCR do HPV-66.** Em azul: o nome do fator de transcrição. Em vermelho: o sítio de ligação predito pelo programa MATCH™. Em verde o *core* do sítio de ligação. A direção da seta indica o sentido da fita de DNA na qual o fator se liga: <----- fita 3'-5'; -----> fita 5'-3'.



**Figura 20. Representação esquemática da LCR dos HPVs-16, -53, -56, -66 e -6** representando os potenciais sítios de ligação para fatores de transcrição. Os potenciais sítios de ligação para fatores de transcrição da LCR dos HPVs de alto risco genótipos -53, -56, -66 e do HPV de baixo risco genótipo -6 foram preditos pelo programa MATCH™. Para o HPV-16 foram localizados sítios de ligação para fatores de transcrição já descritos na literatura.

A figura 20 apresenta os esquemas simplificados da LCR dos HPVs de alto risco genótipos -53, -56 e -66 e do HPV de baixo risco genótipo -6 simulando a ligação dos fatores de transcrição, representados pelos símbolos, aos respectivos sítios preditos pelo programa MATCH™. Para efeito de comparação foi usada a LCR do HPV-16 que já foi bem caracterizada e os fatores usados no esquema estão descritos na literatura (Sibbet e Campo, 1990; O'Connor *et al*, 1995; Abba *et al*, 2004). Comparando-se as LCRs dos HPVs-6, -53, -56 e -66 com a LCR do HPV-16 pode-se observar, em todos os genótipos, a presença dos quatro sítios de ligação para a proteína viral E2, que definem as três regiões da LCR: 5', segmento central e 3'; sítios de ligação para C/EBP, YY1 e Oct-1 e NF-1. Não foram detectados sítios de ligação para receptores de glicocorticóide ou progesterona (GRE) e para TEF-1. Além disso, o HPV-6 não apresentou sítio de ligação para c-Myc e TFIIB, fator de transcrição responsável pelo posicionamento adequado da RNA polimerase II no sítio de início da transcrição por meio do reconhecimento do TATAbox. O HPV-56 não apresentou sítio de ligação para Sp-1 e AP-1 e o HPV-66 não apresentou sítio de ligação para AP-1 e TFIIB.

### 4.3 ANÁLISE DAS SEQÜÊNCIAS

#### 4.3.1 HPV-53

Na análise das seqüências de E6 das amostras do HPV-53, foram detectados seis variantes e 19 variações nucleotídicas pontuais em relação à seqüência de referência do HPV-53 (Tabela 3).

Oito substituições nucleotídicas levaram à troca de aminoácidos (Tabela 3). As variações T52K, A94V e S104L foram não conservativas. A variação A94V foi causada pela substituição nucleotídica de C para T na posição 382. As outras trocas de aminoácidos foram todas conservativas.

**Tabela 3. Variabilidade genética de E6 das amostras do HPV-53.** Em cinza, nucleotídeos sem alteração em relação à seqüência de referência e em azul, variações nucleotídicas não descritas na literatura. Em vermelho estão as alterações de aminoácidos não conservativas.

HPV-53	Posição de variação de nucleotídeos em E6																			
Amostras	1	1	1	1	1	1	1	1	2	2	2	2	2	2	3	3	3	3	4	
	1	1	2	3	5	5	7	7	1	2	3	5	6	6	0	4	5	8	1	
	0	2	9	1	2	8	4	5	5	4	0	6	1	7	5	5	0	2	2	
Referência	T	A	G	A	C	A	A	A	G	G	G	C	C	G	A	T	C	C	C	
DF-33	C			G	T						T		T		G					
DF-128																	T	A		
DF-141	C			G	T					A	T		T		G					
DF-153		G	C	G	T	C	C					T	A	T			C		T	
DF-180															C				T	
DF-231								C	A											
Posição dos códons alterados		4	10				25	25					52				82		94	104
Referência		Q	E				K	K					T				Y		A	S
Aminoácidos alterados		R	Q				Q	T					K				H		E/V	L

Em relação à seqüência de referência do HPV-53, para a LCR também foram detectados seis variantes com um total de 20 variações pontuais e uma inserção de 4pb entre as posições 7530 e 7531 (Tabela 4).

Das vinte variações nucleotídicas observadas na LCR, nove estão localizadas em potenciais sítios de ligação para os fatores de transcrição: YY1, E2, Oct-1, AP-1, NF-1 e TATA, sendo três delas localizadas no *core* dos sítios de ligação para E2 e TATA. As variações nucleotídicas de T para G e T para C nas posições 7618 e 7619, respectivamente, se encontram no *core* do potencial TATAbox (Tabela 4).

**Tabela 4. Variabilidade genética da LCR das amostras do HPV-53.** Em cinza, nucleotídeos sem alteração em relação à sequência de referência e em azul, variações nucleotídicas não descritas na literatura. Potenciais sítios de ligação para fatores de transcrição alterados em preto: a variação nucleotídica ocorreu na sequência próxima ao *core* do potencial sítio de ligação. Em vermelho: a variação ocorreu no *core* do potencial sítio de ligação.

HPV-53	Posição de variação de nucleotídeos na LCR																							
Amostras	7	7	7	7	7	7530-7531	7	7	7	7	7	7	7	7	7	7	7	7	7	7				
	4	4	4	5	5		5	5	6	6	6	6	6	6	6	6	6	7	7	7	7	8		
	2	2	3	0	2	3	3	1	1	3	4	5	7	8	8	5	5	8	9	1				
	2	8	2	8	8	1	8	8	9	4	8	5	6	5	6	6	8	4	7	0				
Referência	T	C	C	G	T		G	T	T	T	T	T	A	C	C	T	C	G	C	A	C			
DF-33	A	T		A	C	T	G	G	G	A	C	G	C	C	A	G	G	T		T	C			
DF-128				A																				
DF-141	A	T		A	C	T	G	G	G	A	C	G	C	C		G	G	T		T	A	T	C	T
DF-153	A	T	G	A	C	T	G	G	G	A	C	G	C	C		G	G	T		T	A	T	C	T
DF-180				A																				
DF-231				A																				
Potencial sítio de ligação alterado	YY1			YY1	E2						Oct-1	TAT		AP-1	NF-1					NF-1			E2	

Foram analisadas cinco amostras do HPV-53 para o gene L1, apresentando três variantes distintos e 12 variações nucleotídicas pontuais. Duas delas levaram à troca de aminoácidos. Uma conservativa (K355E) e outra não conservativa (P430S) (Tabela 5).

**Tabela 5. Variabilidade genética de L1 das amostras do HPV-53.** Em cinza, nucleotídeos sem alteração em relação à sequência de referência e em azul, variações nucleotídicas não descritas na literatura. Em vermelho está a alteração de aminoácidos não conservativa.

HPV-53	Posição de variação de nucleotídeos em L1											
Amostras	6	6	6	6	6	6	6	6	6	7	7	7
	5	6	6	7	7	7	9	9	9	0	0	0
	9	7	8	0	1	4	0	4	9	3	4	4
	6	7	3	4	7	0	8	2	8	4	0	1
Referência	C	T	C	T	A	A	G	C	G	T	C	A
DF-33			T	C		G	A	T	A			
DF-128					G							
DF-153	A	A	T	C		G	A		A	C	A	C
DF-180												
DF-231												
Posição dos códons alterados					355			430				
Referência					K			P				
Aminoácidos alterados					E			S				

#### 4.3.2 HPV-56

Entre as cinco amostras do HPV-56 analisadas para o gene E6, foram detectados três variantes distintos e quatro variações nucleotídicas pontuais. Três substituições de nucleotídeos levaram à troca de aminoácidos (Tabela 6), duas delas não são conservativas: S14R e K54N.

**Tabela 6. Variabilidade genética de E6 das amostras do HPV-56.** Em cinza, nucleotídeos sem alteração em relação à sequência de referência e em azul, variações nucleotídicas não descritas na literatura. Em vermelho estão as alterações de aminoácidos não conservativas.

HPV-56	Posição de variação de nucleotídeos em E6			
	1	2	2	4
Amostras	4	6	7	4
	1	3	9	3
Referência	A	A	G	T
DF-08				
DF-26	C	C		A
DF-80			A	
DF-123	C	C		
DF-230	C	C		
Posição dos códons alterados	14	54	60	
Referência	S	K	D	
Aminoácidos alterados	R	N	N	

Em relação às amostras da LCR analisadas, foram detectados cinco variantes e 17 variações nucleotídicas e uma deleção de 42pb (Tabela 7). Das 17 variações nucleotídicas observadas na LCR, seis estão localizadas em potenciais sítios de ligação para fatores de transcrição e cinco delas localizadas no *core* desses sítios. A deleção dos 42pb inclui um potencial sítio de YY1. As variações nucleotídicas de T para G e T para A nas posições 7383 e 7391, respectivamente, se encontram no mesmo potencial sítio de YY1, sendo a segunda variação no *core* (Tabela 7). Um segundo potencial sítio de ligação para YY1 é alterado pelas variações nas posições 7441 (A para C) e 7446 (deleção de A). A variação na posição 7441 se localizou no *core* de YY1 e a variação na posição 7446 se localizou na sequência próxima ao *core* de YY1 e no *core* do sítio de ligação para E2, que sobrepõe o potencial sítio de YY1 (Tabela 7).

**Tabela 7. Variabilidade genética da LCR das amostras do HPV-56.** Em cinza, nucleotídeos sem alteração em relação à sequência de referência e em azul, variações nucleotídicas não descritas na literatura. Potenciais sítios de ligação para fatores de transcrição alterados em preto: a variação ocorreu na sequência próxima ao *core* do potencial sítio de ligação. Em vermelho: a variação ocorreu no *core* do potencial sítio de ligação.

HPV-56	Posição de variação de nucleotídeos na LCR																	
	7	7	7	7	7	7	7485-7526	7	7	7	7	7	7	7	0	0	0	0
Amostras	3	3	3	4	4	4		5	5	5	5	6	7	8	0	0	0	1
	7	8	9	4	4	6		5	8	9	9	2	7	0	0	3	2	0
	4	3	1	1	6	9		9	1	3	6	8	7	0	9	5	8	0
Referência	T	T	T	A	A	G	42 pb	T	T	G	T	A	G	T	C	G	T	C
DF-08													A					
DF-26				C		T	Deleção dos 42 pb	C	C	A	C	C		G	T	A		T
DF-80	A	G	A				Deleção dos 42 pb		C	A	G	C		G		A	C	
DF-123					-	T	Deleção dos 42 pb	C	C		C	C		G		A		T
DF-230						T	Deleção dos 42 pb	C	C		C	C		G	T	A		T
Potencial sítio de ligação alterado				YY1	YY1	YY1/E2	YY1	deleta sítio de YY1						YY1			E	



Na análise de L1 das cinco amostras do HPV-56, um variante foi detectado em três amostras apresentando duas variações nucleotídicas pontuais em relação à seqüência de referência (Tabela 8). As variações não provocaram troca de aminoácido.

**Tabela 8. Variabilidade genética de L1 das amostras do HPV-56.** Em cinza, nucleotídeos sem alteração em relação à seqüência de referência.

HPV-56	Posição de variação de nucleotídeos em L1	
	Amostras	6 8 2 6
Referência	C	G
DF-08		
DF-26	T	A
DF-80		
DF-123	T	A
DF-230	T	A

#### 4.3.3 HPV-66

Em relação à seqüência de referência do HPV-66, para o gene E6 foi detectado um variante em apenas uma entre as seis amostras analisados, contendo quatro variações pontuais (Tabela 9). A variação de A para G na posição 378 levou à troca não conservativa de aminoácido S93G.

**Tabela 9. Variabilidade genética de E6 das amostras do HPV-66.** Em cinza, nucleotídeos sem alteração em relação à seqüência de referência. Em vermelho estão as alterações de aminoácidos não conservativas.

HPV-66	Posição de variação de nucleotídeos em E6			
	Amostras	2 3 4	2 6 4	3 7 8
Referência	C	C	A	T
DF-24				
DF-41				
DF-107	T	T	G	C
DF-198				
DF-224				
DF-228				
<b>Posição dos códons alterados</b>			93	
Referência			S	
Aminoácidos alterados			G	

Não foram detectadas variações nucleotídicas na LCR do HPV-66 em relação à seqüência de referência.

Das cinco amostras do HPV-66 analisadas para o gene L1, foram detectados dois variantes distintos em relação à seqüência de referência, com 12 variações nucleotídicas pontuais (Tabela 10). Nenhuma das substituições levou à troca de aminoácidos.

**Tabela 10. Variabilidade genética de L1 das amostras do HPV-66.** Em cinza, nucleotídeos sem alteração em relação à seqüência de referência e em azul, variações nucleotídicas não descritas na literatura.

HPV-66	Posição de variação de nucleotídeos em L1											
Amostras	6	6	6	6	6	6	6	6	6	7	7	7
	5	6	7	8	8	8	8	9	9	0	0	0
	9	6	1	4	5	5	7	2	8	1	3	3
	7	0	1	9	5	8	7	7	4	7	5	6
Referência	C	C	C	T	C	A	G	G	A	G	T	A
DF-24												
DF-41												
DF-107	T	T	A	C	A	G		A	G	A	A	C
DF-198												
DF-228							A					

Não foi detectada nenhuma associação entre os variantes encontrados e os diferentes graus de displasia das amostras analisadas neste estudo.

#### 4.4 ANÁLISE FILOGENÉTICA

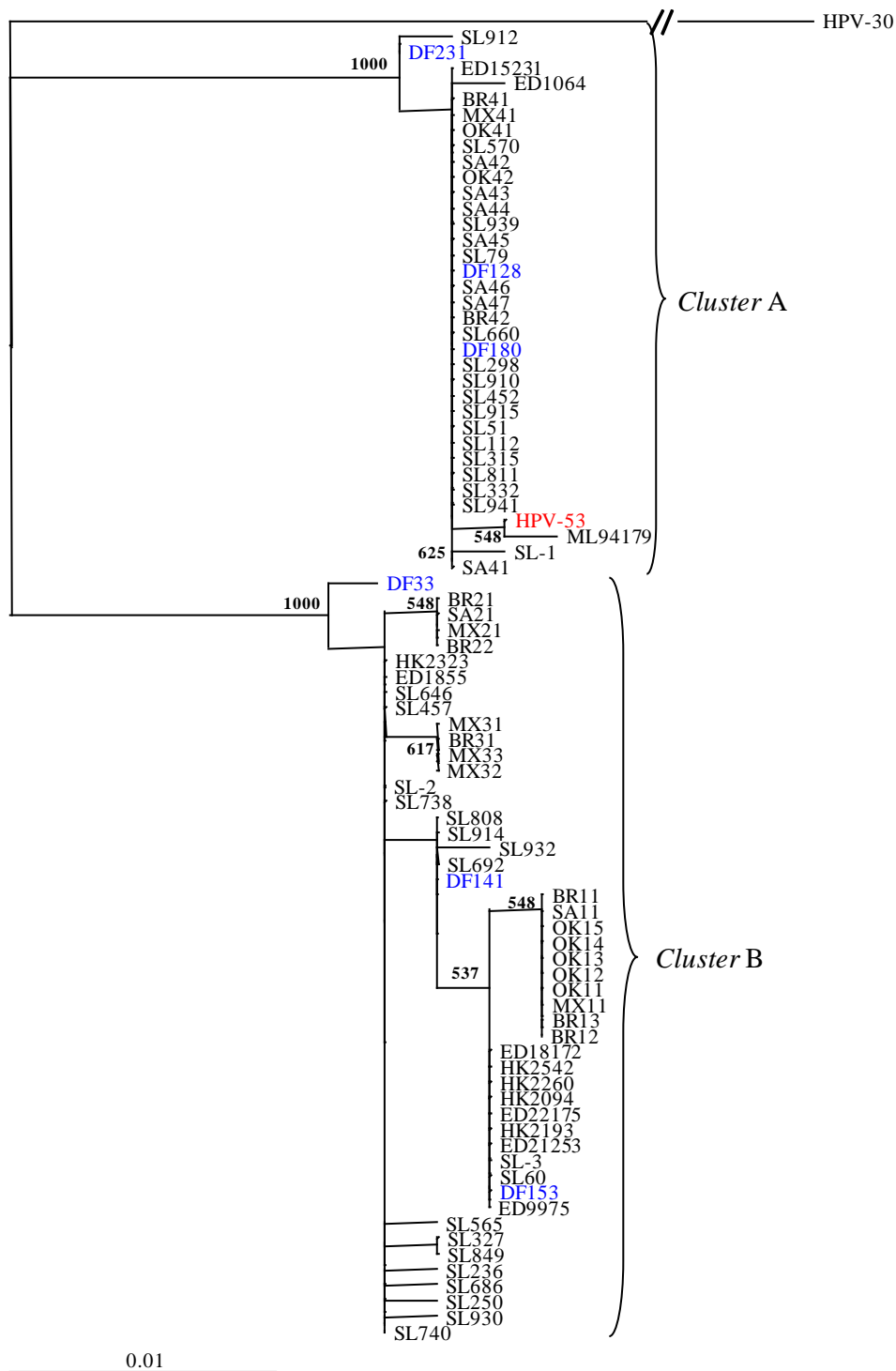
As seqüências obtidas neste trabalho foram usadas, juntamente com as seqüências obtidas por Prado *et al* (2005) e Kocjan *et al* (2007), para a realização da análise filogenética.

##### 4.4.1 HPV-53

A análise filogenética da LCR do HPV-53 foi realizada usando-se um fragmento de 508pb de 120 amostras. Destas, 14 foram provenientes do Brasil, sendo 6 do DF (caracterizadas neste estudo) e 8 de São Paulo; 70 da Eslovênia; 7 dos Estados Unidos; 6 do México; 7 da Escócia; 5 da China; 9 da África do Sul; 1 de Mali e a seqüência de referência do HPV-53.

A análise filogenética revelou a formação de dois agrupamentos separados por uma ramificação dicotômica (Figura 21). Um dos ramos (*cluster* A) apresentou a seqüência de referência do HPV-53 e 48 amostras, entre elas três das seis amostras analisadas neste trabalho. O outro ramo (*cluster* B) conteve as 72 amostras restantes. As amostras do DF, assim como as amostras do Brasil, não formaram agrupamento distinto.

Na figura 21 observa-se a formação de vários agrupamentos de amostras da Eslovênia. Cinco das sete amostras dos Estados Unidos, e quatro de cinco amostras da China se agruparam no *cluster* B. De seis amostras do México, três se agruparam com uma amostra do Brasil formando um pequeno aglomerado geográfico no *cluster* B e sete de nove amostras da África do Sul localizaram-se no *cluster* A.



**Figura 21. Análise filogenética do HPV-53.** Em adição às amostras analisadas neste trabalho (DF, em azul para facilitar a localização) estão presentes amostras dos estudos de Prado *et al* (2005) e Kocjan *et al* (2007). As siglas que antecedem o número das amostras indicam a procedência geográfica: São Paulo, Brasil (BR); Hong Kong (HK); Monte Rei, México (MX), Edimburgo, Escócia (ED); Cidade do Cabo, África do Sul (SA), Oklahoma, Estados Unidos (OK); Mali (ML); Eslovênia (SL) SL-1: amostras eslovenas 101, 158, 244, 293, 296, 336, 353, 412, 413, 596, 602, 681, 911, 933; SL-2: amostras eslovenas 115, 286, 328, 440, 583, 612, 652, 691, 746, 767, 904, 919, 922, 931, 938; SL-3: amostras eslovenas 86, 110, 297, 488, 489, 527, 603, 690, 753, 923. Árvore de máxima verossimilhança enraizada com HPV-30 como grupo externo. Os valores nos ramos correspondem aos valores de *bootstrap* calculados pelo método UPGMA e algoritmo LogDet e 1000 réplicas. São mostrados valores de *bootstrap* acima de 50%. O valor da barra de escala no canto esquerdo inferior de cada figura representa a distância genética em substituições por nucleotídeo do fragmento analisado. Em vermelho, a seqüência de referência.

