



**Ministério da Saúde
Instituto Nacional de Câncer
Coordenação de Pós-graduação**

INSTITUTO NACIONAL DE CÂNCER

Pós-Graduação em Oncologia

SÉRGIO MENEZES AMARO FILHO

DIVERSIDADE DE METILAÇÃO NO DNA DE HPV16 E HPV18 EM CARCINOMAS
CERVICAIS INVASIVOS

Orientador (es): Prof. Dr. Miguel Ângelo Martins Moreira
Prof. Dr. Robert Burk
Prof. Dr. Cibele Rodrigues Bonvicino

RIO DE JANEIRO

2017



**Ministério da Saúde
Instituto Nacional de Câncer
Coordenação de Pós-graduação**

INSTITUTO NACIONAL DE CÂNCER

Pós-Graduação em Oncologia

SÉRGIO MENEZES AMARO FILHO

Diversidade de metilação no DNA de HPV16 e HPV18 em carcinomas cervicais invasivos

Tese apresentada ao Instituto Nacional de Câncer como
parte dos requisitos para obtenção do título de Doutor
em Oncologia

Orientador (es): Prof. Dr. Miguel Ângelo Martins Moreira
Prof. Dr. Robert Burk
Prof. Dr. Cibele Rodrigues Bonvicino

RIO DE JANEIRO

2017

A485d Amaro Filho, Sérgio Menezes.

Diversidade de metilação no DNA de HPV16 e HPV18 em carcinomas cervicais invasivos / Sérgio Menezes Amaro Filho. – Rio de Janeiro, 2017. 113 f.: il. color.

Tese (Doutorado em Oncologia) – Instituto Nacional de Câncer José Alencar Gomes da Silva, 2017.

Orientadores: Miguel Ângelo Martins Moreira; Robert Burk; Cibele Rodrigues Bonvicino.

1. Papillomaviridae. 2. Metilação. 3. Neoplasias do Colo do Útero. I. Moreira, Miguel Ângelo Martins (Orient.). II. Burk, Robert (Orient.). III. Bonvicino, Cibele Rodrigues (Orient.). IV. Instituto Nacional de Câncer José Alencar Gomes da Silva. V. Título.

CDD 616.911



**Ministério da Saúde
Instituto Nacional de Câncer
Coordenação de Pós-graduação**

**INSTITUTO NACIONAL DE CÂNCER
Pós-Graduação em Oncologia**

AUTOR: SÉRGIO MENEZES AMARO FILHO

**DIVERSIDADE DE METILAÇÃO NO DNA DE HPV16 E HPV18 EM CARCINOMAS
CERVICAIS INVASIVOS**

**ORIENTADOR (ES): Prof. Dr. Miguel Ângelo Martins Moreira
Prof. Dr. Robert Burk
Prof. Dr. Cibele Rodrigues Bonvicino**

Aprovada em: ____/____/____

EXAMINADORES:

**Prof. Dr. Marcelo Alves Soares
Prof. Dra. Cláudia Esther Alicia Rocio Hassan
Prof. Dra. Sheila Coelho Soares Lima
Prof. Dra. Laura Sichero
Prof. Dra. (Suplente I) Patrícia Sávio de Araújo Souza
Prof. Dra. (Suplente II) Alcina Frederica Nicol**

**RIO DE JANEIRO
2017**



**Ministério da Saúde
Instituto Nacional de Câncer
Coordenação de Pós-graduação**

INSTITUTO NACIONAL DE CÂNCER

DIVERSIDADE DE METILAÇÃO NO DNA DE HPV16 E HPV18 EM CARCINOMAS CERVICAIS INVASIVOS

RESUMO

TESE DE DOUTORADO

Sérgio Menezes Amaro Filho

A metilação de citosinas em sítios CpG é uma das principais modificações epigenéticas e está frequentemente associada à repressão transcricional. Apesar de estudos demonstrarem que o promotor dos oncogenes *E6/E7* apresenta modulação através da metilação e que existe aumento progressivo nos níveis de metilação em sítios CpG específicos durante a progressão do câncer cervical, pouco se sabe sobre o papel da metilação na regulação da transcrição viral e como características clínicas e biológicas do carcinoma cervical invasivo podem influenciar a metilação. O objetivo principal do trabalho foi caracterizar, em amostras de tumores cervicais invasivos, a diversidade dos níveis de metilação em sítios CpG presentes na *LCR*, *L2* e *L1* de HPV16 e HPV18. Os objetivos específicos: (i) caracterizar os níveis de metilação em sítios CpG na *LCR* de HPV16 e HPV18 por pirosequenciamento; (ii) caracterizar os níveis de metilação em sítios CpG na *LCR*, *L2* e *L1* de HPV16 e HPV18 por NGS; (iii) associar os níveis de metilação às características clínicas e biológicas da doença; e (iv) caracterização dos haplótipos de metilação de HPV16 e HPV18. As amostras foram obtidas de pacientes atendidas no ambulatório do INCA e diagnosticadas com câncer cervical invasivo. Foram selecionadas 229 amostras, sendo 165 infectadas com HPV16 e 64 com HPV18. PCR combinando iniciadores para os genes *E1/E2* dos HPVs 16 e 18 foi utilizada para avaliar o estado físico do DNA viral. Os níveis de metilação foram determinados por conversão por bissulfito sódio das amostras de DNA, seguido de amplificação da região de interesse, *LCR*, *L2* e *L1*, e análise por pirosequenciamento ou NGS. Informações clínicas e biológicas da doença foram extraídas de questionários, banco de dados e estudos prévios. Foram observados maiores níveis de metilação em sítios CpG presentes em *L2* e *L1* de HPV16 e HPV18 quando comparados à *LCR*. Entretanto, o DNA de HPV16 apresentou maior flutuação nos níveis de metilação em sítios CpG próximos e maior amplitude de metilação por amostra em sítios CpG na *LCR*, quando comparado a HPV18. Modelo estatístico utilizando análise bivariada e multivariada demonstrou, por pirosequenciamento, que presença de *E1/E2* e pacientes com idades avançadas estiveram associadas, de forma independente, a maiores níveis de metilação em 3'*LCR* de HPV16 ($p = 0,002$ e $p = 0,021$, respectivamente). Utilizando a mesma estratégia de análise para os dados obtidos por NGS, foi observado que os níveis de metilação nos sítios CpG 5602, 5608 e 5611 presentes em *L1* de HPV16 podem sofrer a influência de fatores como o grau de diferenciação do tumor, tipo histológico, linhagem de HPV e carga viral. Já os níveis de metilação no CpG 7862, localizado na *LCR* de HPV16, está associado ao tipo de tumor, e positivamente à carga viral. Os achados sugerem a metilação no DNA de HPV, principalmente HPV16, é um processo complexo, em que fatores externos podem desempenhar papel fundamental nos níveis de metilação desse vírus e estudos prospectivos e com maior número amostras devem ser feitos a fim de explorar a importância dos achados para o desfecho clínico.