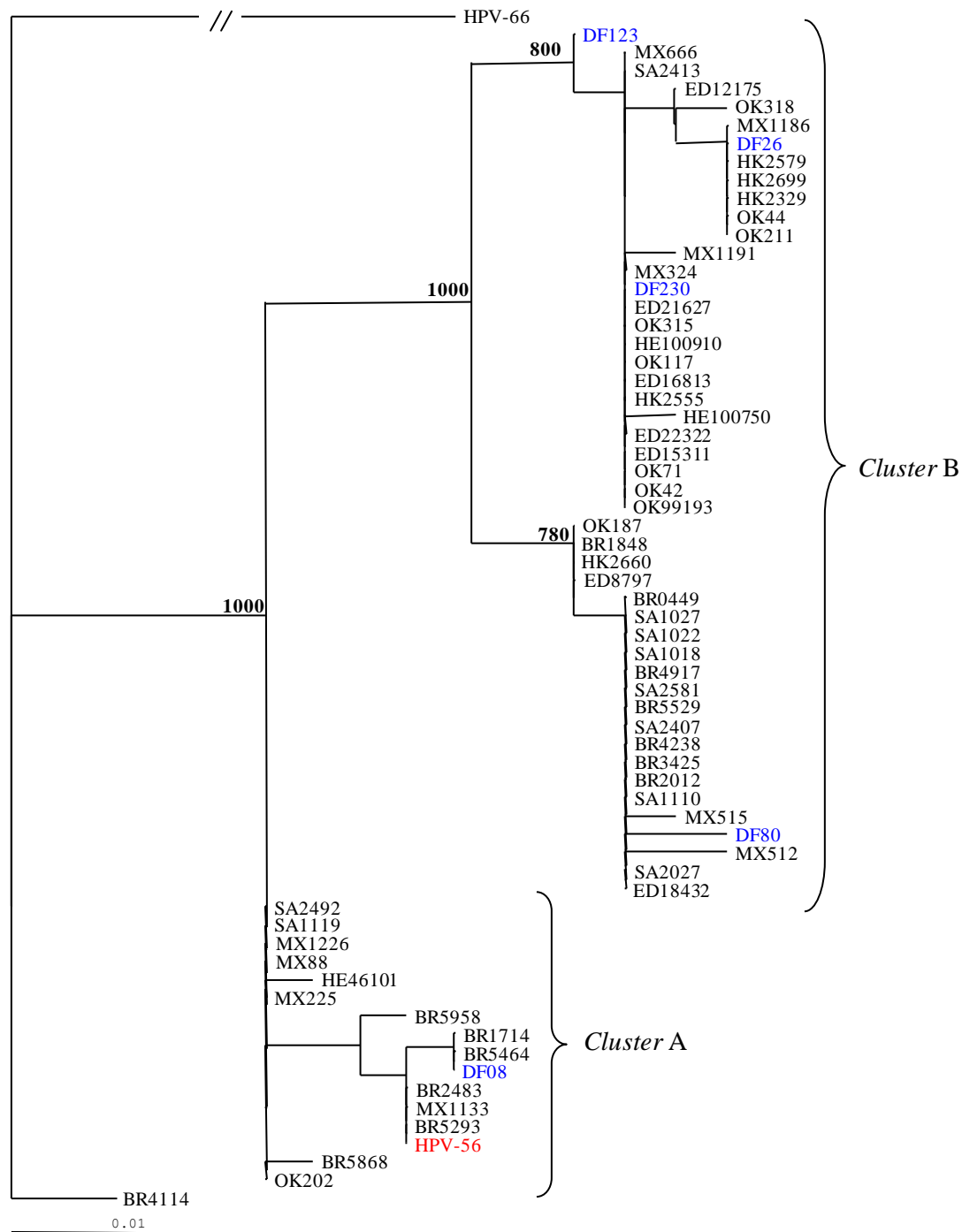
#### 4.4.2 HPV-56

Para a realização da análise filogenética da LCR do HPV-56, foram usadas seqüências de 514pb de 65 amostras. Destas, 19 foram provenientes do Brasil, sendo 5 do Distrito Federal (amostras caracterizados neste estudo) e 14 de São Paulo; 10 dos Estados Unidos; 10 do México; 7 da Escócia; 5 da China; 10 da África do Sul; 3 da Alemanha e a seqüência de referência do HPV-56.

Semelhante ao observado para o HPV-53, foi observada uma ramificação dicotômica na análise filogenética do HPV-56. A dicotomia foi marcada pela presença ou ausência do fragmento de 42pb entre as posições 7485 e 7526. Em um dos *clusters* (*cluster A*), contendo 16 amostras, incluindo a seqüência de referência do HPV-56, as amostras apresentaram o fragmento de 42pb e no outro *cluster* (*cluster B*) as amostras não apresentam esse fragmento (Figura 22). Das 5 amostras analisadas neste estudo, quatro se localizaram no *cluster B*, assim como a maioria das seqüências das amostras de todas as localidades geográficas e todas as amostras da China e da Escócia.

O *cluster B* apresentou aglomerados geográficos de amostras da China e Europa ou amostras provenientes de imigrantes presumidamente Europeus (amostras dos Estados Unidos) (Figura 22).

As amostras do DF não formaram um aglomerado distinto. A amostra DF08, que apresenta o fragmento de 42pb formou um pequeno aglomerado geográfico com 2 outras amostras do Brasil, apesar de o valor de *bootstrap* não ser significativo.



**Figura 22. Análise filogenética do HPV-56.** Em adição às amostras analisadas neste trabalho (DF, em azul para facilitar a localização) estão presentes amostras dos estudos de Prado *et al* (2005). As siglas que antecedem o número das amostras indicam a procedência geográfica: São Paulo, Brasil (BR); Hong Kong (HK); Monte Rei, México (MX), Edimburgo, Escócia (ED); Cidade do Cabo, África do Sul (SA) e Oklahoma, Estados Unidos (OK); Heidelberg, Alemanha (HE). Árvore de máxima verossimilhança enraizada com HPV-66 como grupo externo. Os valores nos ramos correspondem aos valores de *bootstrap* calculados pelo método UPGMA e algoritmo LogDet e 1000 réplicas. São mostrados valores de *bootstrap* acima de 50%. O valor da barra de escala no canto esquerdo inferior de cada figura representa a distância genética em substituições por nucleotídeo do fragmento analisado. Em vermelho, a seqüência de referência.

#### 4.4.3 HPV-66

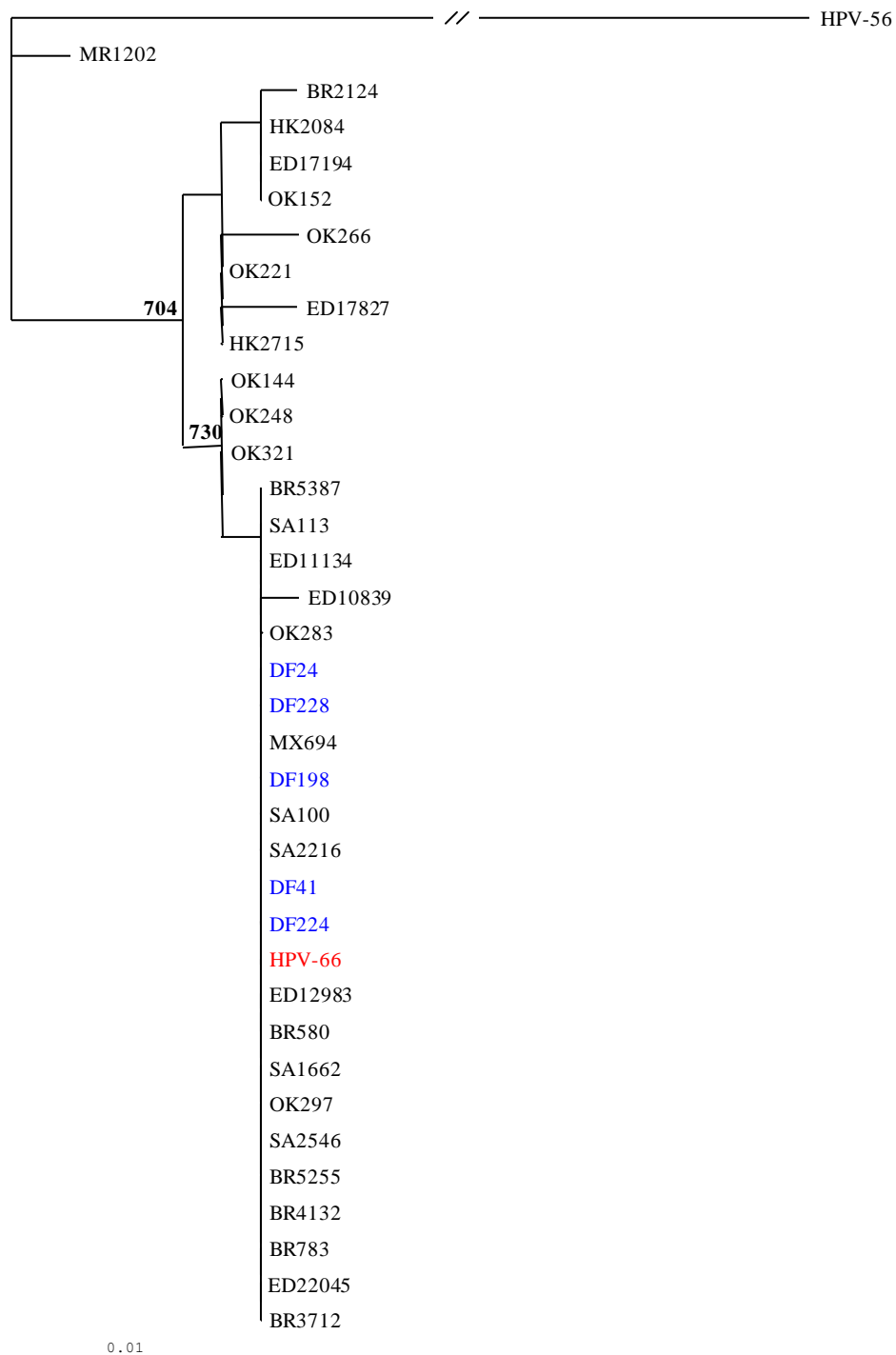
A análise filogenética da LCR do HPV-66 foi realizada utilizando um fragmento de 607pb de um total de 36 amostras, das quais, 12 foram provenientes do Brasil, sendo 5 do Distrito Federal (amostras caracterizados neste estudo) e 7 de São Paulo; 8 dos Estados Unidos; 1 do México; 6 da Escócia; 2 da China; 5 da África do Sul; 1 de Marrocos e a seqüência de referência do HPV-66.

Diferente do observado para o HPV-53 e HPV-56 não houve divisão dicotômica na análise filogenética do HPV-66.

Onze das 12 amostras provenientes do Brasil e todas as amostras do DF se agruparam no mesmo ramo da seqüência de referência HPV-66 (Figura 23). Esse agrupamento das amostras do DF com a seqüência de referência era esperado uma vez que todas as amostras apresentaram seqüência idêntica à seqüência de referência do HPV-66.

As duas amostras da China e seis de oito amostras dos Estados Unidos se situaram em ramos diferentes daquele que apresentou a seqüência de referência, enquanto a maioria das amostras das outras regiões geográficas se agruparam com a seqüência de referência (Figura 23).

O único aglomerado geográfico formado foi de três amostras dos Estados Unidos, OK144, OK248 e OK231 (Figura 23).



**Figura 23. Análise filogenética do HPV-66.** Em adição às amostras analisadas neste trabalho (DF, em azul para facilitar a localização) estão presentes amostras dos estudos de Prado *et al* (2005). As siglas que antecedem o número das amostras indicam a procedência geográfica: São Paulo, Brasil (BR); Hong Kong (HK); Monte Rei, México (MX), Edimburgo, Escócia (ED); Cidade do Cabo, África do Sul (SA) e Oklahoma, Estados Unidos (OK); Marrocos (MR). Árvore de máxima verossimilhança enraizada com HPV-56 como grupo externo. Os valores nos ramos correspondem aos valores de *bootstrap* calculados pelo método UPGMA e algoritmo LogDet e 1000 réplicas. São mostrados valores de *bootstrap* acima de 50%. O valor da barra de escala no canto esquerdo inferior de cada figura representa a distância genética em substituições por nucleotídeo do fragmento analisado. Em vermelho, a seqüência de referência.

## 5 DISCUSSÃO

Neste trabalho foi analisada pela primeira vez a variabilidade genética de três regiões genômicas, LCR, E6 e L1, dos HPVs definidos como sendo de alto risco -53, -56 e -66, detectados no Distrito Federal (DF) e entorno. Para isso, as amostras foram amplificadas, seqüenciadas e as seqüências geradas foram comparadas às seqüências de referência depositadas no GenBank (X74482 - HPVs-53, X74483 – HPV-56 e U31794 – HPV-66) para a análise de variações nucleotídicas e de aminoácidos. Além disso, foi feita a análise filogenética das amostras do DF juntamente com amostras de outras partes do mundo por meio do uso de seqüências depositadas no GenBank.

Para uma das amostras do HPV-66, o DF-107, não foi possível amplificar a região genômica LCR. Foram testados outros protocolos, alterando a concentração de  $MgCl_2$ , dos iniciadores, de DNA e da enzima *Taq* polimerase. Além disso, foram testados programas de amplificação com temperaturas de anelamento mais baixas e com maior número de ciclos. Todas as alterações feitas resultaram na não amplificação do fragmento esperado. Uma explicação possível para isso é que esta amostra apresenta variações na região de anelamento dos iniciadores impedindo a amplificação. Apesar da conservação observada para as amostras do HPV-66, esta amostra foi a única a apresentar variações na seqüência de E6 (Tabela 9) e a que mais apresentou variação na seqüência de L1 (onze trocas nucleotídicas). A outra amostra do HPV-66, DF-228, apresentou somente uma variação na posição 6877 de G para A, enquanto as outras três, DF-24, DF-41 e DF-198 não apresentaram nenhuma variação em relação à seqüência de referência (Tabela 10).

Duas amostras, uma do HPV-53 (DF-141) e uma do HPV-66 (DF-224), não puderam ser caracterizadas para o gene L1. Essas amostras apresentaram co-infecção, detectada pela análise de RFLP durante a genotipagem. Os iniciadores usados na amplificação de L1 (MY09/11) são degenerados e comuns a todos os genótipos do HPV. Assim, todas as seqüências de L1 geradas para essas duas amostras não corresponderam ao genótipo de interesse deste estudo. Mesmo a tentativa de separar os genótipos por meio de clonagem não foi bem sucedida, pois os genótipos detectados por meio da análise das seqüências dos clones selecionados não foram os de interesse para este estudo. Portanto, sugere-se que as amostras apresentavam um maior número de cópias dos outros genótipos presentes na co-infecção que os genótipos de interesse para este estudo.

A análise das seqüências obtidas revelou que o HPV-53 foi o genótipo com maior número de variações em todas as regiões, dentre os genótipos estudados. As amostras dos HPVs-53 e



-56 apresentaram maior variabilidade na LCR, seguida do gene E6 e L1. É esperado a LCR apresente maior variabilidade uma vez que é uma região não codificadora e que L1 apresente mais variações que E6, pois L1 codifica a principal proteína do capsídeo e, possivelmente, sofre pressão seletiva constante para o eficiente escape do sistema imune enquanto que o gene E6 codifica a oncoproteína E6, que tem um papel importante na infecção e patogenicidade viral tendo, portanto, mais restrições a acumular e tolerar mutações. Porém, o par de iniciadores MY09/11, usado para a detecção de todos os genótipos do HPV, amplifica um fragmento conservado de L1, o que poderia explicar a menor variabilidade observada para esse gene. Essa análise não pode ser feita com o HPV de genótipo 66 uma vez que a amostra que apresentou variações em E6 e L1 (DF-107) (Tabelas 9 e 10) não teve sua seqüência amplificada e seqüenciada para LCR.

As variações detectadas no gene E6 das amostras de HPV-53 foram descritas na literatura por Prado *et al* (2005) e Kocjan *et al* (2007) com exceção da variação silenciosa de G por A na posição 224, que está sendo descrita pela primeira vez. Na LCR, na posição 7648, em lugar da variação de T para G, relatada por Kocjan *et al* (2007), foi detectada a variação de T para A, ainda não descrita na literatura. A variação de T para G na posição 7686 também não está descrita na literatura. Em L1 as variações de A para G na posição 6717 e de C para T na posição 6942 estão sendo descritas pela primeira vez.

Entre as quatro substituições nucleotídicas encontradas nas amostras de E6 de HPV-56, duas foram descritas por Prado *et al* (2005). As dezessete variações detectadas na LCR estão sendo descritas pela primeira vez e a deleção de 42pb foi descrita por Prado *et al* (2005), bem como as duas substituições pontuais observadas em L1.

Em relação à análise das seqüências de E6 de HPV-66 as quatro variações detectadas foram descritas por Prado *et al* (2005). Na análise do gene L1 foram observadas doze variações, onze delas estão descritas na literatura (Cerqueira *et al*, 2003; Prado *et al*, 2005; Cerqueira *et al*, 2007).

Das nove trocas de aminoácidos observadas em E6 do HPV-53, três foram não conservativas, T52K, A94V, S104L. Em L1, foram observadas 2 trocas de aminoácidos, sendo uma conservativa, K355E, e outra não, P430S. Em E6 do HPV-56 foram detectadas 3 trocas de aminoácido, sendo duas não conservativas, S14R, K54N. Em L1 não houve trocas de aminoácidos. O HPV-66 foi o genótipo mais conservado, apresentando apenas a variação não conservativa S93G em E6.

Entre as amostras do HPV estudadas, foram detectadas substituições não silenciosas e algumas estavam associadas a trocas não conservativas de aminoácidos. No caso dessas trocas

não conservativas, há alteração das propriedades físico-químicas dos aminoácidos. No entanto, o efeito de cada mutação pontual na atividade da proteína depende da posição e função do aminoácido alterado, além da possível presença de outras variações no genoma. Mesmo que as trocas nucleotídicas provoquem mudanças na conformação das proteínas codificadas, possivelmente não afetam negativamente o desempenho das mesmas, uma vez que foram preservadas.

As regiões LCR dos HPV-53 e -56 apresentaram várias substituições nucleotídicas, algumas delas se encontram em potenciais sítios de ligação para fatores de transcrição e podem, portanto, influenciar a expressão dos oncogenes virais, que estão sob a regulação da LCR. Além disso, as inserções e deleções detectadas modificam a estrutura da fita de DNA, podendo afetar na interação dos fatores de transcrição com seus sítios de ligação localizados na LCR.

Foram detectadas variações em três potenciais sítios de ligação para fatores de transcrição: E2, um importante regulador viral da expressão dos oncogenes que pode funcionar como ativador ou repressor transcricional; TATAbox, região localizada no promotor onde se liga a proteína de ligação ao TATAbox e os fatores de transcrição formando o complexo responsável pelo início da transcrição e YY1, um fator de transcrição multifuncional que ativa ou reprime a transcrição dos genes virais. Essas variações podem, portanto, afetar a transcrição viral e a regulação da expressão dos oncogenes. Estudos para caracterizar as implicações das variações encontradas, tanto nos genes E6 e L1 quanto na LCR, são necessários para avaliar o real impacto que elas proporcionam na infecção viral.

A predição dos sítios de ligação para fatores de transcrição gerada pelo programa MATCH<sup>TM</sup> foi comparada às posições, já bem estabelecidas, dos sítios de ligação presentes no HPV-16 (O'Connor *et al*, 1995). O número de sítios de ligação para os fatores predita pelo MATCH<sup>TM</sup> foi condizente com a descrita para fatores como E2, NF-1, YY1 e Oct-1. No entanto, também foram encontradas divergências como, por exemplo, ausência de fatores como GRE e AP-1 nas LCRs dos HPVs-56 e -66 e presentes na LCR do HPV-16. As características mais conservadas na análise das LCRs são o posicionamento dos sítios de E2 dividindo a LCR nas porções 3', segmento central e 5' e a presença de sítios para NF-1 apenas na porção central da LCR, localizada entre o primeiro e segundo sítios de ligação para E2. Possivelmente, o posicionamento desses fatores represente aspectos importantes na biologia do HPV. É interessante notar que apenas na LCR do HPV-6, considerado um HPV de baixo risco não foi detectado sítio de ligação para c-Myc, fator envolvido no surgimento de diferentes tipos de tumor em humanos capaz de desregular a proliferação celular e inibir a

diferenciação celular. Porém, são necessários estudos que comprovem a existências dos sítios encontrados nas posições preditas.

Não foi estabelecida correlação entre os diferentes variantes detectados e diferentes graus de displasia apresentados pelas amostras analisadas. Curiosamente a amostra DF-107, que apresentou o maior número de variações (4 em E6 e 11 em L1), foi proveniente da única amostra com citologia normal dentre as amostras do HPV-66. Pode-se sugerir, portanto, que as variações nucleotídicas encontradas estariam diminuindo a patogenicidade desse variante ou que a pressão seletiva é intensa nesse indivíduo. Porém essa observação é apenas uma especulação devido ao pequeno número de amostras analisadas neste estudo.

A análise filogenética do HPV-53 confirmou a filogenia dicotômica descrita por Prado *et al* (2005) e Kocjan *et al* (2007). Essa dicotomia também foi observada para o HPV-56 e é similar à observada por Calleja-Macias *et al* (2005) ao analisar pares de subtipos (HPV-44/HPV-45 e HPV-68a/HPV-68b). É sugerido que os HPVs-53 e -56 consistem, de fato, em dois grupos de variantes genômicos distintos que ainda não atingiram distância genômica suficiente para serem classificados como subtipos distintos do HPV e que esses variantes estariam em um processo evolucionário que levaria ao surgimento de novos subtipos e, eventualmente, novos genótipos do HPV (Prado *et al*, 2005) (Kocjan *et al*, 2007).

As análises filogenéticas revelaram alguns aglomerados geográficos. No entanto, com exceção dos ramos formados exclusivamente por amostras eslovenas na análise das amostras do HPV-53 (Figura 21), nenhum outro aglomerado pode ser considerado um grupo de variantes étnicos distintos, como os observados nas análises do HPV-16 e HPV-18, pois, mesmo que os ramos tenham apresentado várias amostras provenientes de uma mesma região, também apresentaram amostras de outras regiões no mesmo ramo (Figuras 21 e 22). Esses agrupamentos de amostras provenientes da Eslovênia também não podem ser considerados um variante étnico, pois 58% das amostras usadas na análise filogenética eram provenientes da Eslovênia (Kocjan *et al*, 2007). São necessários estudos envolvendo mais amostras de todas as outras regiões geográficas para se afirmar que o variante que se repetiu nessas amostras eslovenas é de fato um variante ligado à etnia.

Na análise filogenética do HPV-66, observa-se que a seqüência de referência foi encontrada em várias amostras, tanto da África do Sul como da Escócia e Brasil (Figura 23). Especula-se que os HPVs se originaram nos primatas ancestrais dos humanos (Bernard, 1994). Dessa forma, essa conservação observada para as amostras do HPV-66 pode ser explicada se considerarmos que o genoma do HPV-66 tem uma evolução lenta quando comparado aos outros genótipos do HPV, tendo existido há mais de 100 mil anos (antes da

migração dos seres humanos pelos diferentes continentes) da mesma forma que existe hoje (Prado *et al*, 2005).

A pobre correlação entre variantes e origem etno-geográfica das pacientes foi observada para genótipos do HPV que são raros quando comparados ao HPV-16. Não se sabe o motivo de diferentes genótipos do HPV apresentarem prevalência distinta, mas é provável que genótipos mais abundantes se dispersem mais rápido que os genótipos raros. Com isso, a diversificação dos tipos raros do HPV ocorreria em uma escala de tempo diferente que a do HPV-16, o que não é consistente com a dispersão dos grupos étnicos no mundo (Prado *et al*, 2005).

## 6 CONCLUSÕES

Este foi o primeiro estudo sobre a variabilidade genética dos HPVs -53, -56 e -66 detectados em amostras do Brasil Central, contribuindo para expandir o conhecimento sobre a diversidade desses HPVs, sobre os quais existem poucos estudos publicados.

As conclusões desse trabalho foram:

- Existe variabilidade genética nos HPVs-53, -56 e -66;
- O HPV-66 foi o genótipo que apresentou menor variabilidade e o HPV-53 foi o que apresentou maior variabilidade;
- A região genômica mais variável foi a LCR;
- Na análise da LCR foram detectadas variações nos potenciais sítios de YY1, E2, Oct-1, TATA, AP-1 e NF-1 em amostras do HPV-53 e nos potenciais sítios de YY1 e E2 das amostras do HPV-56;
- Na análise das amostras dos HPVs-53 e -56, o gene E6 foi mais variável que o L1;
- A amostra DF-107, pertencente ao genótipo HPV-66, apresentou maior variabilidade em L1 que em E6;
- Foram detectadas variações nucleotídicas em E6 e L1 que levaram à troca de aminoácidos;
- Não foi detectada associação entre os variantes encontrados e os diferentes graus de displasias apresentados pelas amostras analisadas;
- Na análise filogenética não foi observado agrupamento étnico ou geográfico que pudesse ser correlacionado com os variantes detectados, como observado para HPV-16 e -18;
- Na análise filogenética, os HPVs-53 e -56 apresentaram ramificação dicotômica, anteriormente descrita apenas para subtipos;
- O HPV-66 foi o genótipo que apresentou maior conservação, com amostras de várias regiões se agrupando no mesmo ramo, juntamente com a sequência de referência.

## **7 PERSPECTIVAS**

Mais estudos devem ser conduzidos a fim de esclarecer o impacto das variações encontradas neste trabalho, visto que podem interferir na infectividade e patogenicidade viral.

O desenvolvimento de iniciadores ou sondas em métodos de detecção do HPV ou em pesquisa deve levar em consideração a distribuição dos variantes.

Vacinas contra o HPV estão em desenvolvimento e ainda não se sabe a extensão em que ocorre reatividade cruzada na resposta imune entre variantes intragenótipo do HPV. Dessa forma, o conhecimento sobre a variabilidade genética pode ser importante para o uso e a eficácia dessas vacinas. Além disso, é importante que haja uma vigilância constante dos genótipos do HPV que circulam na população, pois a vacinação pode levar a uma mudança na prevalência dos genótipos uma vez que a frequência dos HPVs prevenidos pelas vacinas irá diminuir, podendo abrir espaço para as infecções por outros genótipos.

## 8 BIBLIOGRAFIA

- Abba, M. C., Laguens, R. M., Dulout, F. N., & Golijow, C. D. (2004). The c-myc activation in cervical carcinomas and HPV 16 infections. *Mutat. Res.* , 557(2), 151-158.
- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic. Acids. Res.* , 25, 3389 - 3402.
- American Cancer Society. (2005). *Cancer facts and figures 2005*. Atlanta: American Cancer Society.
- Anderson, R. G. (1998). The caveolae membrane system. *Annu. Rev. Biochem.* , 67, 199-225.
- Apt, D., Chong, T., Liu, y., & Bernard, H. U. (1993). Nuclear factor I and epithelial cell-specific transcription of human papillomavirus type 16. *J. Virol.* , 67, 4455 - 4463.
- Baseman, J. G., & Koutsky, L. A. (2005). The epidemiology of human papillomavirus infections. *J. Clin. Virol.* , 32 Suppl 1, S16 - S24.
- Bauknecht, T., & Shi, Y. (1998). Overexpression of C/EBPbeta Represses Human Papillomavirus Type 18 Upstream Regulatory Region Activity in HeLa Cells by Interfering with the Binding of TATA-Binding Protein. *J. Virol.* , 72, 2113 - 2124.
- Bernard, H. U. (1994). Coevolution of papillomaviruses with human populations . *Trends in Microbiology* , 2, 140-143.
- Bernard, H. U. (2002). Gene expression of genital human papillomaviruses and considerations on potential antiviral approaches. *Antivir. Ther.* , 7(4), 219-237.
- Bernard, H. U., Chan, S. Y., Manos, M. M., Ong, C. K., Villa, L. L., Delius, H., et al. (1994). Identification and assessment of known and novel human papillomaviruses by polymerase chain reaction amplification, restriction fragment length polymorphisms, nucleotide sequence, and phylogenetic algorithms. *J Infect Dis* , 170(5), 1077-1085.
- Berumen, J., Ordoñez, R. M., Lazcano, E., Salmeron, J., Galvan, S. C., Estrada, R. A., et al. (2001). Asian-American Variants of Human Papillomavirus 16 and Risk for Cervical Cancer: a Case–Control Study. *J. Natl. Cancer. Inst.* , 93, 1325 - 1330.
- Bousarghin, L., Touzé, A., Sizaret, P. Y., & Coursaget, P. (2003). Human Papillomavirus Types 16, 31, and 58 Use Different Endocytosis Pathways To Enter Cells. *J. Virol.* , 77, 3846 - 3850.
- Brestovac, B., Harnett, G. B., Smith, D. W., Shellam, G. R., & Frost, F. A. (2005). Human papillomavirus genotypes and their association with cervical neoplasia in a cohort of Western Australian women. *J. Med. Virol.* , 76(1), 106 - 110.

- Brown, D. R., Bryan, J. T., Schroeder, J. M., Robinson, T. S., Fife, K. H., Wheeler, C. M., et al. (2001). Neutralization of human papillomavirus type 11 (HPV-11) by serum from women vaccinated with yeast-derived HPV-11 L1 virus-like particles: correlation with competitive radioimmunoassay titer. *J. Infect. Dis.* , 184(9), 1183-1186.
- Brown, D. R., Legge, D., & Qadadri, B. (2002). Distribution of human papillomavirus types in cervicovaginal washings from women evaluated in a sexually transmitted diseases clinic. *Sex. Transm. Dis.* , 29(12), 763-768.
- Brown, D. R., Schroeder, J. M., Bryan, J. T., Stoler, M. H., & Fife, K. H. (1999). Detection of Multiple Human Papillomavirus Types in Condylomata Acuminata Lesions from Otherwise Healthy and Immunosuppressed Patients. *J. Clin. Microbiol.* , 37, 3316 - 3322.
- Bryan, J. T., & Brown, D. R. (2000). Association of the human papillomavirus type 11 E1(E4) protein with cornified cell envelopes derived from infected genital epithelium. *Virology* , 277(2), 262-269.
- Burd, E. M. (2003). Human papillomavirus and cervical cancer. *Clin. Microb. Rev.* , 16(1), 1-17.
- Burk, R. D. (1999). Human papillomavirus and the risk of cervical cancer. *Hosp. Pract.* , 34, 103-111.
- Calleja-Macias, I. E., Kalantari, M., Allan, B., Williamson, A. L., Chung, L. P., Collins, R. J., et al. (2005). Papillomavirus Subtypes Are Natural and Old Taxa: Phylogeny of Human Papillomavirus Types 44 and 55 and 68a and -b. *J. Virol.* , 79, 6565 - 6569.
- Camara, G. N., Cerqueira, D. M., Oliveira, A. P., Silva, E. O., Carvalho, L. G., & Martins, C. R. (2003). Prevalence of human papillomavirus types in women with pre-neoplastic and neoplastic cervical lesions in the Federal District of Brazil. *Mem. Inst. Oswaldo Cruz* , 98(7), 879 - 883.
- Castellsague, X., Menendez, C., Loscertales, M. P., Kornegay, J. R., dos Santos, F., Gomez-Olive, F. X., et al. (2001). Human papillomavirus genotypes in rural Mozambique. *Lancet* , 358(9291), 1429-1430.
- Cerqueira, D. M., Camara, G. N., Cruz, M. R., Silva, E. O., Brígido, M. M., Carvalho, L. G., et al. (2003). Variants of Human Papillomavirus Types 53, 58 and 66 Identified in Central Brazil. *Virus Genes* , 26(1), 83 - 87.
- Cerqueira, D. M., Moraes, D. S., Camara, G. N., Amaral, F. A., Oyama, C. N., dos Santos, M. Q., et al. (2007). High HPV genetic diversity in women infected with HIV-1 in Brazil. *Arch. Virol.* , 152(1), 75 - 83.



- Chan, P. K., Chang, A. R., Tam, W. H., Cheung, J. L., & Cheng, A. F. (2002). Prevalence and genotype distribution of cervical human papillomavirus infection: Comparison between pregnant women and non-pregnant controls. *J. Med. Virol.* , 67(4), 583-588.
- Chan, W. K., Klock, G., & Bernard, H. U. (1989). Progesterone and glucocorticoid response elements occur in the long control regions of several human papillomaviruses involved in anogenital neoplasia. *J. Virol.* , 63, 3261 - 3269.
- Clifford, G. M., Gallus, S., Herrero, R., Munoz, N., Snijders, P. J., Vaccarella, S., et al. (2005). Worldwide distribution of human papillomavirus types in cytologically normal women in the International Agency for Research on Cancer HPV prevalence surveys: a pooled analysis. *Lancet* , 366(9490), 991-998.
- Cripe, T. P., Haugen, T. H., Turk, J. P., Tabatabai, F., Schmid, P. G., Durst, M., et al. (1987). Transcriptional regulation of the human papillomavirus-16 E6-E7 promoter by a keratinocyte-dependent enhancer, and by viral E2 trans-activator and repressor gene products: implications for cervical carcinogenesis. *EMBO J* , 6(12), 3745-3753.
- Cuschieri, K. S., & Cubie, H. A. (2005). The role of human papillomavirus testing in cervical screening. *J. Clin. Virol.* , 32 Suppl 1, S34 - S42.
- Day, P. M., Baker, C. C., Lowy, D. R., & Schiller, J. T. (2004). Establishment of papillomavirus infection is enhanced by promyelocytic leukemia protein (PML) expression. *PNAS* , 101, 14252 - 14257.
- de Boer, M. A., Peters, L. A., Aziz, M. F., Siregar, B., Cornain, S., Vrede, M. A., et al. (2004). Human papillomavirus type 16 E6, E7, and L1 variants in cervical cancer in Indonesia, Suriname, and The Netherlands. *Gynecol. Oncol.* , 94(2), 488 - 494.
- de Villiers, E. M., Fauquet, C., Broker, T. R., Bernard, H. U., & zur Hausen, H. (2004). Classification of papillomaviruses. *Virology* , 324(1), 17-27.
- Disbrow, G. L., Hanover, J. A., & Schlegel, R. (2005). Endoplasmic Reticulum-Localized Human Papillomavirus Type 16 E5 Protein Alters Endosomal pH but Not trans-Golgi pH. *J. Virol.* , 79, 5839 - 5846.
- Dong, G., Broker, T. R., & Chow, L. T. (1994). Human papillomavirus type 11 E2 proteins repress the homologous E6 promoter by interfering with the binding of host transcription factors to adjacent elements. *J. Virol.* , 68, 1115 - 1127.
- Dong, X. P., & Pfister, H. (1999). Overlapping YY1- and aberrant SP1-binding sites proximal to the early promoter of human papillomavirus type 16. *J. Gen. Virol.* , 80, 2097 - 2101.
- Doorbar, J. (2006). Molecular biology of human papillomavirus infection and cervical cancer. *Clin Sci (Lond)* , 110(5), 525-541.

- Doorbar, J. (2005). The papillomavirus life cycle. *J Clin Virol* , 32 Suppl 1, S7-S15.
- Durst, M., Glitz, D., Schneider, A., & zur Hausen, H. (1992). Human papillomavirus type 16 (HPV 16) gene expression and DNA replication in cervical neoplasia: analysis by in situ hybridization. *Virology* , 189(1), 132-140.
- Evander, M., Frazer, I. H., Payne, E., Qi, Y. M., Hengst, K., & McMillan, N. A. (1997). Identification of the alpha6 integrin as a candidate receptor for papillomaviruses. *J. Virol.* , 71, 2449 - 2456.
- Evans, T. G., Bonnez, W., Rose, R. C., Koenig, S., Demeter, L., Suzich, J. A., et al. (2001). A Phase 1 study of a recombinant viruslike particle vaccine against human papillomavirus type 11 in healthy adult volunteers. *J. Infect. Dis.* , 183(10), 1485-1493.
- Fauquet, C. M., Mayo, M. A., Maniloff, J., Desselberger, U., & Ball, L. A. (2005). *Virus Taxonomy: The Eighth Report of the International Committee on Taxonomy of Viruses.* (Elsevier, Ed.) Academic Press.
- Felsenstein, J. (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.* , 17(6), 368 - 376.
- Florin, L., Sapp, C., Streeck, R. E., & Sapp, M. (2002). Assembly and Translocation of Papillomavirus Capsid Proteins. *J. Virol.* , 76, 10009 - 10014.
- Franco, E. L., & Harper, D. M. (2005). Vaccination against human papillomavirus infection: a new paradigm in cervical cancer control. *Vaccine* , 23(17-18), 2388-2394.
- Franco, E. L., Villa, L. L., Sobrinho, J. P., Prado, J. M., Rousseau, M. C., Desy, M., et al. (1999). Epidemiology of acquisition and clearance of cervical human papillomavirus infection in women from a high-risk area for cervical cancer. *J Infect Dis* , 180(5), 1415 - 1423.
- Gallahan, D., Müller, M., Schneider, A., Delius, H., Kahn, T., de Villiers, E. M., et al. (1989). Human papillomavirus type 53. *J. Virol.* , 63, 4911 - 4912.
- Galvin, K. M., & Shi, Y. (1997). Multiple mechanisms of transcriptional repression by YY1. *Mol. Cell. Biol.* , 17, 3723 - 3732.
- Genther, S. M., Sterling, S., Duensing, S., Münger, K., Sattler, C., & Lambert, P. F. (2003). Quantitative Role of the Human Papillomavirus Type 16 E5 Gene during the Productive Stage of the Viral Life Cycle. *J. Virol.* , 77, 2832 - 2842.
- Gloss, B., & Bernard, H. U. (1990). The E6/E7 promoter of human papillomavirus type 16 is activated in the absence of E2 proteins by a sequence-aberrant Sp1 distal element. *J. Virol.* , 64, 5577 - 5584.

- Gloss, B., Bernard, H. U., Seedorf, K., & Klock, G. (1987). The upstream regulatory region of the human papilloma virus-16 contains an E2 protein-independent enhancer which is specific for cervical carcinoma cells and regulated by glucocorticoid hormones. *EMBO J* , 6(12), 3735-3743.
- Gonçalves, M. A., Massad, E., Burattini, M. N., & Villa, L. L. (1999). Relationship between human papillomavirus (HPV) genotyping and genital neoplasia in HIV-positive patients of Santos City, Sao Paulo, Brazil. *International Journal of STD & AIDS* , 10, 803 - 807.
- Gravitt, P. E., Peyton, C. L., Alessi, T. Q., Wheeler, C. M., Coutlée, F., Hildesheim, A., et al. (2000). Improved Amplification of Genital Human Papillomaviruses. *J. Clin. Microbiol.* , 38, 357 - 361.
- Harper, D. M., Franco, E. L., Wheeler, C. M., Moscicki, A. B., Romanowski, B., Roteli-Martins, C. M., et al. (2006). Sustained efficacy up to 4.5 years of a bivalent L1 virus-like particle vaccine against human papillomavirus types 16 and 18: follow-up from a randomised control trial. *Lancet* , 367(9518), 1247-1255.
- Harper, D. M., Franco, E. L., Wheeler, C., Ferris, D. G., Jenkins, D., Schuind, A., et al. (2004). Efficacy of a bivalent L1 virus-like particle vaccine in prevention of infection with human papillomavirus types 16 and 18 in young women: a randomised controlled trial. *Lancet* , 364(9447), 1757-1765.
- Harro, C. D., Pang, Y. Y., Roden, R. B., Hildesheim, A., Wang, Z., Reynolds, M. J., et al. (2001). Safety and Immunogenicity Trial in Adult Volunteers of a Human Papillomavirus 16 L1 Virus-Like Particle Vaccine. *J. Natl. Cancer. Inst.* , 93, 284 - 292.
- Hebner, C. M., & Laimins, L. A. (2006). Human papillomaviruses: basic mechanisms of pathogenesis and oncogenicity. *Rev. Med. Virol.* , 16(2), 83-97.
- Hines, C. S., Meghoo, C., Shetty, S., Biburger, M., Brenowitz, M., & Hegde, R. S. (1998). DNA structure and flexibility in the sequence-specific binding of papillomavirus E2 proteins. *J. Mol. Biol.* , 276(4), 809-818.
- Ho, L., Chan, S. Y., Burk, R. D., Das, B. C., Fujinaga, K., Icenogle, J. P., et al. (1993). The genetic drift of human papillomavirus type 16 is a means of reconstructing prehistoric viral spread and the movement of ancient human populations. *J. Virol.* , 67, 6413 - 6423.
- Holmgren, S. C., Patterson, N. A., Ozbun, M. A., & Lambert, P. F. (2005). The Minor Capsid Protein L2 Contributes to Two Steps in the Human Papillomavirus Type 31 Life Cycle. *J. Virol.* , 79, 3938 - 3948.