RIO DE JANEIRO

2017



**Ministério da Saúde
Instituto Nacional de Câncer
Coordenação de Pós-graduação**

INSTITUTO NACIONAL DE CÂNCER

**DIVERSITY OF HPV16 AND HPV18 DNA METHYLATION IN INVASIVE
CERVICAL CARCINOMAS**

ABSTRACT

TESE DE DOUTORADO

Sérgio Menezes Amaro Filho

DNA methylation plays one of the most important epigenetic modifications and often leads to repression of transcription. Despite studies that show oncogene promoters of HPV16 (P97) being modulated by methylation and that methylation in specific CpG sites is related to severity of cervical neoplasia, it is still unclear how methylation contributes to the viral transcription and how clinical and biological factors related with HPV can affect the methylation pattern in invasive cervical cancer (ICC). The main goal was to characterize diversity of methylation level in CpG sites of LCR, L2 and L1 of HPV16 and HPV18 in invasive cervical cancer samples. Specific goals: (i) characterize level of methylation in CpG sites at LCR of HPV16 and HPV18 by pyrosequencing; (ii) characterize level of methylation in CpG sites at LCR, L2 and L1 of HPV16 and HPV18 by NGS; (iii) associate level of methylation with clinical and biological tumor characteristics; and (iv) characterize the methylation haplotype of HPV16 and HPV18. The 229 selected samples, being 165 with HPV16 and 64 with HPV18, were obtained from biopsies of patients attended at INCA ambulatory and diagnosed with invasive cervical cancer (ICC). Level of methylation was assessed by means of bisulfite treatment followed by PCR and pyrosequencing or NGS. PCR combining pairs of primers was performed to assess integrity of E1 and E2 genes for HPV16 and HPV18. CaSki and HeLa cell lines were used as methylation control. Patient and clinical information were obtained from questionnaires, database and previous studies. The methylation level was higher in L2 and L1 CpG sites of HPV16 and HPV18 DNA in comparison with LCR. However, HPV16 DNA displayed a wider range of methylation per sample, and per CpG site nearby in each segment analyzed, when compared to HPV18. Statistical model showed, by means of bivariate and multivariate analysis, that samples with non-disrupted E1/E2 and older patients were associated with higher methylation levels on CpG sites at LCR of HPV16 ($p = 0.002$ e $p = 0.021$, respectively). With the same model, was showed that methylation on CpG 5602, 5608 and 5611 of L1 of HPV16 could be affected by different factors, including tumor grade, type of tumor, HPV lineages and viral load. Level of methylation on CpG 7862, at LCR of HPV16, is also able to be affected by tumor type, and positively associated with viral load. These findings suggest methylation at HPV DNA, mainly on HPV16, as a complex event, which many factors may perform an important role on methylation of HPV DNA, and new prospective studies should be done with a larger sample size in order to explore the importance of these finding for the clinical outcome.

RIO DE JANEIRO

2017

Se, a princípio, a ideia não é absurda, então não há esperança para ela.
Albert Einstein

AGRADECIMENTOS

Agradeço imensamente a todos do Instituto Nacional do Câncer (INCA), incluindo membros da pós-graduação, coordenadores, diretores, pesquisadores, funcionários da recepção e manutenção que vem prestando um serviço de excelência, com competência e satisfação, e que foram fundamentais para esse instituto se tornar uma extensão da minha casa.

Gostaria de agradecer em especial:

À Dra. Cibele Bonvicino pela oportunidade em ingressar no Departamento de Genética do INCA em 2012, por incentivar a minha entrada no doutorado e disponibilidade todo o tempo.

Ao meu orientador Dr. Miguel Ângelo Moreira por toda a paciência, incentivo e parceria durante esses quatro anos. Foram anos de crescimento, amadurecimento e muitas discussões produtivas sobre a biologia do HPV.

Ao grupo de alunos chefiado pelo Miguel que sempre contribuíram com ideias em nossas reuniões semanais: João Vidal, Ayslan Brant, Diogo Lisboa, Shayany Felix, Renan Gomes, Bárbara Soares, Marion Amazonas e Mateus Mota.

Aos amigos de trabalho que estiveram no laboratório nesse período, sempre dispostos a ajudar, compartilhando conhecimento e momentos diários, com muito café e descontração. Em especial: Michele Mellin, Carolina Furtado, Maria Carolina, Ana Lazar, Hanna Soares, Juliana Siqueira, Pricila Spínola e muitos outros.

Aos amigos do departamento de Epidemiologia do INCA que sempre estão dispostos a esclarecer dúvidas e encaminhar nossa pesquisa da maneira correta. Em especial à Dra. Liz Almeida, ao Dr. Luis Martins e à Dra. Neilane Bertoni. Durante esses anos tem ocorrido um intercâmbio de conhecimento entre a pesquisa básica, epidemiologia e assistência, imensamente enriquecedor.

Agradecimento especial à Dra. Neilane Bertoni por ser responsável pelo desenvolvimento e aplicação dos modelos estatísticos que foram aplicados na presente tese.

Ao Dr. Hector Seuánez por todo apoio dentro do Departamento de Genética, conversas e conselhos.

À Dra. Maria do Socorro Pombo por permitir a utilização do pirosequenciamento em seu laboratório e à tecnóloga Gisele Moledo por todo o assessoramento.

Às pacientes que permitiram o uso de suas amostras possibilitando que esse e muitos outros trabalhos fossem realizados.

À Dra. Alcina pela amizade, conselhos e colaborações. O maior ensinamento: “A humildade é a base para a conquista e reconhecimento”.

Aos médicos, enfermeiros e profissionais da área de saúde que atenderam as pacientes, coletaram informações e amostras de estudo, por estarem sempre presentes esclarecendo todas as dúvidas no que diz respeito à esfera clínica. Em especial gostaria de agradecer à cirurgiã Dra. Cláudia Bessa, profissional que compreende a importância maior de toda a pesquisa básica: “gerar conhecimento para salvar vidas”.

Ao Dr. Robert Burk por ter me recebido em seu laboratório no Albert Einstein College of Medicine em Nova Iorque, durante o doutorado sanduíche, disponibilizando recursos e transmitindo conhecimentos. Em seu laboratório foram produzidos grande parte dos resultados apresentados na presente tese. Agradecer também à sua pós-doc Ana Gradíssimo e os demais membros no laboratório que trocaram conhecimento e experiências durante o intercâmbio permitindo o trabalho fosse desenvolvido.

Aos órgãos de pesquisa: CNPq, FAPERJ, INCT e Ministério da Saúde.

SUMÁRIO

RESUMO	v
ABSTRACT	vi
AGRADECIMENTOS	viii
LISTA DE FIGURAS	xii
LISTA DE TABELAS	xiv
LISTA DE SIGLAS	xv
1. TÍTULO	1
2. INTRODUÇÃO	1
2.1. O CÂNCER CERVICAL	1
2.2. O PAPILOMAVÍRUS HUMANO (HPV)	2
2.2.1. REGULAÇÃO DA EXPRESSÃO GÊNICA PELA LCR DE HPV	5
2.2.2. METILAÇÃO NO DNA E CÂNCER.....	10
2.2.3. METILAÇÃO NO GENOMA DE HPV	12
3. OBJETIVOS	16
3.1. OBJETIVO GERAL	16
3.2. OBJETIVOS ESPECÍFICOS	16
4. METODOLOGIA	17
4.1. CASUÍSTICA	17
4.1.1. CRITÉRIOS DE INCLUSÃO DAS AMOSTRAS	17
4.2. TRATAMENTO COM BISSULFITO SÓDIO.....	20
4.3. DETECÇÃO DE METILAÇÃO DE DNA POR PIROSEQUENCIAMENTO	20
4.3.1. REGIÃO DE INTERESSE E REAÇÃO DE AMPLIFICAÇÃO.....	20
4.3.2. PIROSEQUENCIAMENTO	22
4.4. ENSAIOS DE METILAÇÃO DE DNA POR NGS	24
4.4.1. REGIÃO DE INTERESSE E REAÇÃO DE AMPLIFICAÇÃO.....	24
4.4.2. PREPARO DA BIBLIOTECA DE DNA.....	28
4.4.3. BIOINFORMÁTICA.....	30
4.4.4. CONTROLE DE QUALIDADE DOS NÍVEIS DE METILAÇÃO	31
4.5. ANÁLISE DE INTEGRIDADE DAS ORFs E1/E2 DE HPV	31
4.6. CARGA VIRAL DE HPV16.....	34
4.7. ANÁLISE ESTATÍSTICA	34
5. RESULTADOS	36
5.1. ANÁLISES DE METILAÇÃO POR PIROSEQUENCIAMENTO	36
5.1.1. CARACTERIZAÇÃO DA POPULAÇÃO DE ESTUDO.....	36