- Huibregtse, J. M., Scheffner, M., & Howley, P. M. (1991). A cellular protein mediates association of p53 with the E6 oncoprotein of human papillomavirus types 16 or 18. *EMBO J*, *10*(13), 4129-4135.
- Ishiji, T., Lace, M. J., Parkkinen, S., Anderson, R. D., Haugen, T. H., Cripe, T. P., et al. (1992). Transcriptional enhancer factor (TEF)-1 and its cell-specific co-activator activate human papillomavirus-16 E6 and E7 oncogene transcription in keratinocytes and cervical carcinoma cells. *EMBO J*, *11*(6), 2271-2281.
- Jo, H., & Kim, J. W. (2005). Implications of HPV infection in uterine cervical cancer. *Cancer Therapy*, *3*, 419-434.
- Joyce, J. G., Tung, J. S., Przysiecki, C. T., Cook, J. C., Lehman, E. D., Sands, J. A., et al. (1999). The L1 Major Capsid Protein of Human Papillomavirus Type 11 Recombinant Virus-like Particles Interacts with Heparin and Cell-surface Glycosaminoglycans on Human Keratinocytes. *J. Biol. Chem.*, *274*, 5810.
- Kämmer, C., Tommasino, M., Syrjanen, S., Delius, H., Hebling, U., Warthorst, U., et al. (2002). Variants of the long control region and the E6 oncogene in European human papillomavirus type 16 isolates: implications for cervical disease. *Br. J. Cancer.*, *86*(2), 269 - 273.
- Kämmer, C., Warthorst, U., Torrez-Martinez, N., Wheeler, C. M., & Pfister, H. (2000). Sequence analysis of the long control region of human papillomavirus type 16 variants and functional consequences for P97 promoter activity. *J. Gen. Virol.*, *81*, 1975 - 1981.
- Kel, A. E., Gößling, E., Reuter, I., Cheremushkin, E., Kel-Margoulis, O. V., & Wingender, E. (2003). MATCH: a tool for searching transcription factor binding sites in DNA sequences. *Nucleic. Acids. Res.*, *31*, 3576 - 3579.
- Kisseljov, F. L. (2000). Virus-associated human tumors: cervical carcinomas and papilloma viruses. *Biochemistry (Mosc)*, *65*(1), 68-77.
- Kocjan, B. J., Seme, K., Mocilnik, T., Jancar, N., Vrtacnik-Bokal, E., & Poljak, M. (2007). Genomic diversity of human papillomavirus genotype 53 in an ethnogeographically closed cohort of white European women. *J. Med. Virol.*, *79*(4), 431 - 438.
- Koutsky, L. A. (s.d.). A Controlled Trial of a Human Papillomavirus Type 16 Vaccine.
- Lechner, M. S., & Laimins, L. A. (1994). Inhibition of p53 DNA binding by human papillomavirus E6 proteins. *J. Virol.*, *68*, 4262 - 4273.
- Lee, G. Y., Kim, S. M., Rim, S. Y., Choi, H. S., Park, C. S., & Nam, J. H. (2005). Human papillomavirus (HPV) genotyping by HPV DNA chip in cervical cancer and precancerous lesions. *Int. J. Gynecol. Cancer*, *15*(1), 81-87.

- Levi, J. E., Kleter, B., Quint, W. G., Fink, M. C., Canto, C. L., Matsubara, R., et al. (2002). High Prevalence of Human Papillomavirus (HPV) Infections and High Frequency of Multiple HPV Genotypes in Human Immunodeficiency Virus-Infected Women in Brazil. *J. Clin. Microbiol.* , 40, 3341 - 3345.
- Lörincz, A. T., Quinn, A. P., Goldsborough, M. D., McAllister, P., & Temple, G. F. (1989). Human Papillomavirus Type 56: a New Virus Detected in Cervical Cancers. *J. Gen. Virol.* , 70, 3099 - 3104.
- Lowy, D. R., & Schiller, J. T. (2006). Prophylactic human papillomavirus vaccines. *J. Clin. Invest.* , 116, 1167 - 1173.
- Luque, A. E., Messing, S., Rose, R. C., & Reichman, R. C. (2004). Prevalence of Human Papillomavirus (HPV) Genotypes in HIV-Infected Women. *Int. Conf. AIDS* , 15, Abstract nº MoPeB3323.
- Mandic, A., & Vujkov, T. (2004). Human papillomavirus vaccine as a new way of preventing cervical cancer: a dream or the future? *Ann. Onc.* , 15, 197 - 200.
- Manos, M. M., Ting, Y., Wright, D. K., Lewis, A. J., Broker, T. R., & Wolinsky, S. M. (1989). Use of polymerase chain reaction amplification for the detection of genital human papillomaviruses. *Cancer Cells* , 7, 209-214.
- McBride, A. A., Romanczuk, H., & Howley, P. M. (1991). The papillomavirus E2 regulatory proteins. *J. Biol. Chem.* , 266, 18411 - 18414.
- McPhillips, M. G., Ozato, K., & McBride, A. A. (2005). Interaction of Bovine Papillomavirus E2 Protein with Brd4 Stabilizes Its Association with Chromatin. *J. Virol.* , 79, 8920 - 8932.
- Merk. (julho de 2006). *Merk Disease Modules*. Acesso em 28 de março de 2007, disponível em <http://www.merckmedicus.com/ppdocs/us/hcp/diseasemodules/hpvd/images/figure21.jpg>
- Merk. (2006). *Merk Medicus Modules: HPV Disease - Natural History*. Acesso em 28 de março de 2007, disponível em <http://www.merckmedicus.com/pp/us/hcp/diseasemodules/hpvd/natural-history.jsp>
- Ministério da Saúde. (2005). *Estimativa 2006: Incidência de câncer no Brasil*. Acesso em 23 de Abril de 2007, disponível em Site do INCA: <http://www.inca.gov.br/estimativa/2006/versaofinal.pdf>
- Ministério da Saúde. (2003). *Lista Nacional de Agravos de Notificação Compulsória*. Acesso em 23 de Abril de 2007, disponível em Site do Ministério da Saúde: <http://portal.saude.gov.br/portal/arquivos/pdf/portaria2325.pdf>

- Modis, Y., Trus, B. L., & Harrison, S. C. (2002). Atomic model of the papillomavirus capsid. *EMBO J*, *21*(18), 4754-4762.
- Mohan, S., & Ind, T. (2004). Cervical screening in England and Wales: an update. *Curr Opin Obstet Gynecol*, *16*(6), 491 - 496.
- Münger, K., Baldwin, A., Edwards, K. M., Hayakawa, H., Nguyen, C. L., Owens, M., et al. (2004). Mechanisms of human papillomavirus-induced oncogenesis. *J. Virol.*, *78*, 11451-11460.
- Muñoz, N., Bosch, F. X., Sanjosé, S., Herrero, R., Castellsagué, X., Shah, K. V., et al. (2003). Epidemiologic Classification of Human Papillomavirus Types Associated with Cervical Cancer. *N. Engl. J. Med.*, *348*, 518 - 527.
- Munoz, N., Castellsague, X., de Gonzalez, A. B., & Gissmann, L. (2006). HPV in the etiology of human cancer. *Vaccine*, *24 S3*, S1-S10.
- Myers, G., Bernard, H. U., Delius, H., Favre, M., Iconogle, J., van Ranst, M., et al. (1994). *Human papillomaviruses 1994 compendium*. Novo México: Los Alamos National Laboratory.
- Nakahara, T., Peh, W. L., Doorbar, J., Lee, D., & Lambert, P. F. (2005). Human Papillomavirus Type 16 E1{wedge}E4 Contributes to Multiple Facets of the Papillomavirus Life Cycle. *J. Virol.*, *79*, 13150 - 13165.
- Nindl, I., Rindfleisch, K., Lotz, B., Schneider, A., & Durst, M. (1999). Uniform distribution of HPV 16 E6 and E7 variants in patients with normal histology, cervical intra-epithelial neoplasia and cervical cancer. *Int. J. Cancer*, *82*(2), 203 - 207.
- Nucleus Medical Art. (1999-2007). *Nucleus Medical Art*. Acesso em 19 de abril de 2007, disponível em Cervical cancer - Medical Illustration: <http://catalog.nucleusinc.com/generateexhibit.php?ID=4781&ExhibitKeywordsRaw=cervical+cancer&TL=1793&A=2>
- O'Connor, M. J., Tan, S. H., Tan, C. H., & Bernard, H. U. (1996). YY1 represses human papillomavirus type 16 transcription by quenching AP-1 activity. *J. Virol.*, *70*, 6529 - 6539.
- O'Connor, M., & Bernard, H. U. (1995). Oct-1 activates the epithelial-specific enhancer of human papillomavirus type 16 via a synergistic interaction with NFI at a conserved composite regulatory element. *Virology*, *207*(1), 77-88.
- O'Connor, M., Chan, S. Y., & Bernard, H. U. (1995). Transcription Factor Binding Sites in the Long Control Region of Genital HPVs. In: *Human Papillomaviruses 1995 Compendium*.

- Offord, E. A., & Beard, P. (1990). A member of the activator protein 1 family found in keratinocytes but not in fibroblasts required for transcription from a human papillomavirus type 18 promoter. *J. Virol.* , 64, 4792 - 4798.
- Ong, C. K., Chan, S. Y., Campo, M. S., Fujinaga, K., Mavromara-Nazos, P., Labropoulou, V., et al. (1993). Evolution of human papillomavirus type 18: an ancient phylogenetic root in Africa and intratype diversity reflect coevolution with human ethnic groups. *J. Virol.* , 67, 6424 - 6431.
- Patterson, N. A., Smith, J. L., & Ozbun, M. A. (2005). Human Papillomavirus Type 31b Infection of Human Keratinocytes Does Not Require Heparan Sulfate. *J. Virol.* , 79, 6838 - 6847.
- Prado, J. C., Calleja-Macias, I. E., Bernard, H. U., Kalantari, M., Macay, S. A., Allan, B., et al. (2005). Worldwide genomic diversity of the human papillomaviruses-53, 56, and 66, a group of high-risk HPVs unrelated to HPV-16 and HPV-18. *Virology* , 340(1), 95-104.
- Reddy, K. J., Banapour, B., Anderson, D. E., Lee, S. H., Marquez, J. P., Carlos, M. P., et al. (2004). Induction of immune responses against human papillomaviruses by hypervariable epitope constructs. *Immunology* , 112(2), 321-327.
- Richardson, H., Franco, E., Pintos, J., Bergeron, J., Arella, M., & Tellier, P. (2000). Determinants of low-risk and high-risk cervical human papillomavirus infections in Montreal University students. *Sex. Transm. Dis.* , 27(2), 79 - 86.
- Roden, R. B., Ling, M., & Wu, T. C. (2004). Vaccination to prevent and treat cervical cancer. *Hum. Pathol.* , 35(8), 971 - 982.
- Rose, B., Steger, G., Dong, X. P., Thompson, C., Cossart, Y., Tattersall, M., et al. (1998). Point mutations in SP1 motifs in the upstream regulatory region of human papillomavirus type 18 isolates from cervical cancers increase promoter activity. *J. Gen. Virol.* , 79, 1659 - 1663.
- Sarafi, T. R., & McBride, A. A. (1995). Domains of the BPV-1 E1 replication protein required for origin-specific DNA binding and interaction with the E2 transactivator. *Virology* , 211(2), 385-396.
- Scheurer, M. E., Tortolero-Luna, G., & Adler-Storthz, K. (2005). Human papillomavirus infection: biology, epidemiology, and prevention. *Int J Gynecol Cancer* , 15(5), 727-746.
- Schiffman, M. H., & Brinton, L. A. (1995). The epidemiology of cervical carcinogenesis. *Cancer* , 76(10 Suppl), 1888-1901.

- Schmidt, M., Kedzia, W., & Gozdzicka-Jozefiak, A. (2001). Intratype HPV16 sequence variation within LCR of isolates from asymptomatic carriers and cervical cancers. *J. Clin. Virol.* , 23(1-2), 65 - 77.
- Schneider, A. (1994). Natural history of genital papillomavirus infections. *Intervirology* , 37(3-4), 201-214.
- Shafti-Keramat, S., Handisurya, A., Kriehuber, E., Meneguzzi, G., Slupetzky, K., & Kirnbauer, R. (2003). Different Heparan Sulfate Proteoglycans Serve as Cellular Receptors for Human Papillomaviruses. *J. Virol.* , 77, 13125 - 13135.
- Sherman, L., Jackman, A., Itzhaki, A., Stoppler, M. C., Koval, D., & Schlegel, R. (1997). Inhibition of serum- and calcium-induced differentiation of human keratinocytes by HPV16 E6 oncoprotein: role of p53 inactivation. *Virology* , 37(2), 296-306.
- Sibbet, G. J., & Campo, M. S. (1990). Multiple interactions between cellular factors and the non-coding region of human papillomavirus type 16. *J. Gen. Virol.* , 71, 2699 - 2707.
- Stannard, L. M. (1995). *Electron Micrograph Images*. Acesso em 27 de abril de 2007, disponível em Virus Ultra Structure: <http://web.uct.ac.za/depts/mmi/stannard/papillo.html>
- Steenbergen, R. D., de Wilde, J., Wilting, S. M., Brink, A. A., Snijders, P. J., & Meijer, C. J. (2005). HPV-mediated transformation of the anogenital tract. *J. Clin. Virol.* , 32 Suppl 1, S25 - S33.
- Sterling, J. C., & Tyring, S. K. (2001). *Human papillomaviruses: clinical and scientific advances* (1ª ed.). New York: Oxford University.
- Stevens, M. P., Tabrizi, S. N., Quinn, M. A., & Garland, S. M. (2006). Human papillomavirus genotype prevalence in cervical biopsies from women diagnosed with cervical intraepithelial neoplasia or cervical cancer in Melbourne, Australia. *Int. J. Gynecol. Cancer* , 16(3), 1017 - 1024.
- Stoler, M. H., Rhodes, C. R., Whitbeck, A., Wolinsky, S. M., Chow, L. T., & Broker, T. R. (1992). Human papillomavirus type 16 and 18 gene expression in cervical neoplasias. *Hum. Pathol.* , 23(2), 117-128.
- Straight, S. W., Hinkle, P. M., Jewers, R. J., & McCance, D. J. (1993). The E5 oncoprotein of human papillomavirus type 16 transforms fibroblasts and effects the downregulation of the epidermal growth factor receptor in keratinocytes. *J. Virol.* , 67, 4521 - 4532.
- Subramanyam, D., & Krishna, S. (2006). c-Myc substitutes for Notch1-CBF1 functions in cooperative transformation with papillomavirus oncogenes. *Virology* , 347(1), 191-198.



- Tan, S. H., Gloss, B., & Bernard, H. U. (1992). During negative regulation of the human papillomavirus-16E6 promoter, the viral E2 protein can displace Sp1 from a proximal promoter element. *Nucleic Acids Res.* , 20, 251 - 256.
- Tan, S. H., Leong, L. E., Walker, P. A., & Bernard, H. U. (1994). The human papillomavirus type 16 E2 transcription factor binds with low cooperativity to two flanking sites and represses the E6 promoter through displacement of Sp1 and TFIID. *J. Virol.* , 68, 6411 - 6420.
- Tang, S., Tao, M., McCoy, J. P., & Zheng, Z. M. (2006). The E7 Oncoprotein Is Translated from Spliced E6\*I Transcripts in High-Risk Human Papillomavirus Type 16- or Type 18-Positive Cervical Cancer Cell Lines via Translation Reinitiation. *J. Virol.* , 80, 4249 - 4263.
- Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic. Acids. Res.* , 22, 4673 - 4680.
- Trus, B. L., Roden, R. B., Greenstone, H. L., Vrhel, M., Schiller, J. T., & Booy, F. P. (1997). Novel structural features of bovine papillomavirus capsid revealed by a three-dimensional reconstruction to 9 Å resolution. *Nat. Struct. Biol.* , 4(5), 413-420.
- Turek, L. P., & Smith, E. M. (1996). The genetic program of genital human papillomaviruses in infection and cancer. *Obstet. Gynecol. Clin. North. Am.* , 23(4), 735-758.
- Veress, G., Szarka, K., Dong, X. P., Gergely, L., & Pfister, H. (1999). Functional significance of sequence variation in the E2 gene and the long control region of human papillomavirus type 16. *J. Gen. Virol.* , 80, 1035 - 1043.
- Villa, L. L., Costa, R. L., Petta, C. A., Andrade, R. P., Ault, K. A., Giuliano, A. R., et al. (2005). Prophylactic quadrivalent human papillomavirus (types 6, 11, 16, and 18) L1 virus-like particle vaccine in young women: a randomised double-blind placebo-controlled multicentre phase II efficacy trial. *Lancet Oncol* , 6(5), 271-278.
- Villa, L. L., Sichero, L., Rahal, P., Caballero, O., Ferenczy, A., Rohan, T., et al. (2000). Molecular variants of human papillomavirus types 16 and 18 preferentially associated with cervical neoplasia. *J. Gen. Virol.* , 81, 2959 - 2968.
- Wang, Q., Griffin, H., Southern, S., Jackson, D., Martin, A., McIntosh, P., et al. (2004). Functional Analysis of the Human Papillomavirus Type 16 E1{wedge}E4 Protein Provides a Mechanism for In Vivo and In Vitro Keratin Filament Reorganization. *J. Virol.* , 78, 821 - 833.

- Wilson, R., Fehrman, F., & Laimins, L. A. (2005). Role of the E1{wedge}E4 Protein in the Differentiation-Dependent Life Cycle of Human Papillomavirus Type 31. *J. Virol.* , 79, 6732 - 6740.
- Wilson, V. G., West, M., Woytek, K., & Rangasamy, D. (2002). Papillomavirus E1 proteins: form, function, and features. *Virus Genes* , 24(3), 275-290.
- Wingender, E., Dietze, P., Karas, H., & Knuppel, R. (1996). TRANSFAC: a database on transcription factors and their DNA binding sites. *Nucleic. Acids. Res.* , 24, 238 - 241.
- Woodman, C. B., Collins, S. I., & Young, L. S. (2007). The natural history of cervical HPV infection: unresolved issues. *Nat Rev Cancer* , 7(1), 11-22.
- Xi, L. F., Koutsky, L. A., Galloway, D. A., Kuypers, J., Hughes, J. P., Wheeler, C. M., et al. (1997). Genomic variation of human papillomavirus type 16 and risk for high grade cervical intraepithelial neoplasia. *J. Natl. Cancer. Inst.* , 89, 796 - 802.
- Yamada, T., Manos, M. M., Peto, J., Greer, C. E., Munoz, N., Bosch, F. X., et al. (1997). Human papillomavirus type 16 sequence variation in cervical cancers: a worldwide perspective. *J. Virol.* , 71, 2463 - 2472.
- You, J., Croyle, J. L., Nishimura, A., Ozato, K., & Howley, P. M. (2004). Interaction of the bovine papillomavirus E2 protein with Brd4 tethers the viral DNA to host mitotic chromosomes. *Cell* , 117(3), 349-360.
- Yu, T., Ferber, M. J., Cheung, T. H., Chung, T. K., Wong, Y. F., & Smith, D. I. (2005). The role of viral integration in the development of cervical cancer. *Cancer Genet Cytogenet* , 158(1), 27-34.
- Zhao, K. N., Hengst, K., Liu, W. J., Liu, Y. H., Liu, X. S., McMillan, N. A., et al. (2000). BPV1 E2 Protein Enhances Packaging of Full-Length Plasmid DNA in BPV1 Pseudovirions. *Virology* , 272, 382 - 393.
- Zhou, J., Sun, X. Y., Stenzel, D. J., & Frazer, I. H. (1991). Expression of vaccinia recombinant HPV 16 L1 and L2 ORF proteins in epithelial cells is sufficient for assembly of HPV virion-like particles. *Virology* , 185(1), 251 - 257.
- zur Hausen, H. (1991). Human papillomaviruses in the pathogenesis of anogenital cancer. *184*, 9-13.

## 9 ANEXO I

### HPV-53 E6

DF-33

102

ATGGATCGCCAGTTATTTGAAAATACAGAGGAGCGACCACGTACATTGCATCAGCTATGTGAAGTT  
GTGAATAAACCATTGCTGGAGCTGCAACTTGGCTGTGTGTTCTGCAAGAAGGCATTGACAGCTTCA  
GAGGTATATAATTTTGCATATACAGATTTAAGAGTAGTGTATAGAGACGGGTATCCGTATGGAGTG  
TGCAAGTTCTGTTTGCTATTTTATAGTAAGGTCCGAAAATTAAGATATTACAATTGTTTCAGTGTACG  
GGGCTAGCCTGGAAGCACTAACTAAAAAAAAGTTATCTGATTTATCAATAAGGTGCTACAGATGTC  
AACATCCGTTGACACCAGAGGAAAAACAGTTACACTGTGACTATAAGAAACGGTTTCACAAAATTT  
CACATATGTGGACCGGGTTCGTGCCTGACATGCTGGAGACACACAACAGCAACTGAATCAGCAGTA  
TAA 566

DF-128

102

ATGGATCGTCAGTTATTTGAAAATACAGAAGAGCGACCACGTACATTGCACCAGCTATGTGAAGTT  
GTGAATAAACCATTGCTGGAGCTGCAACTTGGCTGTGTGTTCTGCAAGAAGGCATTGACAGCGTCA  
GAGGTATATAATTTTGCATATACAGATCTAAGAGTAGTGTATAGAGACGGGTATCCGTATGGAGTG  
TGCAAATTCTGTTTGCTATTTTATAGTAAGGTCCGAAAATTAAGATATTATAAATTGTTTCAGTGTACG  
GGGCTAGCCTGGAAGAATACTAACTAAAAAAAAGTTATCTGATTTATCAATAAGGTGCTACAGATGTC  
AACATCCGTTGACACCAGAGGAAAAACAGTTACACTGTGACTATAAGAAACGGTTTCACAAAATTT  
CACATATGTGGACCGGGTTCGTGCCTGACATGCTGGAGACACACAACAGC 547

DF-141

102

ATGGATCGCCAGTTATTTGAAAATACAGAGGAGCGACCACGTACATTGCATCAGCTATGTGAAGTT  
GTGAATAAACCATTGCTGGAGCTGCAACTTGGCTGTGTGTTCTGCAAGAAGGCATTAACAGCTTCA  
GAGGTATATAATTTTGCATATACAGATTTAAGAGTAGTGTATAGAGACGGGTATCCGTATGGAGTG  
TGCAAGTTCTGTTTGCTATTTTATAGTAAGGTCCGAAAATTAAGATATTACAATTGTTTCAGTGTACG  
GGGCTAGCCTGGAAGCACTAACTAAAAAAAAGTTATCTGATTTATCAATAAGGTGCTACAGATGTC  
AACATCCGTTGACACCAGAGGAAAAACAGTTACACTGTGACTATAAGAAACGGTTTCACAAAATTT  
CACATATGTGGACCGGGTTCGTGCCTGACATGCTGGAGACACACAACAGCAACTGAATCAGCAGTA  
TAA 566

DF-153

102

ATGGATCGTCGGTTATTTGAAAATACACAGGAGCGACCACGTACATTGCATCAGCTCTGTGAAGTT  
GTGAATCAACCATTGCTGGAGCTGCAACTTGGCTGTGTGTTCTGCAAGAAGGCATTGACAGCTTCA  
GAGGTATATAATTTTGCATATAAAGATTTAAGAGTAGTGTATAGAGACGGGTATCCGTATGGAGTG  
TGCAAATTCTGTTTGCTATTTTATAGTAAGGTCCGAAAATTAAGACATTACAATTGTTTCAGTGTACG  
GGGCTAGCCTGGAAGCACTAACTAAAAAAAAGTTATCTGATTTATTAATAAGGTGCTACAGATGTC  
AACATCCGTTGACACCAGAGGAAAAACAGTTACACTGTGACTATAAGAAACGGTTTCACAAAATTT  
CACATATGTGGACCGGGTTCGTGCCTGACATGCTGGAGACACACAACAGCAACTGAATCAGCAGTA  
TAA 566

DF-180

102

ATGGATCGTCAGTTATTTGAAAATACAGAAGAGCGACCACGTACATTGCACCAGCTATGTGAAGTT  
GTGAATAAACCATTGCTGGAGCTGCAACTTGGCTGTGTGTTCTGCAAGAAGGCATTGACAGCGTCA  
GAGGTATATAATTTTGCATATACAGATCTAAGACTAGTGTATAGAGACGGGTATCCGTATGGAGTG  
TGCAAATTCTGTTTGCTATTTTATAGTAAGGTCCGAAAATTAAGATATTACAATTGTTTCAGTGTACG  
GGGCTAGCCTGGAAGTACTAACTAAAAAAAAGTTATCTGATTTATCAATAAGGTGCTACAGATGTC  
AACATCCGTTGACACCAGAGGAAAAACAGTTACACTGTGACTATAAGAAACGGTTTCACAAAATTT  
CACATATGTGGACCGGGTTCGTGCCTGACATGCTGGAGACACACAACAGCAACTGAATCAGCAGTA  
TAA 518

DF-231

102

ATGGATCGTCAGTTATTTGAAAATACAGAAGAGCGACCACGTACATTGCACCAGCTATGTGAAGTT  
 GTGAATACACCATTGCTGGAGCTGCAACTTGGCTGTGTGTTCTGCAAAAAGGCATTGACAGCGTCA  
 GAGGTATATAATTTTGCATATACAGATCTAAGAGTAGTGTATAGAGACGGGTATCCGTATGGAGTG  
 TGCAAATTTCTGTTTGTATTTTATAGTAAGGTCCGAAAATTAAGATATTACAATTGTTTCAGTGTACG  
 GGGCTAGCCTGGAAGCACTAACTAAAAAAAAGTTATCTGATTTATCAATAAGGTGCTACAGATGTC  
 AACATCCGTTGACACCAGAGGAAAAACAGTTACACTGTGACTATAAGAAACGGTTTCACAAAATTT  
 CACATATGTGGACCGGGTCGTGCCTGACATGCTGGAGACACACAACAGCAACTGAATCAG 558

**HPV-53 LCR**

DF-33

7302

ATAAATATGAATGAGTGTACTTTTTACGCGTGGTTGCATAAACTAAGGCGCGGTGGTGTCCCTAGG  
 CAGTTGGGTGGCATGTTAGGTGGCGTCCCTGTTTATATACCAATCTCCACCCTGTATTATTAATATA  
 TGTACAGTTATTATGCAGTTACCTATACCACATATAGTATAGCTTATAGGTATCCATTTTAGTTTGG  
 CGCCTTTTTGGTACCCTCCATTTTACCTTGCAACCGGTTTCGGTTCGTTGGGACATATCCTGTAAATA  
 CCACAAGCATATTCAGCAGAAGTGTAAATCCTTTGGCATAGTGCCGTTTCCTGTGTTAATGTTTTTT  
 TGTGCTATACTATGCCTCACCTATTAGTCACATACTGTGGGCTGCCAACTATGCTTTTAGCTGCA  
 TACTTTAGTGCTTTTGGGCACACATTTTATTGCAAACATCTGCACTTTAGTGTTTTGGCTTGCAGCA  
 TACGTTTTGCTAGCCAAGTATCTGTCTAGTAACTAGGTGTGCACCTCTTACTCATGCACCGCAACC  
 GTTTACGGTTTTGCAGCAACAGGCCTTTTTTATAATTA 7837

DF-128

7279

TGTTAATAAATATGAATGAGTGTACTTTTTACGCGTGGTTGCATAAACTAAGGCGCGGTGGTGTCC  
 CTAGGCAGTTGGGTGGCATGTTAGGTGGCGTCCCTGTTTATATACCAATCTCCACCCTGTATTATTA  
 ATATATGTACTGTTATCATGCAGTTACCTATACCACATATAGTATAGCTTATAGGTATCCATTTTAG  
 TTTGGCGCCTTTTTGGTACCCTCCATTTTACCTTGCAACCGGTTTCGGTTCGTTGCATATCTTGTAATA  
 CCACAAGCATATTCAGCAGAAGTGTAAATCCTTTGGCATAGTGCCGTTTCCTGTGTTAATGTTTTTT  
 TGTTTTATACACTATGCCTTACCTATTAGTCACTTACTGTAGGCTGCCAACTATGCTTTTACCTGCAT  
 ACCTTAGTGCTTTTGGGCACACATTTTATTGCAAACATCTGCACTTTAGTGTTTTGGCTTGCAGCAT  
 ACGTTCGTAGCCAAGTATCTGTCTAGTAAACCAGGTGTGCACCTATTACTCATGCACCGCAACC  
 GT 7818

DF-141

7284

ATAAATATGAATGAGTGTACTTTTTACGCGTGGTTGCATAAACTAAGGCGCGGTGGTGTCCCTAGG  
 CAGTTGGGTGGCATGTTAGGTGGCGTCCCTGTTTATATACCAATCTCCACCCTGTATTATTAATATA  
 TGTACAGTTATTATGCAGTTACCTATACCACATATAGTATAGCTTATAGGTATCCATTTTAGTTTGG  
 CGCCTTTTTGGTACCCTCCATTTTACCTTGCAACCGGTTTCGGTTCGTTGGGACATATCCTGTAAATA  
 CCACAAGCATATTCAGCAGAAGTGTAAATCCTTTGGCATAGTGCCGTTTCCTGTGTTAATGTTTTTT  
 TGTGCTATACTATGCCTCACCTATTAGTCACTTACTGTGGGCTGCCAACTATGCTTTTACCTGCAT  
 TACTTTAGTGCTTTTGGGCACACATTTTATTGCAAACATCTGCACTTTAGTGTTTTGGCTTGCAGCA  
 TACGTTTTACTAGCCAAGTATCTGTCTAGTAACTAGGTGTGCACCTCTTACTCATGCACCTGCAACC  
 GTTTACGGTTTTGCAGCAACAGGCCTTTTTTATAATTATTGAAAGTAACA 010

DF-153

7284

ATAAATATGAATGAGTGTACTTTTTACGCGTGGTTGCATAAACTAAGGCGCGGTGGTGTCCCTAGG  
 CAGTTGGGTGGCATGTTAGGTGGCGTCCCTGTTTATATACCAATCTCCACCCTGTATTATTAATATA  
 TGTACAGTTATTATGGAGTTACCTATACCACATATAGTATAGCTTATAGGTATCCATTTTAGTTTGG  
 CGCCTTTTTGGTACCCTCCATTTTACCTTGCAACCGGTTTCGGTTCGTTGGGACATATCCTGTAAATA  
 CCACAAGCATATTCAGCAGAAGTGTAAATCCTTTGGCATAGTGCCGTTTCCTGTGTTAATGTTTTTT  
 TGTGCTATACTATGCCTCACCTATTAGTCACTTACTGTGGGCTGCCAACTATGCTTTTACCTGCAT  
 TACTTTAGTGCTTTTGGGCACACATTTTATTGCAAACATCTGCACTTTAGTGTTTTGGCTTGCAGCA

TACGTTTTACTAGCCAAGTATCTGTCTAGTAACTAGGTGTGCACCTCTTACTCATGCACTGCAACC  
GTTTACGGTTTTGCAGCAACAGGCCTTTTTTATAATTATTGAAAGTAACAATCCTACT 018

**DF-180**

7313

GTGGTTGCATAAACTAAGGCGCGGTGGTGTCCCTAGGCAGTTGGGTGGCATGTTAGGTGGCGTCCC  
TGTTTATATACCAATCTCCACCCTGTATTATTAATATATGTACTGTTATCATGCAGTTACCTATACCA  
CATATAGTATAGCTTATAGGTATCCATTTTGTGGCGCTTTTTGGTACCCTCCATTTTACCTTGC  
AACCGTTTTCGGTTGTGCATATCTTGTAAATACCACAAGCATATTCAGCAGAAGTGTAAATCCTTTG  
GCATAGTGCCGTTTCTGTGTTTAAATGTTTTTTGTTTTATACACTATGCCTTACCTATTAGTCACTT  
ACTGTAGGCTGCCAACTATGCTTTTACCTGCATACCGTAGTGCTTTTGGGCACACATTTTTATTGCA  
AACATCTGCACCTTAGTGTTTTGGCTTGCAGCATACGTTCTGCTAGCCAAGTATCTGTCTAGTAAAC  
CAGGTGTGCACCTATFACTCATGCACCGCAACCGTTTTACGGTTTTGCAGCAACAGGCCTTTTTTATA  
ATTATTGAAAGTAACAATCCTACT 018

**DF-231**

7278

TTGTTAATAAATATGAATGAGTGTTACTTTTACGCGTGGTTGCATAAACTAAGGCGCGGTGGTGTCC  
CTAGGCAGTTGGGTGGCATGTTAGGTGGCGTCCCTGTTTATATACCAATCTCCACCCTGTATTATTA  
ATATATGTACTGTTATCATGCAGTTACCTATACCACATATAGTATAGCTTATAGGTATCCATTTTGT  
TTTGGCGCTTTTTGGTACCCTCCATTTTACCTTGCACCGGTTTTCGGTTGTGCATATCTTGTAAATA  
CCACAAGCATATTCAGCAGAAGTGTAAATCCTTTGGCATAGTGCCGTTTCTGTGTTTAAATGTTTTTT  
TGTTTTATACACTATGCCTCACCTATTAGTCACTTACTGTAGGCTGCCAACTATGCTTTTACCTGCAT  
ACCTTAGTGCTTTTGGGCACACATTTTTATTGCAAACATCTGCACCTTAGTGTTTTGGCTTGCAGCAT  
ACGTTCTGCTAGCCAAGTATCTGTCTAGTAAACCAGGTGTGCACCTATTACTCATGCACCGCAACC  
GTTTACGGTTTTGCAGCAA 7835

**HPV-53 L1**

**DF-33**

6661

GGAATACAAACATGACTCTTTCTGCAACCACACAGTCTATGTCCACATATAATTCAAAGCAAATTA  
AACAGTATGTTAGGCATGCAGAGGAATATGAATTACAATTTGTGTTTCAACTATGTAAAATATCCC  
TGCTGCTGAGGTTATGGCCTATTTACATACTATGAATTCTACCTTACTGGAAGACTGGAATATAGG  
TTTGTGCGCTCCTGTTGCCACTAGCTTAGAGGACAAATACAGATATGTAAAAAGTGCAGCTATAAC  
CTGTCAAAGGATCAGTCCCTCCTGAAAAGCAGGACCCACTATCTAAATATAAATTTTGGGAGGT  
CAATTTACAAAACAGTTTTTCTGCTGATTTGGATCAGTT 7030

**DF-128**

6681

TCCGCAACCACACAGTCTATGTCTACATATAATTACAGAGCAAATTAACAGTATGTTAGACATGCA  
GAGGAATATGAATTACAATTTGTGTTTCAACTATGTAAAATATCCCTGTCTGCTGAGGTTATGGCCT  
ATTTACATACTATGAATTCTACCTTACTGGAAGACTGGAATATAGGTTTGTGCGCTCCTGTTGCCAC  
TAGCTTAGAGGACAAATACAGATATGTGAAAAGTGCAGCTATAACCTGTCAAAGGATCAGCCCC  
CTCCTGAAAAGCAGGACCCACTATCTAAATATAAATTTTGGGAGG 6990

**DF-153**

6594

GCACAGGGACATAATAATGGCATCTGTTGGAACAATCAGTTATTTGTAAGTGTGTTGGATACCACC  
AGGAATACAAACATGACACTTTCTGCAACCACACAGTCTATGTCCACATATAATTCAAAGCAAATT  
AAACAGTATGTTAGGCATGCAGAGGAATATGAATTACAATTTGTGTTTCAACTATGTAAAATATCC  
CTGTCTGCTGAGGTTATGGCCTATTTACATACTATGAATTCTACCTTACTGGAAGACTGGAATATAG  
GTTTGTGCGCTCCTGTTGCCACTAGCTTAGAGGACAAATACAGATATGTAAAAGTGCAGCTATAA  
CCTGTCAAAGGATCAGCCCCCTCCTGAAAAGCAGGACCCACTATCTAAATATAAATTTTGGGAGG  
TCAATTTACAAAACAGTTTTTCTGCTGATTTGGATCAGTTTCCCTTGGACG 7042

DF-180

6641

AACTGTTGTGGATACCACCAGGAATACAAACATGACTCTTTCCGCAACCACACAGTCTATGTCTAC  
 ATATAATTCAAAGCAAATTAACAGTATGTTAGACATGCAGAGGAATATGAATTACAATTTGTGTT  
 TCAACTATGTAAAATATCCCTGTCTGCTGAGGTTATGGCCTATTTACATACTATGAATTCTACCTTA  
 CTGGAAGACTGGAATATAGGTTTGTGCGCTCCTGTTGCCACTAGCTTAGAGGACAAATACAGATAT  
 GTGAAAAGTGCAGCTATAACCTGTCAAAGGATCAGCCCCCTCCTGAAAAGCAGGACCCACTATCT  
 AAATATAAATTTT 6985

DF-231

6610

ATGGCATCTGTTGGAACAATCAGTTATTTGTAAGTGTGGATACCACCAGGAATACAAACATGA  
 CTCTTTCCGCAACCACACAGTCTATGTCTACATATAATTCAAAGCAAATTAACAGTATGTTAGAC  
 ATGCAGAGGAATATGAATTACAATTTGTGTTTCAACTATGTAAAATATCCCTGTCTGCTGAGGTTAT  
 GGCCTATTTACATACTATGAATTCTACCTTACTGGAAGACTGGAATATAGGTTTGTGCGCTCCTGTT  
 GCCACTAGCTTAGAGGACAAATACAGATATGTGAAAAGTGCAGCTATAACCTGTCAAAGGATCA  
 GCCCCCTCCTGAAAAGCAGGACCCACTATCTAAATATAAATTTTGGGAGGTCAATTTGCAAACAG  
 TTTTCTGC 7015

**HPV-56 E6**

DF-08

102

ATGGAGCCACAATTCAACAATCCACAGGAACGTCCACGAAGCCTGCACCACTTGAGTGAGGTATTA  
 GAAATACCTTTAATTGATCTTAGATTATCATGTGTATATTGCAAAAAAGAAGTAACACGTGCTGAG  
 GTATATAATTTTGCATGCACTGAATTAATAATTAGTGTATAGGGATGATTTTCCTTATGCAGTGTGCA  
 GAGTATGTTTATTGTTTTATAGTAAAGTTAGAAAATATAGGTATTATGACTATTCAGTGTATGGAGC  
 TACTAGAAAAGTATAACTAAAAACAGTTATGTGATTTATTAATAAGGTGCTACAGATGTCAAAG  
 TCCGTTAACTCCGGAGGAAAAGCAATTGCATTGTGACAGAAAAAGACGATTTTCATCTAATAGCACA  
 TGGTTGGACCGGGTCATGTTTGGGGTGCTGGAGACAAACATCTAGAGAACCTAGAGAATCTACAGT  
 A 566

DF-26

102

ATGGAGCCACAATTCAACAATCCACAGGAACGTCCACGACGCCTGCACCACTTGAGTGAGGTATTA  
 GAAATACCTTTAATTGATCTTAGATTATCATGTGTATATTGCAAAAAAGAAGTAACACGTGCTGAG  
 GTATATAATTTTGCATGCACTGAATTAATAATTAGTGTATAGGGATGATTTTCCTTATGCAGTGTGCA  
 GAGTATGTTTATTGTTTTATAGTAAAGTTAGAAAATATAGGTATTATGACTATTCAGTGTATGGAGC  
 TACTAGAAAAGTATAACTAAAAACAGTTATGTGATTTATTAATAAGGTGCTACAGATGTCAAAG  
 TCCGTTAACTCCGGAGGAAAAGCAATTGCATTGTGACAGAAAAAGACGATTTTCATCTAATAGCAC  
 ATGGTTGGACCGGGTCATGTTTGGGGTGCTGGAGACAAACATCTAGAGAACCTAGAGAATCTACA  
 GTA 566

DF-80

102

ATGGAGCCACAATTCAACAATCCACAGGAACGTCCACGAAGCCTGCACCACTTGAGTGAGGTATTA  
 GAAATACCTTTAATTGATCTTAGATTATCATGTGTATATTGCAAAAAAGAAGTAACACGTGCTGAG  
 GTATATAATTTTGCATGCACTGAATTAATAATTAGTGTATAGGGATAATTTTCCTTATGCAGTGTGCA  
 GAGTATGTTTATTGTTTTATAGTAAAGTTAGAAAATATAGGTATTATGACTATTCAGTGTATGGAGC  
 TACTAGAAAAGTATAACTAAAAACAGTTATGTGATTTATTAATAAGGTGCTACAGATGTCAAAG  
 TCCGTTAACTCCGGAGGAAAAGCAATTGCATTGTGACAGAAAAAGACGATTTTCATCTAATAGCACA  
 TGGTTGG 506

DF-123

102

ATGGAGCCACAATTCAACAATCCACAGGAACGTCCACGACGCCTGCACCACTTGAGTGAGGTATTA  
 GAAATACCTTTAATTGATCTTAGATTATCATGTGTATATTGCAAAAAAGAAGTAACACGTGCTGAG  
 GTATATAATTTTGCATGCACTGAATTAATAATTAGTGTATAGGGATGATTTTCCTTATGCAGTGTGCA

GAGTATGTTTATTGTTTTATAGTAAAGTTAGAAAATATAGGTATTATGACTATTCAGTGTATGGAGC  
TACTAGAAAAGTATAACTAAAAACAGTTATGTGATTTATTAATAAGGTGCTACAGATGTCAAAG  
TCCGTTAACTCCGGAGGAAAAGCAATTGCATTGTGACAGAAAAAGACGATTCATCTAATAGCACA  
TGGTTGGACCGGGTCATGTTTGGGGTGCTGGAGACAAACATCTAGAGAACCTAGAGAATCT 560

DF-230

102

ATGGAGCCACAATTCAACAATCCACAGGAACGTCCACGACGCCTGCACCACTTGAGTGAGGTATTA  
GAAATACCTTTAATTGATCTTAGATTATCATGTGTATATTGCAAAAAAGAACTAACACGTGCTGAG  
GTATATAATTTGCATGCACTGAATTAAGTTAGTGTATAGGGATGATTTTCCTTATGCAGTGTGCA  
GAGTATGTTTATTGTTTTATAGTAAAGTTAGAAAATATAGGTATTATGACTATTCAGTGTATGGAGC  
TACTAGAAAAGTATAACTAAAAACAGTTATGTGATTTATTAATAAGGTGCTACAGATGTCAAAG  
TCCGTTAACTCCGGAGGAAAAGCAATTGCATTGTGACAGAAAAAGACGATTCATCTAATAGCACA  
TGGTTGGACCGGGTCATGTTTGGGGTGCTGGAGACAAACATCTAGAGAACCTAGAGAATCTACAGT  
ATAA 569