5.1.2. INTEGRIDADE DE E1/E2 DE HPV.....	36
5.1.3. DESCRIÇÃO DOS NÍVEIS DE METILAÇÃO EM 3'LCR DE HPV16 E HPV18.....	37
5.1.4. ASSOCIAÇÃO DOS NÍVEIS DE METILAÇÃO NA LACR DE HPV16 E HPV18 COM O TIPO DE TUMOR, O ESTÁGIO TUMORAL, A VARIANTE DE HPV E A INTEGRIDADE DE E1/E2.....	39
5.1.5. CORRELAÇÃO DOS NÍVEIS DE METILAÇÃO NA LCR DE HPV16 E HPV18 COM A IDADE DA PACIENTE.....	41
5.1.6. ANÁLISE BIVARIADA E MULTIVARIADA UTILIZANDO MODELO DE QUASI-VEROSSIMILHANÇA.....	42
5.2. ANÁLISES DE METILAÇÃO POR NGS.....	42
5.2.1. CARACTERIZAÇÃO DA POPULAÇÃO DE ESTUDO.....	44
5.2.2. INTEGRIDADE DE E1/E2 DE HPV.....	44
5.2.3. REPRODUTIBILIDADE DOS DADOS DE NGS.....	44
5.2.4. DESCRIÇÃO DOS NÍVEIS DE METILAÇÃO EM L2, L1 E LCR DE HPV16 E HPV18.....	45
5.2.5. DESCRIÇÃO DOS HAPLÓTIPOS DE METILAÇÃO EM L2, L1 E 3'LCR DE HPV16 E HPV18.....	48
5.2.6. ASSOCIAÇÃO DOS NÍVEIS DE METILAÇÃO EM L2, L1 E LCR DE HPV16 E HPV18 COM O TIPO DE TUMOR, A LINHAGEM DE HPV E O ESTÁGIO TUMORAL.....	51
5.2.7. ASSOCIAÇÃO DOS NÍVEIS DE METILAÇÃO EM L2, L1 E LCR DE HPV16 E HPV18 COM A INTEGRIDADE DE E1/E2.....	54
5.2.8. ANÁLISE BIVARIADA E MULTIVARIADA UTILIZANDO O MODELO DE QUASI-VEROSSIMILHANÇA.....	56
5.2.9. CORRELAÇÃO ENTRE OS NÍVEIS DE METILAÇÃO POR PIROSEQUENCIAMENTO E NGS.....	59
6.DISCUSSÃO.....	60
7.CONCLUSÕES.....	70
8.REFERÊNCIAS.....	71
9. ANEXOS.....	84
9.1. Anexo 1. Registro do comitê de ética em pesquisa (CEP).....	84
9.2. Anexo 2. Prevalência de HPV na população do estudo principal.....	85
9.3. Anexo 3. Artigo publicado sobre o projeto de doutorado.....	86
9.4. Anexo 4. Artigos publicados em colaboração.....	92

9.5. Anexo 5. Informações totais das amostras HPV16 positivas analisadas por pirosequenciamento	94
9.6. Anexo 6. Informações totais das amostras HPV18 positivas analisadas por pirosequenciamento	97
9.7. Anexo 7. Informações totais das amostras HPV16 positivas analisadas por NGS para os dados de metilação do ensaio HPV16_L2	99
9.8. Anexo 8. Informações totais das amostras HPV16 positivas analisadas por NGS para os dados de metilação do ensaio HPV16_L1	102
9.9. Anexo 9. Informações totais das amostras HPV16 positivas analisadas por NGS para os dados de metilação do ensaio HPV16_LCR	105
9.10. Anexo 10. Informações totais das amostras HPV18 positivas analisadas por NGS para os dados de metilação do ensaio HPV18_L2	108
9.11. Anexo 11. Informações totais das amostras HPV18 positivas analisadas por NGS para os dados de metilação do ensaio HPV18_L1.....	110
9.12. Anexo 12. Informações totais das amostras HPV18 positivas analisadas por NGS para os dados de metilação do ensaio HPV18_LCR.....	112

LISTA DE FIGURAS

Figura 2.1. Genoma de HPV16.	4
Figura 2.2. Caracterização da LCR com ênfase em 3'LCR.....	6
Figura 2.3. Distribuição dos sítios CpG nos genomas de HPV16 e HPV18.	9
Figura 4.1. Fluxograma com o número de amostras selecionadas para as análises de metilação por pirosequenciamento e NGS	19
Figura 4.2. Etapa de checagem da conversão total por bissulfito de sódio.....	21
Figura 4.3. Princípio da análise de metilação e pirosequenciamento.....	24
Figura 4.4. Identificação dos sítios CpG analisados nos genomas de HPV16 e HPV18.....	29
Figura 4.5. Princípio das análises de metilação por NGS.....	29
Figura 4.6. Representação da análise de integridade de E1 e E2 de HPV16	32
Figura 5.1. Distribuição dos níveis de metilação analisados por pirosequenciamento em 5 sítios CpG presentes na LCR de HPV16 e HPV18 em amostras tumorais	39
Figura 5.2. Associação dos níveis de metilação analisados por pirosequenciamento em 10 sítios CpG presentes na LCR de HPV16 e HPV18 por variáveis de estudo.	40
Figura 5.3. Distribuição e correlação dos níveis de metilação, analisados por pirosequenciamento, e a idade da paciente no momento do diagnóstico.	41
Figura 5.4. Controle de reprodutibilidade: correlação entre os níveis de metilação entre as duplicatas de amostras para os diferentes ensaios.....	45
Figura 5.5. Distribuição dos níveis de metilação analisados por NGS em sítios CpG presentes em L2, L1 e LCR de HPV16 e HPV18.	47
Figura 5.6. Frequência dos principais haplótipos de metilação presentes em ensaios de HPV16 por NGS.	49

Figura 5.7. Frequência dos principais haplótipos de metilação presentes em ensaios de HPV18 por NGS.	50
Figura 5.8. Distribuição dos níveis de metilação analisados por NGS em sítios CpG presentes em L2, L1 e LCR de HPV16 por tipo histológico, variante de HPV e estágio tumoral.....	52
Figura 5.9. Distribuição dos níveis de metilação analisados por NGS em sítios CpG presentes em L2, L1 e LCR de HPV18 por tipo histológico, variante de HPV e estágio tumoral.....	53
Figura 5.10. Associação entre os níveis de metilação e a integridade das ORFs E1 e E2 de HPV16 e HPV18, analisados por NGS.	55

LISTA DE TABELAS

Tabela 4.1. Informações dos iniciadores utilizados nos ensaios de análise de metilação por pirosequenciamento e NGS.	26
Tabela 4.2. Informações sobre os iniciadores utilizados nos ensaios de análise da integridade das ORFs E1 e E2 de HPV16 e HPV18.	33
Tabela 5.1. Características das amostras selecionadas da população de estudo agrupadas por técnica de análise de metilação e tipo de HPV.	38
Tabela 5.2. Análises bivariadas e multivariadas para associação dos níveis de metilação por variáveis de estudo utilizando pirosequenciamento.	43
Tabela 5.3. Análises bivariadas para os níveis de metilação por sítio CpG utilizando modelo de quasi-verossimilhança	57
Tabela 5.4. Análises multivariadas para os níveis de metilação por sítio CpG utilizando modelo de quasi-verossimilhança	58
Tabela 5.5. Correlação entre os dois métodos de análise de metilação, pirosequenciamento e NGS, para os ensaios HPV16_LCR e HPV18_LCR	59

LISTA DE SIGLAS

ADP	Adenosina difosfato
Alu	<i>Arthrobacter luteus</i> (Família Alu de elementos de DNA repetitivo)
ADC	Adenocarcinoma
ATP	Adenosina trifosfato
AUC	Área sobre a curva (<i>area under the curve</i>)
CaSki	Linhagem celular imortalizada infectada com HPV16
CCE	Carcinoma de células escamosas
CEP	Comitê de ética em pesquisa
Chi2	Teste de Qui-Quadrado
COBRA	Análise de restrição combinada com bissulfite (<i>Combined Bisulfite Restriction Analysis</i>)
CpG	Citosina seguido de uma guanina no sentido 5'-3'
CT	Ciclo de referência, limiar de detecção (<i>cycle threshold</i>)
DNMT	DNA metil transferase
dNTP	Desoxirribonucleotídeo trifosfato (<i>Deoxyribonucleoside Triphosphates</i>)
DST	Doença sexualmente transmissível
E1/E2	genes E1 e/ou E2
E2BS	Sítio de ligação para a proteína E2 de HPV
FIGO	Federação Internacional de Ginecologia e Obstetrícia (<i>The International Federation of Gynecology and Obstetrics</i>)
HAT	Histona acetiltransferase
HDAC	Histona deacetilase
HeLa	Linhagem celular imortalizada infectada com HPV18
HPV-AR	Papilomavírus humano de alto risco
HSIL	Lesão intraepitelial escamosa de alto grau (<i>high-grade squamous intraepithelial lesions</i>)
INCA	Instituto Nacional do Câncer
INCA/MS	Instituto Nacional do Câncer/Ministério da Saúde
Kb	Pares de quilobases
LCR	Região controladora longa (<i>long control region</i>)
LINES	Elementos nucleares intercalares longos (<i>Long Interspersed Elements</i>)
LSIL	Lesão intraepitelial escamosa de baixo grau (<i>low-grade squamous intraepithelial lesions</i>)
M	Molar (concentração)
MeCP	Complexo de proteínas ligadoras de sítio CpG metilado
MgCl²	Cloreto de magnésio
mM	Milimolar
MS-PCR	PCR específica para metilação (<i>methylation-specific PCR</i>)
ng	Nanograma
NGS	Sequenciamento de nova geração (<i>next generation sequencing</i>)
nt	Nucleotídeo
ORF	Quadro aberto de leitura (<i>Open reading frame</i>)
P105	Promotor precoce para HPV18
p53	Proteína de 53kD
P97	Promotor precoce para HPV16
PAE	Poliadenilação de transcritos virais precoces
PAL	Poliadenilação de transcritos virais tardios

PAVE	Banco de dados eletrônico genômico do papiloma vírus (<i>Papilloma virus genome database</i>)
Pb	Pares de bases
PCR	Reação em Cadeia da Polimerase (<i>Polymerase Chain Reaction</i>)
pmol	Picomoles
pmol/μL	Picomoles por microlitro
Ppi	Pirofosfato inorgânico
pRb	Proteína do retinoblastoma
qPCR	PCR em tempo real
rpm	Rotação por minuto
SINES	Elementos Intercalares Curtos (<i>Short Interspersed Elements</i>)
Taq	<i>Thermus aquaticus</i>
URR	Região controladora <i>upstream</i> (<i>upstream regulatory region</i>)
μL	Microlitro
3'LCR	Segmento 3' da região controladora longa (<i>long control region</i>)

1. TÍTULO

DIVERSIDADE DE METILAÇÃO NO DNA DE HPV16 E 18 EM CARCINOMAS CERVICAIS INVASIVOS

2. INTRODUÇÃO

2.1. O CÂNCER CERVICAL

Apesar dos resultados promissores alcançados nas últimas décadas com o rastreamento de mulheres assintomáticas pelo teste de Papanicolaou e mais recentemente com o advento das vacinas contra o agente etiológico do câncer cervical, nem todos os casos por esse tipo de câncer podem ser prevenidos e, ainda ocorrem por ano no mundo, segundo as últimas estimativas mundiais para o ano de 2012, cerca de 530.000 novos casos de câncer cervical com aproximadamente 275.000 mortes (Jemal *et al.*, 2011). No Brasil, dados do Instituto Nacional de Câncer (INCA) e Ministério da Saúde estimam, para o ano de 2016, 16.340 casos novos de mulheres acometidas por essa neoplasia (Inca/Ms, 2015). Excluindo os tumores de pele não melanoma, estima-se o câncer de colo uterino como tumor mais incidente na região Norte e o segundo mais frequente nas regiões Nordeste e Centro-Oeste. Na região Sudeste as estimativas apontam a terceira posição, enquanto a região Sul a quarta (Inca/Ms, 2015).

Essa variação poderia ser explicada pela grande extensão territorial do país e pelas desigualdades sócio-econômicas entre as áreas, visto que a incidência do câncer de colo do útero está relacionada a vários fatores comportamentais e epidemiológicos. Dentre os principais, destacam-se a baixa escolaridade e renda, o início precoce da atividade sexual, multiplicidade de parceiros sexuais, uso prolongado de contraceptivos orais, multiparidade, história de doenças sexualmente transmissíveis (DSTs) e a ineficiência na implementação das estratégias de rastreamento. Além desses, outros fatores de risco como imunossupressão, fatores genéticos e tabagismo estão supostamente envolvidos no desenvolvimento do câncer cervical (Daling *et al.*, 1996; Bayo *et al.*, 2002; Pinto *et al.*, 2002; Inca/Ms, 2015).

O câncer de colo uterino é uma doença que ocorre a partir de anormalidades intra epiteliais as quais evoluem lentamente podendo adquirir um caráter maligno, levando a um processo invasivo. As lesões intraepiteliais precursoras do câncer cervical podem ser de baixo grau (LSIL, do inglês *Low-Grade Squamous*

Intraepithelial Lesion) que são lesões intraepiteliais correspondendo a uma displasia leve, e de alto grau (HSIL, do inglês *High-Grade Squamous Intraepithelial Lesion*) que se associa à displasia moderada, displasia severa e ao carcinoma *in situ* (Apgar *et al.*, 2003).

A infecção persistente pelo papilomavírus humano (HPV) compreende o fator de risco necessário para o aparecimento de lesões intraepiteliais de alto grau e do câncer do colo do útero (Durst *et al.*, 1983; Walboomers *et al.*, 1999; Vaccarella *et al.*, 2008). Durante as últimas décadas, com o esclarecimento da importância da infecção pelo HPV para a carcinogênese cervical, ampliou-se consideravelmente o conhecimento sobre a biologia molecular do papilomavírus, seu ciclo de vida e o seu papel no desenvolvimento não só do câncer de colo do útero mas também em tumores envolvendo outros sítios anatômicos como a região da orofaringe, anal e peniana (Mckaig *et al.*, 1998; Doorbar, 2006; Backes *et al.*, 2009; Forman *et al.*, 2012).

Entretanto, fatores envolvendo a persistência e clareamento da infecção pelo HPV ainda não foram elucidados (Trottier e Franco, 2006; Goodman *et al.*, 2008). Sabe-se que apenas pequena parcela (<1%) de mulheres inicialmente infectadas por HPV de alto risco oncogênico desenvolverão câncer cervical, havendo necessidade em identificar marcadores adicionais capazes de prever as infecções de HPV16 que progredirão para as lesões de alto grau e câncer, e também para os casos que precisam de uma intervenção terapêutica. O Papanicolaou é o teste tradicional de rastreio baseado em citologia, e apresenta baixa taxa relativa de sensibilidade (Schiffman *et al.*, 2005). Já os dados sobre o teste para presença de HPV de alto risco indicam alta sensibilidade, porém baixa especificidade (Schiffman *et al.*, 2005). Assim, novos biomarcadores são necessários para diferenciar mulheres infectadas com HPV com risco significativo de desenvolvimento de lesões de alto risco e carcinoma *in situ*.