## HPV-56 LCR

DF-08

7342

TTGTTTTGTAAGTTATTGTGTACAGTGTACTATGTGTATTGTGCATACATATATATACCATAACATA  
CTCCATTTTGTGTTTTCCGCCATTTGTACATGCAACCGAATTCGGTTGCATGGCCTAGTGCCATT  
ATTTAACTAAAAGGAATTCGGTTGCATGGCCTAGTGCCATTATTTAAACCAAAGGCCCTTTTCA  
GCAGAACAGTTAATCCTTTGGCATAATTGCCGTTTCCTGTGTTTTATACTTGAATTATGTACAGTACC  
GCACCCTGTATTACTCACAGGTAATGACTGCCAACTATGCTTTTATCTGCATACTTTAGTGCTGT  
TGGGCACACATTTTTATACATGTGTCTGCAACTTTGGTGTTTTGGCTTGCAGAATACACTATGTAGG  
CCAAGTATCTGTCAGTATCTGTTTTGCAAACATATAACATAACAATTAATCACTTTTTAAAACCGTTT  
ACGGTCGTGCAAAAACAGGTTTCTTTAATTGTTGAAAGTTTCAATCATACTTTTATATATTGGGAG  
TGACCGAAAAGGGTTTAAGACCGAAAACGGTACATATAAAAGGCAGCTTATTCTGTGTGGACATAT  
CC 0101

DF-26

7332

ATTGTGGCTTTTGTGTTTTGTAAGTTATTGTGTACAGTGTACTATGTGTATTGTGCATACATATATAC  
CATAACATACTCCATTTTGTGTTTTCCGCCATTTGTACCTGCAACCGAATTCGGTTGCATGGCCT  
ATTGCCATTATTTAAACCAAAGGCCCTTTTACAGCAGAACAGTTAATCCCTTTGGCATAATTGCCGTTT  
CCTGCGTTTTATACTTAAAGTATGTACAGTACCGCACCTGTATTACTACCGGTAATGACTGCC  
AACTATGCTTTTATCTGCATACTTTAGTGCTGTTGGGCACACATTTTTATACATGTGTCTGCAACTTT  
GGTGTGTTTTGGCTTGCAGAATACACTATGTAGGCCAAGTATCTGTCAGTATCTGTTTTGCAAACATGT  
AACATAACAATTAATCACTTTTTGAAAACCGTTTACGGTCGTGCAAAAACAGGTTTCTTTAATTGTTG  
AAAGTTTTAATCATACTTTTATATATTGGGAGTAACCGAAAAGGGTTTAAGACCG 056

DF-80

7333

TTGTGGCTTTTGTGTTTTGTAAGTTATTGTGTACAGTGTACTAAGTGTATTGGGCATACAAATATATAC  
CATAACATACTCCATTTTGTGTTTTCCGCCATTTGTACATGCAACCGAATTCGGTTGCATGGCCT  
AGTGCCATTATTTAAACCAAAGGCCCTTTTACAGCAGAACAGTTAATCCCTTTGGCATAATTGCCGTTT  
CCTGCGTTTTATACTTAAAGTATGTACAGTACCGCACCTGTATTACTACCGGTAATGACTGCC  
AACTATGCTTTTATCTGCATACTTTAGTGCTGTTGGGCACACATTTTTATACATGTGTCTGCAACTTT  
GGTGTGTTTTGGCTTGCAGAATACACTATGTAGGCCAAGTATCTGTCAGTATCTGTTTTGCAAACATGT  
AACATAACAATTAATCACTTTTTGAAAACCGTTTACGGTCGTGCAAAAACAGGTTTCTTTAATTGTTG  
AAAGTTTCAATCATACTTTTATATACTGGGAGTAACCGAAAAGGGTTCAAGACCGAAAACCGGTACA  
TATAAAAGGCAGCTTATTCTGTGTGGACATATCC 101

DF-123

7387

TACATATATATACCATAACATACTCCATTTTGTGTTTTCCGCCATTTTGTACATGCA-  
CCGAATTCGGTTGCATGGCCTATTGCCATTATTTAAACCAAAGGCCCTTTTACAGCAGAACAGTTA  
ATCCCTTTGGCATAATTGCCGTTTCCTGCGTTTTATACTTGAATATGTACAGTACCGCACCTGTATTA

CTCACCGTACTATGACTGCCAACTATGCTTTTATCTGCATACTTTAGTGCTGTTGGGCACACATTT  
 TTATACATGTGTCTGCAACTTTGGTGTTTTGGCTTGCAGAATACTATGTAGGCCAAGTATCTGTC  
 AGTATCTGTTTTGCAAACATGTAACATAACAATTACTCATTTTTGAAAACCGTTTACGGTCGTGCAAA  
 AACAGTTTTCTTTAATTGTTGAAAGTTTCAATCATACTTTTATATATTGGGAGTAACCGAAAAGGG  
 TTTAAGACCGAAAACGGTACATATAAAAGGCAGCTTATTCTGTGTGGACATATTC 101

DF-230

7364

CAGTGTACTATGTGTATTGTGCATACATATATATACCATAACATACTCCATTTTGTGTTTTTCCGCC  
 ATTTTGTACATGCAACCGAATTCGGTTGCATGGCCTATTGCCATTATTTAAACCAAAGGCCCTTTT  
 CAGCAGAACAGTTAATCCCTTGGCATATTGCCGTTTCTGCGTTTTATACTTGAACATGTACAGTA  
 CCGCACCTGTATTACTCACCGTACTATGACTGCCAACTATGCTTTTATCTGCATACTTTAGTGCT  
 GTTGGGCACACATTTTTATACATGTGTCTGCAACTTTGGTGTTTTGGCTTGCAGAATACTATGTA  
 GGCCAAGTATCTGTCAGTATCTGTTTTGCAAACATGTAACATAACAATTACTCATTTTTGAAAACCGT  
 TTACGGTCGTGCAAAAACAGGTTTCTTTAATTGTTGAAAGTTTTAATCATACTTTTATATATTGGG  
 AGTAACCGAAAAGGGTTAAGACCGAAAACGGTACATATAAAAGGCAGCTTATTCTGTGTGGACA  
 TATTC 101

## HPV-56 L1

DF-08

6584

GTTACTGTAGTAGATACTACTAGAAAGTACTAACATGACTATTAGTACTGCTACAGAACAGTTAAGT  
 AAATATGATGCACGAAAAATTAATCAGTACCTTAGACATGTGGAGGAATATGAATTACAATTTGTT  
 TTCAATTATGCAAAATTACTTTGTCTGCAGAGGTTATGGCATATTTACATAATATGAATGCTAACCC  
 TACTGGAGGACTGGAATATTGGGTTATCCCCGCCAGTGGCCACCAGCCTAGAAGATAAATATAGAT  
 ATGTTAGAAGCACAGCTATAACATGTCAACGGGAACAGCCACCAACAGAAAAACAGGACCCATTA  
 GCTAAATATAAATTTTGGGATGTAACTTACAGGACAGTTTTTCTACAGACCT 6967

DF-26

6548

CATAATAATGGCATTGCTGGGGTAATCAATTATTTGTTACTGTAGTAGATACTACTAGAAAGTACTA  
 ACATGACTATTAGTACTGCTACAGAACAGTTAAGTAAATATGATGCACGAAAAATTAATCAGTACC  
 TTAGACATGTGGAGGAATATGAATTACAATTTGTTTTCAATTATGCAAAATTACTTTGTCTGCAGA  
 GGTTATGGCATATTTACATAATATGAATGCTAACCTACTGGAGGACTGGAATATTGGGTTATCCCC  
 GCCAGTGGCCACTAGCCTAGAAGATAAATATAGATATGTTAGAAGCACAGCTATAACATGTCAAC  
 GAGAACAGCCACCAACAGAAAAACAGGACCCATTAGCTAAATATAAATTTTGGGATGTAACTTA  
 CA 6946

DF-80

6629

ACTGCTACAGAACAGTTAAGTAAATATGATGCACGAAAAATTAATCAGTACCTTAGACATGTGGAG  
 GAATATGAATTACAATTTGTTTTCAATTATGCAAAATTACTTTGTCTGCAGAGGTTATGGCATATT  
 TACATAATATGAATGCTAACCTACTGGAGGACTGGAATATTGGGTTATCCCCGCCAGTGGCCACCA  
 GCCTAGAAGATAAATATAGATATGTTAGAAGCACAGCTATAACATGTCAACGGGAACAGCCACCA  
 ACAGAAAAACAGGACCCATTAGCTAAATATAAATTTTGGGATGTAACTTACAGGACAGTTTTTCT  
 ACAGACCTGGATCA 6972

DF-123

6591

TAGTAGATACTACTAGAAAGTACTAACATGACTATTAGTACTGCTACAGAACAGTTAAGTAAATATG  
 ATGCACGAAAAATTAATCAGTACCTTAGACATGTGGAGGAATATGAATTACAATTTGTTTTCAAT  
 TATGCAAAATTACTTTGTCTGCAGAGGTTATGGCATATTTACATAATATGAATGCTAACCTACTGGA  
 GGACTGGAATATTGGGTTATCCCCGCCAGTGGCCACTAGCCTAGAAGATAAATATAGATATGTTAG  
 AAGCACAGCTATAACATGTCAACGAGAACAGCCACCAACAGAAAAACAGGACCCATTAGCTAAAT  
 ATAAATTTTGGGATGTAACTTACAGGACAGTTTTTCTACAGACCTGGATCA 6972

DF-230



6625

TTAGTACTGCTACAGAACAGTTAAGTAAATATGATGCACGAAAAATTAATCAGTACCTTAGACATG  
 TGGAGGAATATGAATTACAATTTGTTTTCAATTATGCAAAAATTACTTTGTCTGCAGAGGTTATGGC  
 ATATTTACATAATATGAATGCTAACCTACTGGAGGACTGGAATATTGGGTTATCCCCGCCAGTGGC  
 CACTAGCCTAGAAGATAAATATAGATATGTTAGAAGCACAGCTATAACATGTCAACGAGAACAGC  
 CACCAACGAAAAACAGGACCCATTAGCTAAATATAAATTTTGGGATGTTAACTTACAGGACAGTT  
 TTTCTACAGACCTGGATCA 6973

**HPV-66 E6**

DF-24

102

ATGGATTCCATATTCAGCAATACACAGGAACGTCCACGAAGCCTGCACCATCTGAGCGAGGTATTA  
 CAAATACCTTTACTTGATCTTAGATTATCATGTGTATACTGCAAAAAGGAACTTACAAGTTTAGAGC  
 TATATAGGTTTGCATGTATTGAGTTAAAAGTATATAGAAACAATTGGCCATATGCAGTATGTA  
 GGGTATGTTTATTGTTTTATAGTAAGGTTAGAAAATATAGGTACTATAAATATTCAGTGTATGGGGC  
 AACATTAGAAAGTATAACTAAAAACAGTTATCTGATTTATCAATAAGGTGCTACCGATGTCAATG  
 TCCGTTAACACCGGAGGAAAAACAATTGCACTGTGAACATAAAAGACGATTTTATTATATAGCATA  
 TGCATGGACCGGGTCATGTTTGCAGTGTGGAGACATACGAGTAGACAAGCTACAGAATCTACAGT  
 ATAA 569

DF-41

102

ATGGATTCCATATTCAGCAATACACAGGAACGTCCACGAAGCCTGCACCATCTGAGCGAGGTATTA  
 CAAATACCTTTACTTGATCTTAGATTATCATGTGTATACTGCAAAAAGGAACTTACAAGTTTAGAGC  
 TATATAGGTTTGCATGTATTGAGTTAAAAGTATATAGAAACAATTGGCCATATGCAGTATGTA  
 GGGTATGTTTATTGTTTTATAGTAAGGTTAGAAAATATAGGTACTATAAATATTCAGTGTATGGGGC  
 AACATTAGAAAGTATAACTAAAAACAGTTATCTGATTTATCAATAAGGTGCTACCGATGTCAATG  
 TCCGTTAACACCGGAGGAAAAACAATTGCACTGTGAACATAAAAGACGATTTTATTATATAGCATA  
 TGCATGGACCGGGTCATGTTTGCAGTGTGGAGACATACGAGTAGACAAGCTACAGAATCTACAGT  
 ATAA 569

DF-107

102

ATGGATTCCATATTCAGCAATACACAGGAACGTCCACGAAGCCTGCACCATCTGAGCGAGGTATTA  
 CAAATACCTTTACTTGATCTTAGATTATCATGTGTATACTGCAAAAAGGAACTTACAAGTTTAGAGT  
 TATATAGGTTTGCATGTATTGAGTTAAAATTAGTATATAGAAACAATTGGCCATATGCAGTATGTA  
 GGGTATGTTTATTGTTTTATAGTAAGGTTAGAAAATATAGGTACTATAAATATTCAGTGTATGGGGC  
 AACATTAGAAGGTATAACTAAAAACAGTTATCTGATTTATCAATAAGGTGCTACCGATGTCAATG  
 TCCGTTAACACCGGAGGAAAAACAATTGCACTGTGAACATAAAAGACGATTTTATTATATAGCATA  
 TGCATGGACCGGGTCATGTTTGCAGTGTGGAGACATACGAGTAGACAAGCCACAGAATCTACAGT  
 ATAA 569

DF-198

102

ATGGATTCCATATTCAGCAATACACAGGAACGTCCACGAAGCCTGCACCATCTGAGCGAGGTATTA  
 CAAATACCTTTACTTGATCTTAGATTATCATGTGTATACTGCAAAAAGGAACTTACAAGTTTAGAGC  
 TATATAGGTTTGCATGTATTGAGTTAAAAGTATATAGAAACAATTGGCCATATGCAGTATGTA  
 GGGTATGTTTATTGTTTTATAGTAAGGTTAGAAAATATAGGTACTATAAATATTCAGTGTATGGGGC  
 AACATTAGAAAGTATAACTAAAAACAGTTATCTGATTTATCAATAAGGTGCTACCGATGTCAATG  
 TCCGTTAACACCGGAGGAAAAACAATTGCACTGTGAACATAAAAGACGATTTTATTATATAGCATA  
 TGCATGGACCGGGTCATGTTTGCAGTGTGGAGACATACGAGTAGACAAGCTACAGAATCTACAGT  
 ATAA 569

DF-224

102

ATGGATTCCATATTCAGCAATACACAGGAACGTCCACGAAGCCTGCACCATCTGAGCGAGGTATTA  
 CAAATACCTTTACTTGATCTTAGATTATCATGTGTATACTGCAAAAAGGAACTTACAAGTTTAGAGC

TATATAGGTTTGCATGTATTGAGTTAAAAGTATATAGAAAACAATTGGCCATATGCAGTATGTA  
GGGTATGTTTATTGTTTTATAGTAAGGTTAGAAAATATAGGTTACTATAAATATTCAGTGTATGGGGC  
AACATTAGAAAAGTATAACTAAAAACAGTTATCTGATTTATCAATAAGGTGCTACCGATGTCAATG  
TCCGTTAACACCGGAGGAAAAACAATTGCACTGTGAACATAAAAAGACGATTCATTATATAGCATA  
TGCATGGACCGGGTCATGTTTGCAGTGTGGAGACATACGAGTAGACAAGCTACAGAATCTACAGT  
ATAA 569

DF-228

102

ATGGATTCCATATTCAGCAATACACAGGAACGTCCACGAAGCCTGCACCATCTGAGCGAGGTATTA  
CAAATACCTTTACTTGATCTTAGATTATCATGTGTATACTGCAAAAAGGAACTTACAAGTTTAGAGC  
TATATAGGTTTGCATGTATTGAGTTAAAAGTATATAGAAAACAATTGGCCATATGCAGTATGTA  
GGGTATGTTTATTGTTTTATAGTAAGGTTAGAAAATATAGGTTACTATAAATATTCAGTGTATGGGGC  
AACATTAGAAAAGTATAACTAAAAACAGTTATCTGATTTATCAATAAGGTGCTACCGATGTCAATG  
TCCGTTAACACCGGAGGAAAAACAATTGCACTGTGAACATAAAAAGACGATTCATTATATAGCATA  
TGCATGGACCGGGTCATGTTTGCAGTGTGGAGACATACGAGTAGACAAGCTACAGAATCTACAGT  
ATAA 569

## HPV-66 LCR

DF-24

7171

TGTTGTATGTATTGTATGGTTGTGCTTGTACTGTATGTTTTGTGTATGTTTATGTATTTTATAAATTGT  
GTATGTGCTATGTGTATGTATGACTGTATGTATGTGTAATGTTTTGTGTGTATGTAATAAACATGCA  
TGGTTACTTTTACGCGTGGTTGCATAAACTAAGGTGCGGTAGTATCCTTGGGCAGTGTGTGTCAGGT  
TAGGTGGTGTTCCTTACTGTTAATGTTATATTAATAAGGTTGTTTGTATGCACTATAGTAACACAC  
CAAACCTCATTTTAGTGCTGTACGCCATTTTATGCATGCAACCGAATTCGGTTGCCTAGCCTTTTGT  
CCTTATTTAAACCCAAAACGACTTTTCAGCAAAAACAGTTAATCCTTTGGCATATTGCCGTTTCCTGT  
TGTATGATTCAGGTATGTACACTGCCTTACCCTGTATTACTCACCTGTATTTCTGTGCCAACTATGCT  
TTTATCTGCATACTTTGGCGCTGTTGGGCATATGTTTTTATGCAGGTGTTTGCAATATATTTTGTGG  
CGTGTAGCCCTTATTGTATAAGCCAAGTATCTGTCTTGCAAATATGTAACCATATACTTACTCATT  
TACAAAACCGTTTACGGTCGTGCTAAAACAGGTTTCTTTTAATTGTTGAAAGTTTCAATCATACTTT  
ATTATATTGGGAGTAACCGAAATGGGTTTAGGACCGAAAACGGTACATATAAAAAGGCAGCCTGTT  
GTGCCTGTAGATATCC 101

DF-41

7171

TGTTGTATGTATTGTATGGTTGTGCTTGTACTGTATGTTTTGTGTATGTTTATGTATTTTATAAATTGT  
GTATGTGCTATGTGTATGTATGACTGTATGTATGTGTAATGTTTTGTGTGTATGTAATAAACATGCA  
TGGTTACTTTTACGCGTGGTTGCATAAACTAAGGTGCGGTAGTATCCTTGGGCAGTGTGTGTCAGGT  
TAGGTGGTGTTCCTTACTGTTAATGTTATATTAATAAGGTTGTTTGTATGCACTATAGTAACACAC  
CAAACCTCATTTTAGTGCTGTACGCCATTTTATGCATGCAACCGAATTCGGTTGCCTAGCCTTTTGT  
CCTTATTTAAACCCAAAACGACTTTTCAGCAAAAACAGTTAATCCTTTGGCATATTGCCGTTTCCTGT  
TGTATGATTCAGGTATGTACACTGCCTTACCCTGTATTACTCACCTGTATTTCTGTGCCAACTATGCT  
TTTATCTGCATACTTTGGCGCTGTTGGGCATATGTTTTTATGCAGGTGTTTGCAATATATTTTGTGG  
CGTGTAGCCCTTATTGTATAAGCCAAGTATCTGTCTTGCAAATATGTAACCATATACTTACTCATT  
TACAAAACCGTTTACGGTCGTGCTAAAACAGGTTTCTTTTAATTGTTGAAAGTTTCAATCATACTTT  
ATTATATTGGGAGTAACCGAAATGGGTTTAGGACCGAAAACGGTACATATAAAAAGGCAGCCTGTT  
GTGCCTGTAGATATCC 101

DF-198

7171

TGTTGTATGTATTGTATGGTTGTGCTTGTACTGTATGTTTTGTGTATGTTTATGTATTTTATAAATTGT  
GTATGTGCTATGTGTATGTATGACTGTATGTATGTGTAATGTTTTGTGTGTATGTAATAAACATGCA  
TGGTTACTTTTACGCGTGGTTGCATAAACTAAGGTGCGGTAGTATCCTTGGGCAGTGTGTGTCAGGT  
TAGGTGGTGTTCCTTACTGTTAATGTTATATTAATAAGGTTGTTTGTATGCACTATAGTAACACAC  
CAAACCTCATTTTAGTGCTGTACGCCATTTTATGCATGCAACCGAATTCGGTTGCCTAGCCTTTTGT  
CCTTATTTAAACCCAAAACGACTTTTCAGCAAAAACAGTTAATCCTTTGGCATATTGCCGTTTCCTGT  
TGTATGATTCAGGTATGTACACTGCCTTACCCTGTATTACTCACCTGTATTTCTGTGCCAACTATGCT