2.2. O PAPILOMAVÍRUS HUMANO (HPV)

Os HPVs são vírus pequenos (~ 55 nm) não envelopados, contendo um genoma circular em fita dupla com aproximadamente 8 pares de quilobases (Kb) envolvidos por um capsídeo icosaédrico. Mais de 40 genótipos de HPV já foram identificados infectando a mucosa genital e 13 destes HPV (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59 and 68) (Iarc, 2012) já estão bem estabelecidos como HPV de

alto risco (HPV-AR) devido a sua alta prevalência em cânceres cervicais (>90%) e evidências moleculares (Munoz *et al.*, 2003).

O genoma de HPV é dividido em três regiões: (1) região de ~ 4 Kb com genes que codificam proteínas necessárias para a replicação do DNA viral, regulação da expressão gênica de HPV e envolvidos no processo de carcinogênese (genes E1, E2, E4, E5, E6 e E7), (2) região codificante das proteínas estruturais do capsídeo (genes L1 e L2) com ~ 3kb do genoma e, (3) uma região não codificante denominada LCR (do inglês *long control region*) ou URR (do inglês *upstream regulatory region*) de ~ 1Kb que contém sequências conservadas necessárias para expressão gênica e início da replicação do DNA viral (Schwarz *et al.*, 1985; Munger *et al.*, 2004; Doorbar, 2006; Munoz *et al.*, 2006) (Figura 2.1).

A fim de refletir o padrão de expressão gênica *in vivo*, em uma infecção produtiva, o genoma de HPV pode ser agrupado em genes exclusivamente de expressão precoce (E6 e E7), genes precoces e tardios (E1, E2, E4 e E5) e genes exclusivamente tardios (L1 e L2) expressos nas camadas superficiais do epitélio (Johansson e Schwartz, 2013). Durante a progressão do câncer cervical é observada uma mudança no padrão de expressão dos genes virais com ausência de expressão dos genes tardios, L1 e L2, e superexpressão dos oncogenes E6 e E7 (Doorbar, 2006).

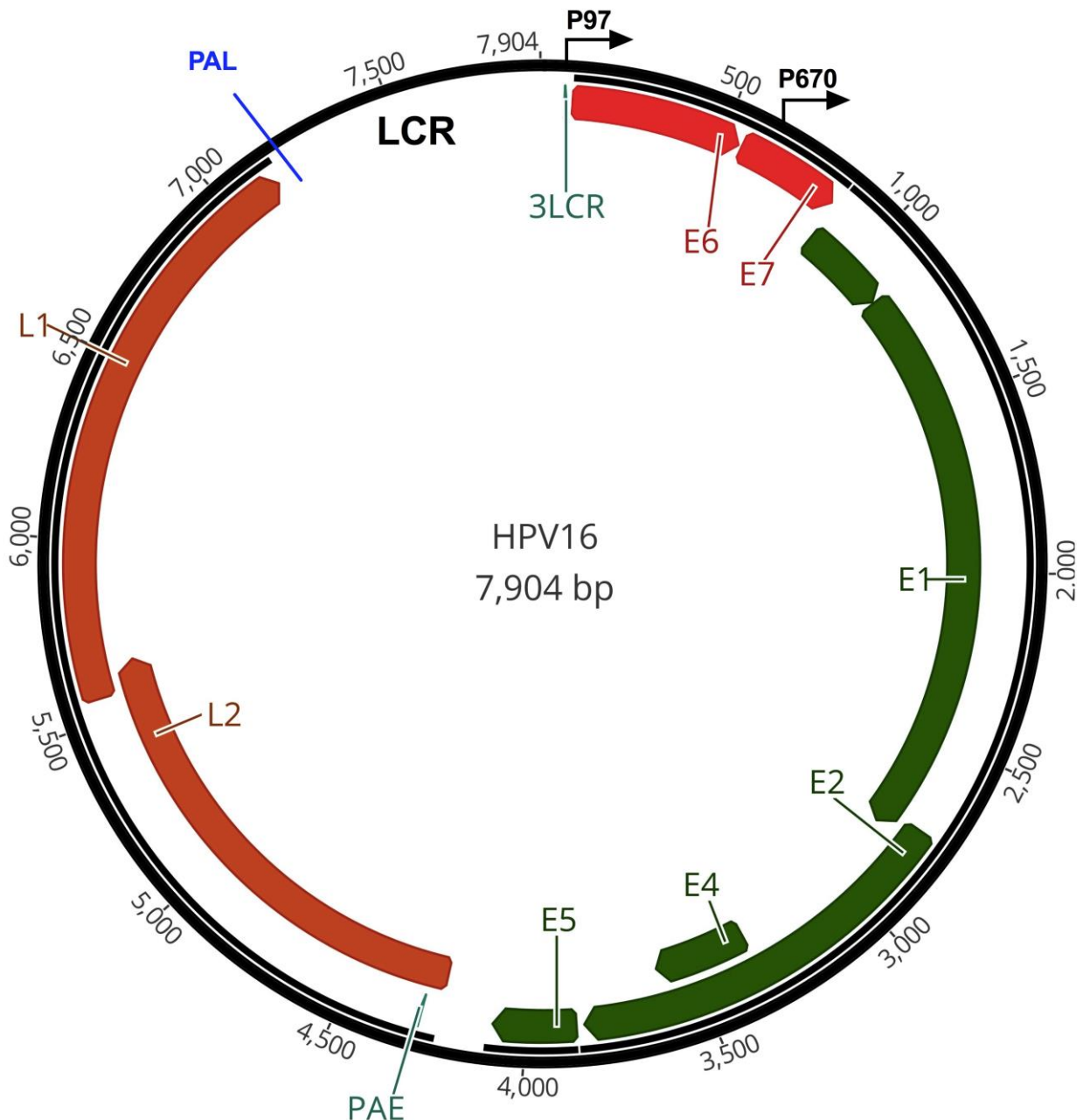


Figura 2.1. Genoma de HPV16. A figura identifica o posicionamento no genoma de HPV16 de cada ORF (fase aberta de leitura, do inglês *open reading frame*) e da região da LCR. Em vermelho, identificando os principais oncogenes, E6 e E7. Em verde, os genes que regulam a atividade gênica e necessários para replicação (E1, E2, E4 e E5) e, em laranja, os genes estruturais que formam o capsídeo viral (L1 e L2). Na porção superior se encontra a LCR, entre L1 e E6, destacando a região 3'LCR, promotora. P97 é o promotor precoce utilizado para transcrição dos genes *E6* e *E7* e P670 promotor dos genes tardios, L1 e L2. PAE e PAL são os sítios de poliadenilação dos genes precoces e tardios, respectivamente. Figura desenhada com base no genoma de referência de HPV16 (número de acesso K02718.1), utilizando o programa Geneious 10.0.9.

2.2.1. REGULAÇÃO DA EXPRESSÃO GÊNICA PELA LCR DE HPV

Como nos organismos eucariotos, acredita-se que a cromatina de HPV está complexada com histonas e estruturada em nucleossomos (Favre, 1975; Favre *et al.*, 1977). Conseqüentemente, estes nucleossomos podem sofrer a influência de enzimas celulares como as histonas acetil transferases (HATs) e histonas deacetilases (HDACs) permitindo ou bloqueando o acesso da maquinaria proteica de transcrição. Além disso, citosinas em dinucleotídeos CpGs (citosinas seguidas de guaninas no sentido 5'-3') presentes no genoma viral, são capazes de sofrer metilação através da ação das enzimas DNA metiltransferases (DNMT) celulares, influenciando a expressão gênica de HPV (Badal *et al.*, 2004; Kalantari *et al.*, 2004; Bernard, 2013).

A regulação da transcrição dos genes de HPV ocorre principalmente na região da LCR. A LCR de HPV16, que pode ser utilizada como modelo de LCR para todos os HPV genitais, é uma região de aproximadamente 850 pares de base (pb) dividida em três segmentos: 5'LCR, LCR central e 3'LCR (Figura 2.2 A) (O'Connor, 1995). O segmento 5'LCR possui ~ 300pb, localizado entre o códon terminal de L1 e o sítio de ligação para proteína viral E2 na posição nt 35 a 46 (E2BS#1), é a região de ancoragem à matriz nuclear, contendo sítios de término da transcrição e de poliadenilação para os transcritos tardios. O segmento central, também denominado, em inglês, *enhancer*, apresenta ~ 400pb e está flanqueado por dois sítios de ligação de E2 (E2BS#1 e E2BS#2) (Figura 2.2 A). Nesse segmento, diversos fatores de transcrição celulares, como por exemplo AP-1, NFI, YY1, Oct-1, TF-1, TEF-2, receptores para glicocorticóides e receptor para progesterona, se ligam, sendo responsáveis pelo tropismo dos HPV por epitélios específicos e pela transativação de promotores gênicos virais (Desaintes e Demeret, 1996; Bernard, 2013). O segmento 3'LCR, também chamado de promotor por incluir nessa região o promotor dos genes precoces, apresenta ~ 140pb, sendo flanqueado por ESBS#2 e o gene *E6*. Esse segmento contém uma origem de replicação na porção *upstream*, localizada no sítio de ligação para proteína viral E1, uma região *downstream* ao sítio de E1 para ligação do fator Sp1, seguido de dois sítios de ligação para E2 (E2BS#3 e E2BS#4) e uma região TATA box de ligação do fator TFIID (Figura 2.2 A, B e C) (Tan *et al.*, 1994).

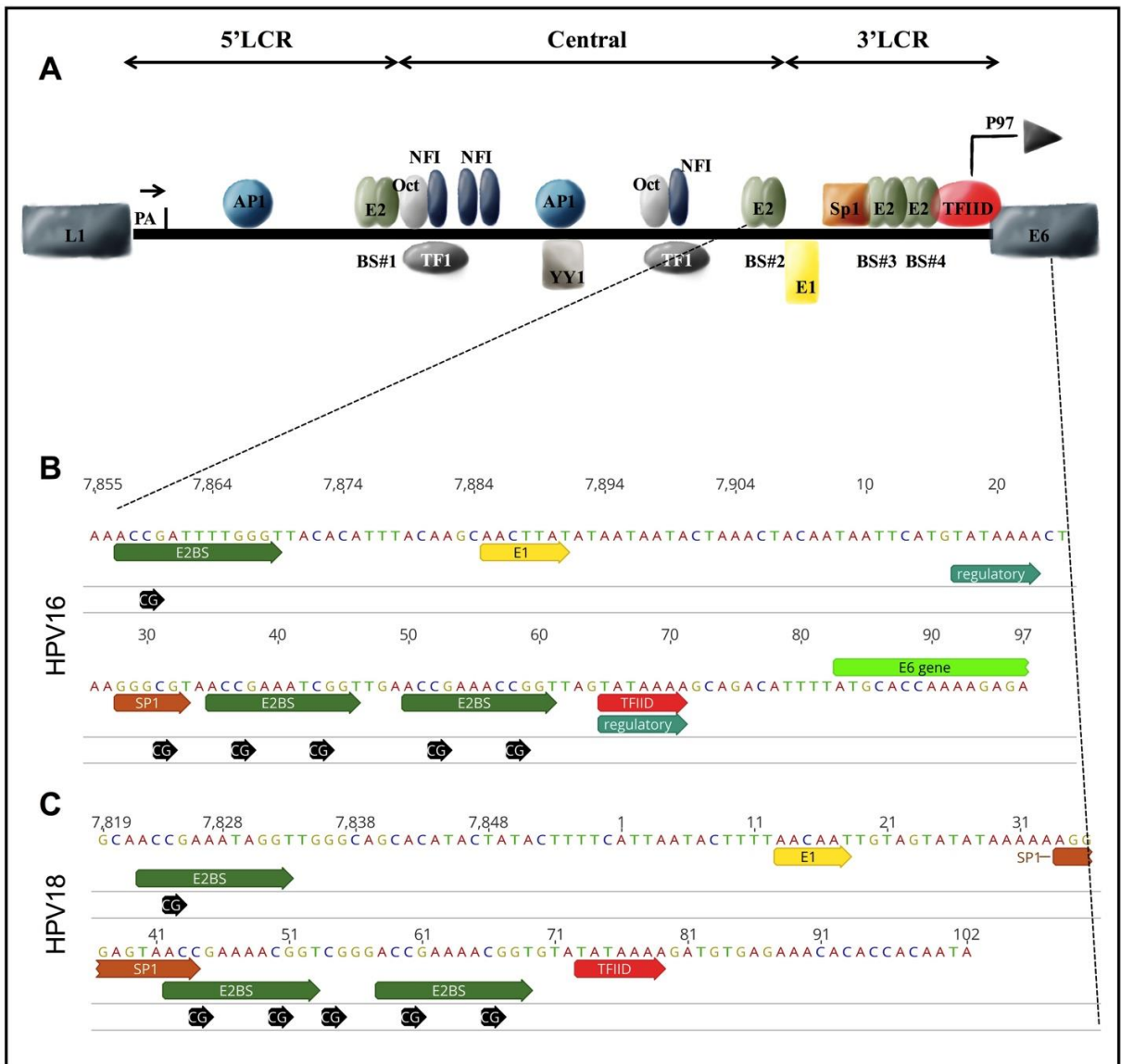


Figura 2.2. Caracterização da LCR com ênfase em 3'LCR. Em A, acima, observa-se uma ilustração da região da LCR de HPV16 e seus três segmentos, 5'LCR, central e 3'LCR, delimitados pelos genes E6 (à direita) e L1 (à esquerda). Alguns fatores de transcrição (Sp1, E2, E1, TFIID, Oct, NFI, AP1), o promotor dos genes precoces (P97) e a região de poliadenilação identificando o final da transcrição (PA) estão representados. Estão identificados também os quatro sítios de ligação para a proteína E2 de HPV16 (E2BS#1 - #4). Em B e C, estão representadas as sequências de nucleotídeos do segmento 3'LCR de HPV16 e HPV18, respectivamente, mostrando o posicionamento dos sítios CpG (em preto), e as sequências conservadas (*motifs*) de ligação das proteínas importantes para regulação da transcrição e início da atividade do promotor P97 e P105 (Sp1, E2 e TFIID). Motifs: AACNAT (E1), ACCG(n4)CGGT (E2), GGGCGT (SP1, porção principal, em comum com HPV16 e HPV18), TATAAAA (TFIID). Figura baseada nos trabalhos de O'CONNOR *ET AL.* (1995) e TAN *ET AL.* (1994) e o desenho realizado com os programas Geneious 10.0.9.

Juntamente com as proteínas celulares, as proteínas E2 de HPV regulam a atividade transcricional na LCR, desempenhando papel fundamental durante a carcinogênese induzida pelo HPV. Além disso, E2 interage diretamente com E1 e está envolvida na replicação do DNA viral dependente de E1 e manutenção do genoma (Abbate *et al.*, 2004).