TTTATCTGCATACTTTGGCGCTGTTGGGCATATGTTTTTATGCAGGTGTTTGCAATATATTTTTGTTGG  
CGTGTAGCCCTTATTGTATAAGCCAAGTATCTGTCTTGCAAATATGTAACCATATACTTACTCATT  
TACAAAACC 7786

DF-224

7171

TGTTGTATGTATTGTATGGTTGTGCTTGTACTGTATGTTTTTGTGTATGTTTATGTATTTTATAAATTGT  
GTATGTGCTATGTGTATGTATGACTGTATGTATGTGTAATGTTTTGTGTGTATGTAATAACATGCA  
TGGTTACTTTTACGCGTGGTTGCATAAACTAAGGTGCGGTAGTATCCTTGGGCAGTGTGTGTCAGGT  
TAGGTGGTGTTCCTTACTGTTTAAATGTTATATTAATAGGTTGTTTGTATGCACTATAGTAACACAC  
CAAACCTCCATTTTAGTGCTGTACGCCATTTTATGCATGCAACCGAATTCGGTTGCCTAGCCTTTTGT  
CCTTATTTAAACCCAAAACGACTTTTCAGCAAAAACAGTTAATCCTTTGGCATATTGCCGTTTCCTGT  
TGTATGATTCAGGTATGTACACTGCCTTACCCTGTATTACTCACCTGTATTTCTGTGCCAACTATGCT  
TTTATCTGCATACTTTGGCGCTGTTGGGCATATGTTTTTATGCAGGTGTTTGCAATATATTTTTGTTGG  
CGTGTAGCCCTTATTGTATAAGCCAAGTATCTGTCTTGCAAATATGTAACCATATACTTACTCA 7774

DF-228

7171

TGTTGTATGTATTGTATGGTTGTGCTTGTACTGTATGTTTTTGTGTATGTTTATGTATTTTATAAATTGT  
GTATGTGCTATGTGTATGTATGACTGTATGTATGTGTAATGTTTTGTGTGTATGTAATAACATGCA  
TGGTTACTTTTACGCGTGGTTGCATAAACTAAGGTGCGGTAGTATCCTTGGGCAGTGTGTGTCAGGT  
TAGGTGGTGTTCCTTACTGTTTAAATGTTATATTAATAGGTTGTTTGTATGCACTATAGTAACACAC  
CAAACCTCCATTTTAGTGCTGTACGCCATTTTATGCATGCAACCGAATTCGGTTGCCTAGCCTTTTGT  
CCTTATTTAAACCCAAAACGACTTTTCAGCAAAAACAGTTAATCCTTTGGCATATTGCCGTTTCCTGT  
TGTATGATTCAGGTATGTACACTGCCTTACCCTGTATTACTCACCTGTATTTCTGTGCCAACTATGCT  
TTTATCTGCATACTTTGGCGCTGTTGGGCATATGTTTTTATGCAGGTGTTTGCAATATATTTTTGTTGG  
CGTGTAGCCCTTATTGTATAAGCCAAGTATCTGTCTTGCAAATATGTAACCATATACTTACTCATT  
TACA 7772

## HPV-66 L1

DF-24

6691

ACATTAACATAAATATGATGCCCGTGAAATCAATCAATACCTTCGCCATGTGGAGGAATATGAACTA  
CAGTTTGTGTTTCAACTTTGTAAAATAACCTTAACTGCAGAAGTTATGGCATATTTGCATAAATATGA  
ATAATACTTTATTAGACGATTGGAATATTGGCTTATCCCCACCAGTTGCAACTAGCTTAGAGGATA  
AATATAGGTATATTAAGCACAGCTATTACATGTCAGAGGGAACAGCCCCCTGCAGAAAAGCAG  
GATCCCCTGGCTAAATATAAGTTTTGGGA 6982

DF-41

6551

AATAATGGCATATGCTGGGGTAATCAGGTATTTGTTACTGTTGTGGATACTACCAGAAGCACCAAC  
ATGACTATTAATGCAGCTAAAAGCACATTAATAAATATGATGCCCGTGAAATCAATCAATACCTT  
CGCCATGTGGAGGAATATGAACTACAGTTTGTGTTTCAACTTTGTAAAATAACCTTAACTGCAGAA  
GTTATGGCATATTTGCATAAATATGAATAATACTTTATTAGACGATTGGAATTGGCTTATCCCCACC  
AGTTGCAACTAGCTTAGAGGATAAATATAGGTATATTAAGCACAGCTATTACATGTCAGAGGGA  
ACAGCCCCCTGCAGAAAAGCAGGATCCCCTGGCTAAATATAAGTTTTGGGAAGTTAATTTACAGGA  
CAGTTTTTCTGCAGACCTGGATCAGTTT 6976

DF-107

6589

GCACAGGGTCATAAATAATGGCATATGCTGGGGTAATCAGGTATTTGTTACTGTTGTGGATACTACC  
AGAAGTACCAACATGACTATTAATGCAGCTAAAAGCACATTAATAAATATGATGCACGTGAAATC  
AATCAATACCTTCGCCATGTGGAGGAATATGAACTACAGTTTGTGTTTCAACTTTGTAAAATAACCT  
TAACTGCAGAAGTTATGGCATATTTGCATAAATATGAATAATACTTTATTAGACGATTGGAACATTG  
GATTGTCCCCACCAGTTGCAACTAGCTTAGAGGATAAATATAGGTATATTAAGCACAGCTATTA  
CATGTCAAAGGGAACAGCCCCCTGCAGAAAAGCAGGATCCCCTGGCTAAATATAAGTTTTGGGAG  
GTTAATTTACAGGACAGCTTTTTCTGCAGACCTAGATCAGTTTCCTTTGGGACG 7037

## DF-198

6602

ATAATGGCATATGCTGGGGTAATCAGGTATTTGTTACTGTTGTGGATACTACCAGAAGCACCAACA  
TGACTATTAATGCAGCTAAAAGCACATTA ACTAAATATGATGCCCGTGAAATCAATCAATACCTTC  
GCCATGTGGAGGAATATGAACTACAGTTTGTGTTTCAACTTTGTAAAATAACCTTAACTGCAGAAG  
TTATGGCATATTTGCATAATATGAATAATACTTTATTAGACGATTGGAATATTGGCTTATCCCCACC  
AGTTGCAACTAGCTTAGAGGATAAATATAGGTATATTAAGCACAGCTATTACATGTCAGAGGGA  
ACAGCCCCCTGCAGAAAAGCAGGATCCCCTGGCTAAATATAAGTTTTGGGAAGTTAATTTACAGGA  
CAGCTTTTCTGCAGACCTGG 7018

## DF-228

6584

GTTACTGTTGTGGATACTACCAGAAGCACCAACATGACTATTAATGCAGCTAAAAGCACATTA ACT  
AAATATGATGCCCGTGAAATCAATCAATACCTTCGCCATGTGGAGGAATATGAACTACAGTTTGTG  
TTTCAACTTTGTAAAATAACCTTAACTGCAGAAGTTATGGCATATTTGCATAATATGAATAATACTT  
TATTAGACGATTGGAATATTGGCTTATCCCCACCAGTTGCAACTAGCTTAGAGGATAAATATAGGT  
ATATTAAGCACAGCTATTACATGTCAAAGGGAACAGCCCCCTGCAGAAAAGCAGGATCCCCTG  
GCTAAATATAAGTTTTGGGAAGTTAATTTACAGGACAGCTTTTCTGCAGACCTGGATCAGTTTCC  
7028

## 10 ANEXO II

## Predição de sítios de ligação HPV-53 pelo programa MATCH™

matrix identifier	position (strand)	core match	matrix match	sequence (always the (+)-strand is shown)	factor name
<a href="#">V\$YY1 01</a>	19 (-)	1.000	0.911	gtgtatgtATGGTtgta	<a href="#">YY1</a>
<a href="#">V\$OCT1 Q6</a>	111 (-)	1.000	0.876	aatgTTTGCatggtt	<a href="#">Oct-1</a>
<a href="#">V\$YY1 01</a>	113 (+)	0.906	0.904	tgttTGCAAtgtgtaa	<a href="#">YY1</a>
<a href="#">V\$YY1 01</a>	127 (-)	0.922	0.899	taataaatATGAAtgag	<a href="#">YY1</a>
<a href="#">V\$MYCMAX 02</a>	152 (-)	0.901	0.923	ttacGCGTGgtt	<a href="#">c-Myc/Max</a>
<a href="#">V\$CEBPB 01</a>	159 (-)	0.970	0.957	tgGTTGCataaact	<a href="#">C/EBPbeta</a>
<a href="#">V\$CEBPB 02</a>	159 (+)	1.000	0.971	tggTTGCATAaact	<a href="#">C/EBPbeta</a>
<a href="#">V\$SP1 01</a>	173 (+)	0.924	0.910	aaGGCGCggt	<a href="#">Sp1</a>
<a href="#">V\$YY1 01</a>	200 (+)	0.904	0.895	gggtGGCATgttaggtg	<a href="#">YY1</a>
<a href="#">V\$SP1 01</a>	202 (+)	0.921	0.874	gtGGCATgtt	<a href="#">Sp1</a>
<a href="#">V\$YY1 01</a>	253 (-)	0.906	0.901	attaatatATGTAActgt	<a href="#">YY1</a>
<a href="#">V\$YY1 01</a>	310 (+)	1.000	0.977	ggtaTCCATtttagttt	<a href="#">YY1</a>
<a href="#">V\$YY1 01</a>	341 (+)	1.000	0.980	acccTCCATtttgcctt	<a href="#">YY1</a>
<a href="#">V\$E2 Q6</a>	359 (+)	0.984	0.964	caaccggtttCGGTTg	<a href="#">E2</a>
<a href="#">V\$E2 01</a>	359 (-)	0.986	0.983	cAACCggtttcggtt	<a href="#">E2</a>
<a href="#">V\$E2 Q6</a>	359 (-)	0.984	0.983	cAACCggtttcggtt	<a href="#">E2</a>
<a href="#">V\$E2 01</a>	359 (+)	0.986	0.965	caaccggtttCGGTTg	<a href="#">E2</a>
<a href="#">V\$OCT1 Q6</a>	379 (+)	0.909	0.808	tatcttGTAAAtacc	<a href="#">Oct-1</a>
<a href="#">V\$OCT1 02</a>	393 (-)	1.000	0.940	cacaaGCATAttcag	<a href="#">Oct-1</a>
<a href="#">V\$YY1 01</a>	393 (+)	0.906	0.893	cacaAGCATattcagca	<a href="#">YY1</a>
<a href="#">V\$OCT1 02</a>	396 (+)	0.992	0.937	aagcaTATTCagcag	<a href="#">Oct-1</a>
<a href="#">V\$OCT1 Q6</a>	420 (-)	0.883	0.802	tcctTTGGCatagtg	<a href="#">Oct-1</a>
<a href="#">V\$NF1 Q6</a>	422 (+)	1.000	0.984	ctTTGGCatagtgccgtt	<a href="#">NF-1</a>
<a href="#">V\$TATA C</a>	461 (-)	1.000	0.944	gTTTTAtaca	<a href="#">TATA</a>
<a href="#">V\$AP1 Q2</a>	483 (-)	1.000	0.985	tattAGTCAct	<a href="#">AP-1</a>
<a href="#">V\$NF1 Q6</a>	494 (-)	1.000	0.969	tactgtaggctGCCAAct	<a href="#">NF-1</a>
<a href="#">V\$YY1 01</a>	543 (+)	0.902	0.893	ggcaCACATttttattg	<a href="#">YY1</a>
<a href="#">V\$NF1 Q6</a>	580 (+)	1.000	0.981	ttTTGGCttgcagcatac	<a href="#">NF-1</a>
<a href="#">V\$CEBP C</a>	581 (-)	0.996	0.920	tttggCTTGCagcatacg	<a href="#">C/EBP</a>
<a href="#">V\$CEBPB 02</a>	584 (+)	1.000	0.901	ggcTTGCAGcatac	<a href="#">C/EBPbeta</a>
<a href="#">V\$NF1 Q6</a>	596 (-)	1.000	0.966	acgttctgctaGCCAAgt	<a href="#">NF-1</a>
<a href="#">V\$E2 Q6</a>	651 (-)	1.000	0.933	gCACCGcaaccgttta	<a href="#">E2</a>
<a href="#">V\$E2 Q6</a>	657 (-)	0.984	0.981	cAACCggtttacggttt	<a href="#">E2</a>
<a href="#">V\$E2 Q6</a>	657 (+)	0.984	0.962	caaccggtttaCGGTTt	<a href="#">E2</a>
<a href="#">V\$E2 01</a>	657 (-)	0.986	0.979	cAACCggtttacggttt	<a href="#">E2</a>
<a href="#">V\$E2 01</a>	657 (+)	0.986	0.966	caaccggtttaCGGTTt	<a href="#">E2</a>
<a href="#">V\$OCT1 Q6</a>	667 (-)	1.000	0.802	cggTTTGCagcaac	<a href="#">Oct-1</a>
<a href="#">V\$CEBPB 02</a>	669 (+)	1.000	0.896	gttTTGCAGcaaca	<a href="#">C/EBPbeta</a>
<a href="#">V\$CEBPB 02</a>	670 (-)	0.996	0.892	ttttgcAGCAAcag	<a href="#">C/EBPbeta</a>
<a href="#">V\$TATA C</a>	718 (-)	1.000	1.000	cTTTTAtaga	<a href="#">TATA</a>
<a href="#">V\$E2 01</a>	735 (-)	0.986	0.954	tAACCgaaataggttt	<a href="#">E2</a>
<a href="#">V\$E2 Q6</a>	735 (-)	0.984	0.958	tAACCgaaataggttt	<a href="#">E2</a>
<a href="#">V\$E2 Q6</a>	752 (-)	0.989	0.974	gGACCgaaaacggttac	<a href="#">E2</a>
<a href="#">V\$E2 01</a>	752 (+)	1.000	0.992	ggaccgaaaaCGGTAC	<a href="#">E2</a>
<a href="#">V\$E2 01</a>	752 (-)	0.980	0.968	gGACCgaaaacggttac	<a href="#">E2</a>
<a href="#">V\$E2 Q6</a>	752 (+)	0.998	0.997	ggaccgaaaaCGGTAC	<a href="#">E2</a>

```

1          <-----V$YY1_01(0.91)                                <-----...V$OCT1_Q6(0.88)
2          <-----...V$YY1_01(0.90)
atgttggtgctgtgattgtgtatggttgattatgtactatttattttgtgtatactgtatgtgtgtatttggctgtgtttgtatgactgtatgtatgtgtaatgtttgca 120
1-----V$OCT1_Q6(0.88)          <-----V$MYCMAX_02(0.92)          -----V$YY1_01(0.90)
2----->V$YY1_01(0.90)          <-----V$CEBPPB_01(0.96)          ----->V$SP1_01(0.87)
3          <-----V$YY1_01(0.90)          ----->V$SP1_01(0.91)
4          ----->V$CEBPPB_02(0.97)
tgttgtaataaatatgaatgagtggttacttttacgcgtggttcataaaactaaggcgcggtggtgtccctaggcagttgggtggcatggttagtggcgtccctgtttatataccaatct 240
1          <-----V$YY1_01(0.90)                                ----->V$YY1_01(0.98)          ---...V$E2_Q6(0.96)
2          ----->V$YY1_01(0.98)
3          <---...V$E2_01(0.98)
4          <---...V$E2_Q6(0.98)
5          ---...V$E2_01(0.96)
ccacctgtattattaatataatgtactgttatcatgcagttacctataaccacatatagtatagcttataggtatccatttagttggcgcctttttgttacccctccattttgccttgca 360
1----->V$E2_Q6(0.96)          <-----V$OCT1_Q2(0.94)                                <-----V$TATA_C(0.94)
2          ----->V$OCT1_Q6(0.81)          <-----V$OCT1_Q6(0.80)
3-----V$E2_01(0.98)          ----->V$YY1_01(0.89)
4-----V$E2_Q6(0.98)          ----->V$OCT1_Q2(0.94)
5----->V$E2_01(0.96)                                ----->V$NF1_Q6(0.98)
accggttcggtgtgcataatcctgtaataaccacaagcatattcagcagaactgtaatacctttggcatagtgccgtttcctgtgttaatgtttttttatacactatgcotta 480
1          <-----V$NF1_Q6(0.97)                                ----->V$YY1_01(0.89)          ----->V$NF1_Q6(0.98)
2 <-----V$AP1_Q2(0.98)          ----->V$NF1_Q6(0.97)          ----->V$CEBP_C(0.92)
3          ----->V$CEBPPB_02(0.90)
4          <-----...V$NF1_Q6(0.97)
cctattagtcacttactgttaggctgccaactatgcttttacctgcataccttagtgcttttggcacacattttattgcaaacatctgcaacttagtgcttttggcttgagcagcagtt 600
1          <-----V$E2_Q6(0.93)                                <---...V$TATA_C(1.00)
2          <-----V$E2_Q6(0.98)
3          ----->V$E2_Q6(0.96)
4-----V$NF1_Q6(0.97)          <-----V$E2_01(0.98)
5          ----->V$E2_01(0.97)
6          <-----V$OCT1_Q6(0.80)
7          ----->V$CEBPPB_02(0.90)
8          <-----V$CEBPPB_02(0.89)
ctgctagccaagtatctgtctagtaaacagggtgtgcacctattactcatgcaccgcaaccggttttacggttttgagcaaacaggcctttttataattattgaaagtaacaatcctactt 720
1-----V$TATA_C(1.00)          <-----V$E2_Q6(0.97)
2          <-----V$E2_01(0.95)
3          <-----V$E2_Q6(0.96)
4          ----->V$E2_01(0.99)
5          <-----V$E2_01(0.97)
6          ----->V$E2_Q6(1.00)
ttatagacagggagtaaccgaaataggttttaggaccgaaaacggtacatataaaagcactgtgtacaacaccaggacatcc 802