E2 apresenta uma ação dose-dependente, e dependendo da localização do sítio de ligação, pode atuar tanto promovendo, diminuindo ou reprimindo a atividade transcricional da LCR (revisado por Ottinger *et al.*, 2009 e McBride, 2013) (Thierry e Yaniv, 1987; Romanczuk *et al.*, 1990; Tan *et al.*, 1994; Ottinger *et al.*, 2009; McBride, 2013). Quatro sítios conservados de ligação para E2 (E2BS#1, E2BS#2, E2BS#3 e E2BS#4) na LCR, com uma sequência palindrômica consenso ACCG(N)₄CGGT, são capazes de regular a região promotora dos genes E6 e E7, denominado P97 para HPV16 e P105 para HPV18 (Steger e Corbach, 1997). Nucleotídeos internos e sequências flanqueadoras são capazes de influenciar a afinidade e estabilidade de ligação da proteína E2 e a magnitude dos seus efeitos (Sanders e Maitland, 1994).

O sítio E2BS#1 de HPV16 e HPV18, localizado na porção 5'LCR, apresenta forte afinidade à proteína E2, os sítios E2BS#3 e E2BS#4 têm uma moderada afinidade e o sítio E2BS#2 fraca afinidade. Sendo assim, baixas concentrações da proteína E2 são capazes de ativar o sítio de ligação E2BS#1, promovendo a transativação de P97 e P105. Já em elevadas concentrações, E2 se liga aos sítios E2BS#3 e #4, deslocando os fatores de transcrição celular Sp1 e TFIID, consequentemente diminuindo ou reprimindo a transcrição a partir desses promotores (Thierry e Yaniv, 1987; Thierry e Howley, 1991; Tan *et al.*, 1994). Altas concentrações de E2 também aumentam a atividade transcricional ao ocupar o sítio E2BS#2 (Sanders e Maitland, 1994; Steger e Corbach, 1997; McBride, 2013).

Assim, a ausência ou baixos níveis da proteína E2 estariam associados à superexpressão dos oncogenes E6 e E7, e à progressão do câncer. Isso porque, como bem descrito na literatura, altos níveis das oncoproteínas E6 e E7 são capazes de alterar mecanismos do ciclo celular. Dentre outras funções, a proteína E6 é capaz de induzir a degradação da fosfoproteína nuclear P53 através da via de ubiquitinação (Tommasino *et al.*, 2003; Munger *et al.*, 2004). Já a proteína E7 interage e inativa a fosfoproteína nuclear Retinoblastoma (pRb1), que desempenha um papel chave na regulação do ciclo celular, controlando negativamente a

passagem da fase G1 para S (Milde-Langosch e Riethdorf, 2003; Munger *et al.*, 2004).

Em lesões pré-malignas e no câncer cervical, a ausência ou diminuição da concentração da proteína E2 está associada ao processo de ruptura do DNA circular de HPV e integração ao DNA da célula hospedeira, geralmente em locais mais susceptíveis à instabilidade genômica (Duensing e Munger, 2002; Doorbar, 2006). A integração ocorre mais frequentemente com a ruptura do genoma viral nos genes E1 e E2, o que permite a expressão das oncoproteínas E6 e E7 mas não de E2 (Berumen *et al.*, 1994; Xu *et al.*, 2013; Liu *et al.*, 2016).

Entretanto, em aproximadamente 40% dos tumores cervicais tem sido observada a presença de formas epissomais ou concomitantes (epissomal e integrado) do DNA de HPV16 (Cullen *et al.*, 1991; Cricca *et al.*, 2009; Liu *et al.*, 2016), sugerindo a possibilidade da presença da proteína E2 funcional. Além disso, E2 intacto é também frequente em tumores com integração com múltiplas cópias virais em *tandem*. Diferentemente, em tumores cervicais infectados por HPV18, observa-se quase a totalidade das amostras com o genoma viral integrado (Cullen *et al.*, 1991; Park *et al.*, 1997).

Além do fator integração, evidências sugerem diferenças biológicas entre HPV16 e HPV18 durante a progressão do câncer cervical. Baseado em revisões e meta análises disponíveis na literatura (Clifford *et al.*, 2003; Clifford *et al.*, 2005; Clifford *et al.*, 2006), HPV16 é o tipo carcinogênico mais comum, detectado em aproximadamente 25% das lesões intra-epiteliais de baixo grau (LSIL) e em 50% das lesões de alto grau (HSIL) e câncer cervical em todo o mundo. Apesar de HPV18 ser o segundo tipo de HPV mais prevalente em câncer cervical, é observada baixa prevalência deste tipo de HPV em LSIL e HSIL, sugerindo evolução relativamente rápida da doença. Além disso, HPV18 está associado a uma maior proporção de adenocarcinomas (Clifford *et al.*, 2003; Clifford *et al.*, 2005; Clifford *et al.*, 2006).

As diferenças entre os dois tipos de HPV também são notadas na distribuição dos sítios CpG no genoma. Baseado nos genomas de referência de ambos, HPV16 possui 111 sítios CpG e HPV18 168 sítios CpG dispersos no genoma. Todos os genes de HPV18 apresentam maior número de sítios CpG quando comparado aos mesmos genes de HPV16, exceto em E5, 3 vs 5 sítios CpG, respectivamente (Figura 2.3).