```

## Predição de sítios de ligação HPV-56 pelo programa MATCH™

matrix identifier	position (strand)	core match	matrix match	sequence (always the (+)-strand is shown)	factor name
<a href="#">V\$CEBPB 02</a>	13 (-)	0.888	0.895	tgtttgTGTAActg	<a href="#">C/EBPbeta</a>
<a href="#">V\$YY1 01</a>	37 (-)	1.000	0.977	ttgtatatATGGTatgt	<a href="#">YY1</a>
<a href="#">V\$TATA C</a>	69 (-)	1.000	0.944	aTTTTAtact	<a href="#">TATA</a>
<a href="#">V\$OCT1 Q6</a>	74 (-)	0.838	0.842	atacTTGTatgtgt	<a href="#">Oct-1</a>
<a href="#">V\$CEBPB 02</a>	95 (-)	0.888	0.893	tgtttgTGTAAtg	<a href="#">C/EBPbeta</a>
<a href="#">V\$OCT1 Q6</a>	96 (+)	0.909	0.802	gtttGTAAAtggt	<a href="#">Oct-1</a>
<a href="#">V\$YY1 01</a>	98 (-)	0.906	0.900	ttgtgtaaATGTTtgtg	<a href="#">YY1</a>
<a href="#">V\$YY1 01</a>	180 (-)	0.922	0.898	taataaacATGAAtgag	<a href="#">YY1</a>
<a href="#">V\$MYCMax 02</a>	202 (-)	0.901	0.923	ttacGCGTGgtt	<a href="#">c-Myc/Max</a>
<a href="#">V\$CEBPB 01</a>	209 (-)	0.970	0.957	tgGTTGCataaact	<a href="#">C/EBPbeta</a>
<a href="#">V\$CEBPB 02</a>	209 (+)	1.000	0.971	tggTTGCataaact	<a href="#">C/EBPbeta</a>
<a href="#">V\$CEBPB 01</a>	246 (+)	0.965	0.957	ttgttttGTAAGtt	<a href="#">C/EBPbeta</a>
<a href="#">V\$YY1 01</a>	287 (+)	0.906	0.895	tgcaTACATatataatac	<a href="#">YY1</a>
<a href="#">V\$YY1 01</a>	310 (+)	1.000	0.965	atacTCCATtttgtgt	<a href="#">YY1</a>
<a href="#">V\$YY1 01</a>	329 (+)	0.998	0.965	ttccGCCATtttgtaca	<a href="#">YY1</a>
<a href="#">V\$YY1 01</a>	337 (-)	0.906	0.894	ttttgtacATGCAaccg	<a href="#">YY1</a>
<a href="#">V\$E2 01</a>	348 (+)	0.986	0.965	caaccgaattCGGTTg	<a href="#">E2</a>
<a href="#">V\$E2 01</a>	348 (-)	0.986	0.965	cAACCGaattcggttg	<a href="#">E2</a>
<a href="#">V\$E2 Q6</a>	348 (+)	0.984	0.966	caaccgaattCGGTTg	<a href="#">E2</a>
<a href="#">V\$E2 Q6</a>	348 (-)	0.984	0.966	cAACCGaattcggttg	<a href="#">E2</a>
<a href="#">V\$YY1 01</a>	371 (+)	0.998	0.915	tagtGCCATtatttaaa	<a href="#">YY1</a>
<a href="#">V\$YY1 01</a>	413 (+)	0.998	0.915	tagtGCCATtatttaaa	<a href="#">YY1</a>
<a href="#">V\$OCT1 Q6</a>	460 (-)	0.883	0.806	tcctTTGGCatattg	<a href="#">Oct-1</a>
<a href="#">V\$OCT1 02</a>	462 (-)	1.000	0.911	ctttgGCATAttgcc	<a href="#">Oct-1</a>
<a href="#">V\$NF1 Q6</a>	462 (+)	1.000	0.984	ctTTGGCatattgccgtt	<a href="#">NF-1</a>
<a href="#">V\$TATA C</a>	486 (-)	1.000	0.944	gTTTTAtact	<a href="#">TATA</a>
<a href="#">V\$NF1 Q6</a>	534 (-)	1.000	0.964	gtactatgactGCCAAct	<a href="#">NF-1</a>
<a href="#">V\$YY1 01</a>	583 (+)	0.902	0.903	ggcaCACATttttatac	<a href="#">YY1</a>
<a href="#">V\$OCT1 Q6</a>	587 (+)	0.928	0.941	cacatTTTTAtaca	<a href="#">Oct-1</a>
<a href="#">V\$TATA C</a>	591 (-)	1.000	0.947	tTTTTAtaca	<a href="#">TATA</a>
<a href="#">V\$YY1 01</a>	592 (-)	0.902	0.895	ttttatacATGTGtctg	<a href="#">YY1</a>
<a href="#">V\$NF1 Q6</a>	620 (+)	1.000	0.966	ttTTGGCttgcagaatac	<a href="#">NF-1</a>
<a href="#">V\$CEBP C</a>	621 (-)	0.996	0.931	tttggCTTGcagaatac	<a href="#">C/EBP</a>
<a href="#">V\$CEBPB 01</a>	624 (-)	0.985	0.963	ggCTTGcagaatac	<a href="#">C/EBPbeta</a>
<a href="#">V\$CEBPB 02</a>	624 (+)	1.000	0.957	ggCTTGcagaatac	<a href="#">C/EBPbeta</a>
<a href="#">V\$CEBPB 01</a>	666 (+)	0.976	0.956	ctgttttGCAAACA	<a href="#">C/EBPbeta</a>
<a href="#">V\$CEBPB 02</a>	666 (-)	1.000	0.973	ctgttttGCAAACA	<a href="#">C/EBPbeta</a>
<a href="#">V\$OCT1 Q6</a>	666 (-)	1.000	0.843	ctgtTTTGCAaacat	<a href="#">Oct-1</a>
<a href="#">V\$YY1 01</a>	671 (-)	0.906	0.896	ttgcaaacATGTAacat	<a href="#">YY1</a>
<a href="#">V\$YY1 01</a>	691 (+)	0.918	0.913	attaCTCATtttttaaa	<a href="#">YY1</a>
<a href="#">V\$E2 Q6</a>	706 (-)	0.984	0.976	aAACCGtttacgggtcg	<a href="#">E2</a>
<a href="#">V\$E2 01</a>	706 (+)	0.980	0.964	aaaccgtttaCGGTCg	<a href="#">E2</a>
<a href="#">V\$E2 Q6</a>	706 (+)	0.989	0.969	aaaccgtttaCGGTCg	<a href="#">E2</a>
<a href="#">V\$E2 01</a>	706 (-)	0.986	0.978	aAACCGtttacgggtcg	<a href="#">E2</a>
<a href="#">V\$CEBPB 02</a>	716 (-)	1.000	0.913	cgggtcgTGCAAAAA	<a href="#">C/EBPbeta</a>
<a href="#">V\$OCT1 Q6</a>	717 (+)	1.000	0.801	gggtcgTGCAAAAA	<a href="#">Oct-1</a>
<a href="#">V\$OCT1 Q6</a>	761 (+)	0.928	0.933	catacTTTTAtata	<a href="#">Oct-1</a>
<a href="#">V\$E2 01</a>	782 (-)	0.980	0.952	tGACCGaaaagggttt	<a href="#">E2</a>
<a href="#">V\$E2 Q6</a>	782 (-)	0.989	0.960	tGACCGaaaagggttt	<a href="#">E2</a>
<a href="#">V\$E2 Q6</a>	799 (-)	0.989	0.964	aGACCGaaaacgggtac	<a href="#">E2</a>
<a href="#">V\$E2 Q6</a>	799 (+)	0.998	0.996	agaccgaaaaCGGTAC	<a href="#">E2</a>
<a href="#">V\$E2 01</a>	799 (+)	1.000	0.987	agaccgaaaaCGGTAC	<a href="#">E2</a>
<a href="#">V\$E2 01</a>	799 (-)	0.980	0.966	aGACCGaaaacgggtac	<a href="#">E2</a>

```

1      <-----V$CEBPB_02(0.90)      <-----V$TATA_C(0.94)  ----->V$OCT1_Q6(0.80)
2      <-----V$YY1_01(0.98)      <-----V$OCT1_Q6(0.84)
3      <-----V$CEBPB_02(0.89)
4      <-----V$YY1_01(0.90)
TgtgtgtgtgtgtgtttgtgtaactgtgtttgtgtgtgtatatatggatgtttgtgtaTgtgctttattttatactttgtatgtgtatgtttgtgtttgtgtaaatgtttgtgtgaaat  120
1      <-----V$YY1_01(0.90)
2      <-----V$MYCMAx_02(0.92)
3      <-----V$CEBPB_01(0.96)
4      ----->V$CEBPB_02(0.97)
GtttgtgtgtgtattcattgtatgtatgactgtatatatgtgtaatgtttgtgtgtctgtAataaacatgaatgagtgcttttacgcgtgttcataaaactaagggtgtgtcattattgt  240
1      ----->V$CEBPB_01(0.96)      ----->V$YY1_01(0.90)      ----->V$YY1_01(0.96)      ----->V$YY1_01(0.96)
2      ----->V$YY1_01(0.96)      ----->V$YY1_01(0.96)      ----->V$YY1_01(0.96)      ----->V$YY1_01(0.96)
3      ----->V$YY1_01(0.96)      ----->V$YY1_01(0.96)      ----->V$YY1_01(0.96)      ----->V$YY1_01(0.96)
4      ----->V$YY1_01(0.96)      ----->V$YY1_01(0.96)      ----->V$YY1_01(0.96)      ----->V$YY1_01(0.96)
5      ----->V$YY1_01(0.96)      ----->V$YY1_01(0.96)      ----->V$YY1_01(0.96)      ----->V$YY1_01(0.96)
6      ----->V$YY1_01(0.96)      ----->V$YY1_01(0.96)      ----->V$YY1_01(0.96)      ----->V$YY1_01(0.96)
GcgttttgtttgtgaaattattgtgtacagtgtactatgtgtattgtgcatacatatataTaccataacatactccattttgttgttttccgccattttgtacatgcaaccgaattcgg  360
1      ----->V$YY1_01(0.92)      ----->V$YY1_01(0.92)      ----->V$YY1_01(0.92)      ----->V$YY1_01(0.92)
2-->V$E2_Q6(0.97)      ----->V$YY1_01(0.92)      ----->V$YY1_01(0.92)      ----->V$YY1_01(0.92)
3      ----->V$YY1_01(0.92)      ----->V$YY1_01(0.92)      ----->V$YY1_01(0.92)      ----->V$YY1_01(0.92)
4---V$E2_Q6(0.97)
5-->V$E2_Q6(0.97)
6---V$E2_Q6(0.97)
TtgcattggtcattatttaaactaaaaggaattcgggtgcatggcctagtgccaTattttaaaccaaaagcccttttccagcagaacagttaatccctttggcatattgcccgttt  480
1      <-----V$NF1_Q6(0.96)      ----->V$YY1_01(0.90)      ----->V$YY1_01(0.90)
2      ----->V$YY1_01(0.90)      ----->V$YY1_01(0.90)      ----->V$YY1_01(0.90)      ----->V$YY1_01(0.90)
3      ----->V$YY1_01(0.90)      ----->V$YY1_01(0.90)      ----->V$YY1_01(0.90)      ----->V$YY1_01(0.90)
4      <-----V$TATA_C(0.94)      ----->V$YY1_01(0.90)      ----->V$YY1_01(0.90)      ----->V$YY1_01(0.90)
CctgtgttttatacttgaattatgtacagtaccgcaccctgtattactcacaggtactatGactgccaactatgcttttatctgcatacttttagtgctgttgggacacattttataca  600
1      ----->V$NF1_Q6(0.97)      ----->V$CEBPB_01(0.96)      ----->V$CEBPB_01(0.96)      ----->V$CEBPB_01(0.96)
2      <-----V$CEBPB_C(0.93)      <-----V$CEBPB_02(0.97)      ----->V$YY1_01(0.90)      ----->V$YY1_01(0.90)
3      <-----V$CEBPB_01(0.96)      <-----V$OCT1_Q6(0.84)      ----->V$YY1_01(0.90)      ----->V$YY1_01(0.90)
4-----V$YY1_01(0.90)      ----->V$CEBPB_02(0.96)      ----->V$YY1_01(0.91)      ----->V$YY1_01(0.91)
5      ----->V$YY1_01(0.91)      ----->V$YY1_01(0.91)      ----->V$YY1_01(0.91)      ----->V$YY1_01(0.91)
6      ----->V$YY1_01(0.91)      ----->V$YY1_01(0.91)      ----->V$YY1_01(0.91)      ----->V$YY1_01(0.91)
7      ----->V$YY1_01(0.91)      ----->V$YY1_01(0.91)      ----->V$YY1_01(0.91)      ----->V$YY1_01(0.91)
TgtgtctgcaactttggtgttttgcttgcagaatacactatgtaggccaagtatctgtcAgatctgttttgcacaatgtaacatacaattactcattttttaaaaccggtttacggtc  720
1-V$E2_Q6(0.98)      ----->V$OCT1_Q6(0.93)      ----->V$E2_Q6(0.96)      ----->V$E2_Q6(0.96)
2>V$E2_Q6(0.96)      ----->V$E2_Q6(0.96)      ----->V$E2_Q6(0.96)      ----->V$E2_Q6(0.96)
3>V$E2_Q6(0.97)      ----->V$E2_Q6(0.96)      ----->V$E2_Q6(0.96)      ----->V$E2_Q6(0.96)
4-V$E2_Q6(0.98)      ----->V$E2_Q6(1.00)      ----->V$E2_Q6(1.00)      ----->V$E2_Q6(1.00)
5      ----->V$E2_Q6(1.00)      ----->V$E2_Q6(1.00)      ----->V$E2_Q6(1.00)      ----->V$E2_Q6(1.00)
6-----V$CEBPB_02(0.91)      ----->V$E2_Q6(1.00)      ----->V$E2_Q6(1.00)      ----->V$E2_Q6(1.00)
7----->V$OCT1_Q6(0.80)      ----->V$E2_Q6(1.00)      ----->V$E2_Q6(1.00)      ----->V$E2_Q6(1.00)
GtgcaaaacaggtttcttttaattgttgaagtttcaatcatacttttatatttggaGtgaccgaaaagggtttaaagaccgaaaacggtacataaaaaggcagcttattctgtgtg  840
Gacatatcc  849

```



## Predição de sítios de ligação HPV-66 pelo programa MATCH™

matrix identifier	position (strand)	core match	matrix match	sequence (always the (+)-strand is shown)	factor name
<a href="#">V\$YY1_01</a>	20 (-)	1.000	0.893	tgtattgtATGGTtgtg	<a href="#">YY1</a>
<a href="#">V\$YY1_01</a>	136 (-)	0.906	0.891	taataaacATGCAtggt	<a href="#">YY1</a>
<a href="#">V\$YY1_01</a>	140 (-)	1.000	0.895	aaacatgcATGGTtact	<a href="#">YY1</a>
<a href="#">V\$MYCMAX_02</a>	158 (-)	0.901	0.923	ttacGCGTGggt	<a href="#">c-Myc/Max</a>
<a href="#">V\$CEBPB_02</a>	165 (+)	1.000	0.971	tggTTGCAtaaact	<a href="#">C/EBPbeta</a>
<a href="#">V\$CEBPB_01</a>	165 (-)	0.970	0.957	tgTTGCAtaaact	<a href="#">C/EBPbeta</a>
<a href="#">V\$SP1_01</a>	197 (+)	0.972	0.891	tgGGCAGtgt	<a href="#">Sp1</a>
<a href="#">V\$OCT1_Q6</a>	255 (-)	0.838	0.826	gttgTTGTatgcac	<a href="#">Oct-1</a>
<a href="#">V\$YY1_01</a>	284 (+)	1.000	0.977	aaacTCCATtttagtgc	<a href="#">YY1</a>
<a href="#">V\$YY1_01</a>	302 (+)	0.998	0.967	gtacGCCATtttatgca	<a href="#">YY1</a>
<a href="#">V\$E2_Q6</a>	321 (-)	0.984	0.966	cAACCGaattcggttg	<a href="#">E2</a>
<a href="#">V\$E2_01</a>	321 (+)	0.986	0.965	caaccgaattCGGTTg	<a href="#">E2</a>
<a href="#">V\$E2_Q6</a>	321 (+)	0.984	0.966	caaccgaattCGGTTg	<a href="#">E2</a>
<a href="#">V\$E2_01</a>	321 (-)	0.986	0.965	cAACCGaattcggttg	<a href="#">E2</a>
<a href="#">V\$CEBPB_02</a>	371 (-)	0.996	0.947	cttttcAGCAAaac	<a href="#">C/EBPbeta</a>
<a href="#">V\$OCT1_Q6</a>	391 (-)	0.883	0.806	tcctTTGGCAtattg	<a href="#">Oct-1</a>
<a href="#">V\$OCT1_Q2</a>	393 (-)	1.000	0.911	ctttgGCATAttgcc	<a href="#">Oct-1</a>
<a href="#">V\$NF1_Q6</a>	393 (+)	1.000	0.984	ctTTGGCAtattgccggtt	<a href="#">NF-1</a>
<a href="#">V\$SP1_01</a>	435 (-)	0.972	0.876	acaCTGCctt	<a href="#">Sp1</a>
<a href="#">V\$NF1_Q6</a>	462 (-)	1.000	0.964	tgtatttctgtGCCAAact	<a href="#">NF-1</a>
<a href="#">V\$NF1_Q6</a>	547 (+)	1.000	0.990	tgTTGGCgttagccctt	<a href="#">NF-1</a>
<a href="#">V\$NF1_Q6</a>	563 (-)	1.000	0.961	ttattgtataaGCCAAgt	<a href="#">NF-1</a>
<a href="#">V\$OCT1_Q6</a>	584 (+)	1.000	0.849	tgtcttGCAAAtatg	<a href="#">Oct-1</a>
<a href="#">V\$YY1_01</a>	588 (-)	0.906	0.906	ttgcaaatATGTAacca	<a href="#">YY1</a>
<a href="#">V\$YY1_01</a>	597 (+)	1.000	0.976	tgtaACCATatacttac	<a href="#">YY1</a>
<a href="#">V\$YY1_01</a>	609 (+)	0.918	0.912	cttaCTCATtttataaaa	<a href="#">YY1</a>
<a href="#">V\$CEBPB_02</a>	616 (+)	0.888	0.898	attTTACAAAaccg	<a href="#">C/EBPbeta</a>
<a href="#">V\$E2_Q6</a>	624 (+)	0.989	0.969	aaaccgtttaCGGTCg	<a href="#">E2</a>
<a href="#">V\$E2_01</a>	624 (+)	0.980	0.964	aaaccgtttaCGGTCg	<a href="#">E2</a>
<a href="#">V\$E2_Q6</a>	624 (-)	0.984	0.976	aAACCGtttacggtcg	<a href="#">E2</a>
<a href="#">V\$E2_01</a>	624 (-)	0.986	0.978	aAACCGtttacggtcg	<a href="#">E2</a>
<a href="#">V\$E2_Q6</a>	700 (-)	0.984	0.958	tAACCGaaatgggttt	<a href="#">E2</a>
<a href="#">V\$YY1_01</a>	700 (-)	0.996	0.954	taaccgaaATGGGttta	<a href="#">YY1</a>
<a href="#">V\$E2_01</a>	700 (-)	0.986	0.955	tAACCGaaatgggttt	<a href="#">E2</a>
<a href="#">V\$E2_Q6</a>	717 (+)	0.998	0.997	ggaccgaaaaCGGTAc	<a href="#">E2</a>
<a href="#">V\$E2_01</a>	717 (-)	0.980	0.968	gGACCGaaaaacggtac	<a href="#">E2</a>
<a href="#">V\$E2_Q6</a>	717 (-)	0.989	0.974	gGACCGaaaaacggtac	<a href="#">E2</a>
<a href="#">V\$E2_01</a>	717 (+)	1.000	0.992	ggaccgaaaaCGGTAc	<a href="#">E2</a>

```

1          <-----V$YY1_01(0.89)
TgtgtgtgtgtgtgtgtgtgtatgtattgtatggttgtgcttgtactgtatgtttttgtgtatGtttatgtattttataattgtgtatgtgctatgtgtatgtatgactgtatgtatgttaa 120
1          <-----V$YY1_01(0.89)          ----->V$SP1_01(0.89)
2          <-----V$YY1_01(0.90)
3          <-----V$MYCMAX_02(0.92)
4          ----->V$CEBPPB_02(0.97)
5          <-----V$CEBPPB_01(0.96)
TgtttttgtgtatgtaataaacatgcatggttacttttacgcggtggttgcataaaactaaGgtgcggttagtatccttgggcagtggtgtgcaggttaggtggtgttccttactgtttaat 240
1          <-----V$OCT1_Q6(0.83)          ----->V$YY1_01(0.97)
2          ----->V$YY1_01(0.98)          <-----V$E2_Q6(0.97)
3          ----->V$E2_Q1(0.96)
4          ----->V$E2_Q6(0.97)
5          <-----V$E2_Q1(0.96)
GttatattaataggttgtttgtatgcaactatagtaacacaccaaactccatthtagtgcTgtacgccatthtagcatgcaaccgaattcggttgccctagcctthtgccttattttaa 360
1          <-----V$CEBPPB_02(0.95)          ----->V$SP1_01(0.88)          <-----V$NF1_Q6(0.96)
2          <-----V$OCT1_Q6(0.81)
3          <-----V$OCT1_Q2(0.91)
4          ----->V$NF1_Q6(0.98)
CcaaaaacgacttttcagcaaaaacagttaatcctttggcatattgccgtttcctgttgtaTgattcaggtatgtacactgccttacctgtattactcacctgtatttctgtgccaacta 480
1          ----->V$NF1_Q6(0.99)          ----->V$OCT1_Q6(0.85)
2          <-----V$NF1_Q6(0.96)
3          <-----...V$YY1_01(0.91)
4          -----...V$YY1_01(0.98)
TgtttttatctgcatactttggcgctgttggcagatgtttttatgcaggtgtttgcaatAtatthtggcgtgtagcccttattgtataagccaagtatctgtcttgcaaatatgta 600
1          ----->V$E2_Q6(0.97)          <-----V$E2_Q6(0.96)
2          ----->V$YY1_01(0.91)          <-----V$YY1_01(0.95)
3-----V$YY1_01(0.91)          ----->V$E2_Q1(0.96)          <-----V$E2_Q1(0.95)
4----->V$YY1_01(0.98)          -----...V$E2_Q6(1.00)
5          ----->V$CEBPPB_02(0.90)          <-----...V$E2_Q1(0.97)
6          <-----V$E2_Q6(0.98)          <-----...V$E2_Q6(0.97)
7          <-----V$E2_Q1(0.98)          -----...V$E2_Q1(0.99)
AccatatacttactcattttacaaaacggtttacggctgtgctaaaaacaggtttcttttaAttgttgaaagtttcaatcatactttattatattgggagtaaccgaaatgggtttaggac 720
1
2
3
4----->V$E2_Q6(1.00)
5-----V$E2_Q1(0.97)
6-----V$E2_Q6(0.97)
7----->V$E2_Q1(0.99)
cgaaaacggtacatataaaaaggcagcctgttgcctgtagatatcc 767

```