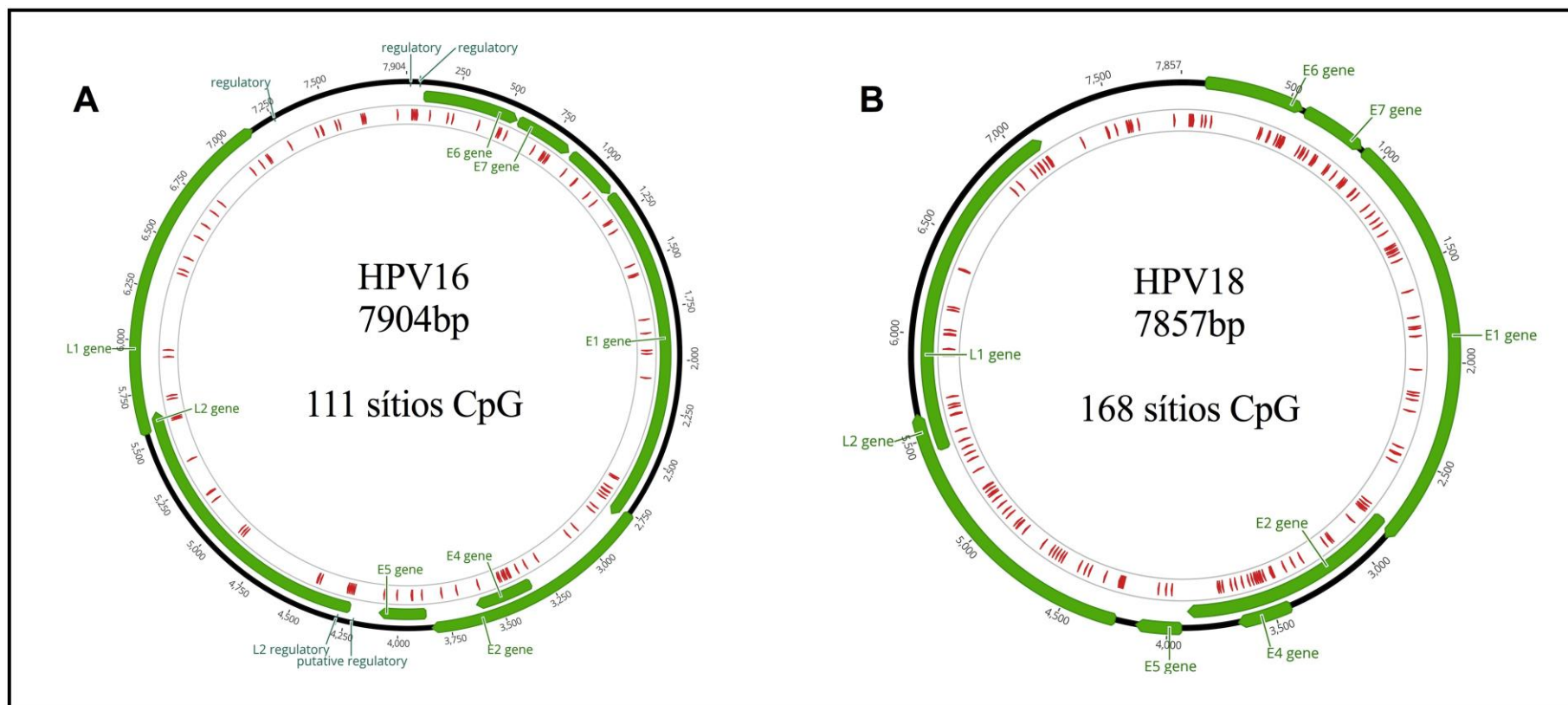


Figura 2.3. Distribuição dos sítios CpG nos genomas de HPV16 e HPV18. A distribuição e posicionamento dos genes e sítios CpG nos genomas de HPV16 (A) e HPV18 (B) estão representados em verde e vermelho, respectivamente. Desenho e contagem dos sítios através do programa Geneious 10.0.9. Número de acesso dos genomas de referência: K02718.1 (HPV16) e X05015.1 (HPV18).

Acredita-se que a metilação no DNA viral em E2BS#3 e E2BS#4 (Figura 2.2 A e B), nos casos com DNA epissomal ou em *tandem*, poderia contribuir regulando a transcrição dos oncogenes através do bloqueio da ligação da proteína E2, levando a um aumento da expressão de E6 e E7. Em 1996, foi mostrado por Thain e colaboradores que a metilação em determinados sítios CpG inibem de forma direta a ligação da proteína E2 de HPV16 (Thain *et al.*, 1996). Em seguida, Kim e colaboradores mostraram que a atividade transcricional de E2 pode ser também afetada pela metilação do DNA nas sequências de seus sítios de ligação (Kim *et al.*, 2003). Chaiwongkot *et al.*, reportaram esses achados em carcinomas de células escamosas com HPV16, sugerindo que altos níveis de metilação em E2BS#3 e #4 contribuem com a superexpressão dos transcritos de E6 e E7 por bloquearem a ligação de E2 (Chaiwongkot *et al.*, 2013). Entretanto, esses estudos não avaliaram a presença da proteína E2.

2.2.2. METILAÇÃO NO DNA E CÂNCER

Metilação de citosinas em dinucleotídeos CpG é uma das principais modificações epigenéticas e está frequentemente associada à repressão transcricional (Bird, 2002). É a única modificação covalente conhecida no DNA de eucariotos e desempenha um importante papel durante o desenvolvimento e diferenciação celular. A metilação do DNA consiste na adição de um radical metil em citosinas seguidas de guanina no sentido 5'-3' (sítio CpG) catalisada por enzimas chamadas DNA metiltransferases (DNMTs), incluindo DNMT1, DNMT3a e DNMT3b. As DNMT1, são enzimas denominadas de manutenção devido a sua preferência, em mamíferos, por substratos parcialmente metilados (hemimetilados), enquanto as DNMT3a e DNMT3b estão envolvidas na metilação *de novo* (Baylin, 2005).

Foram descritos diferentes mecanismos de como a metilação pode impedir a ligação do fator de transcrição e levar ao silenciamento gênico: (1) através do bloqueio físico e direto devido à presença do grupamento metil, (2) de forma indireta, através da competição dos fatores de transcrição com os complexos de proteínas ligadoras de CpG metilados (MeCP1 e 2, do inglês *methyl-CpG binding protein complexes*), e (3) através do recrutamento de histona deacetilases (HDACs) pelos complexos MeCP causando deacetilação de histonas adjacentes e, conseqüente, condensação da cromatina e silenciamento gênico (dados revisados por Leonhardt e Cardoso, 2000).

A maior parte dos sítios CpG estão localizados em regiões denominadas ilhas CpG. Essas são regiões pequenas, cerca 500 pares de base, que apresentam mais de 55% de sequências CpG e estão frequentemente associadas a regiões promotoras gênicas, locais de início da transcrição (Bird, 2002; Antequera, 2003; Herman e Baylin, 2003).

Em tecidos normais não embrionários de humanos, a metilação do DNA é encontrada em aproximadamente 70% dos sítios CpG, sendo caracterizada como um padrão global de hipermetilação (Bird, 1986; Gronbaek *et al.*, 2007). Entretanto, essa metilação ocorre em sítios CpG localizados em regiões não promotoras de genes, regiões pericentroméricas, regiões heterocromáticas e intergênicas, promovendo maior estabilidade genômica (Bird, 1986; Chandler e Jones, 1988; Rollins *et al.*, 2006; Herrera *et al.*, 2008).

Fatores como a idade e componentes nutricionais da dieta, além de outros fatores ambientais e psicológicos, podem alterar as taxas de metilação de DNA a nível sistemático ou em alvos específicos em um indivíduo, podendo afetar a estabilidade genômica e transcrição de supressores tumorais e oncogenes (Christensen *et al.*, 2009; Lim e Song, 2012; Kim *et al.*, 2016). Em dados revisados por Liu e colaboradores (2003), o envelhecimento está associado à diminuição dos níveis globais de metilação em humanos, mas também em outros organismos como camundongos, ratos e salmões, supostamente devido à diminuição na expressão da enzima DNMT1. Já uma dieta deficiente em nutrientes doadores do grupamento metil, como folato e vitamina B12, podem afetar a via bioquímica do processo de metilação (Liu *et al.*, 2003). Piyathilake *et al.* (2014) observaram que elevados níveis desses nutrientes no plasma podem apresentar papel importante na manutenção dos níveis de metilação em sítios CpG presentes na LCR de HPV16. Além disso, mulheres com altas concentrações de folato e alta metilação de HPV16 ou aquelas com alta concentração de vitamina B12 e alta metilação de HPV16, eram menos prováveis de serem diagnosticadas com NIC2+ (Piyathilake *et al.*, 2014).

No câncer existe uma mudança no padrão de metilação passando para um estado de hipometilação global, com perdas de metilação nos corpos dos genes e, especialmente, em regiões com DNA repetitivo (sequências *Alu*, *LINES* e *SINES*), e uma hipermetilação focal em regiões específicas, como em ilhas CpG (Gronbaek *et al.*, 2007; Christensen *et al.*, 2009). A perda global nos níveis de metilação do DNA está associada à instabilidade genômica, uma vez que facilita quebras e

translocações cromossômicas (Hoffmann e Schulz, 2005). Já o aumento de metilação em ilhas CpG, frequentemente em promotores gênicos, está associado à diminuição da atividade de genes supressores tumorais (Gronbaek *et al.*, 2007).

2.2.3. METILAÇÃO NO GENOMA DE HPV

No que se refere à metilação no DNA de HPV, sabe-se, entretanto, que o genoma de HPV está submetido à ação das DNMT celulares, especialmente as DNMT1 provavelmente como mecanismo de resposta inata do hospedeiro (Badal *et al.*, 2004). Assim, a metilação no genoma viral pode ser, em parte, um mecanismo pelo qual o hospedeiro tenta suprimir a expressão dos genes virais e dessa forma sua patogenicidade.

De outro modo, foi observado que as oncoproteínas E6 e E7 de HPV são capazes de modular a maquinaria epigenética celular tanto alterando a expressão e atividade das DNMTs quanto de enzimas modificadoras de histonas (Au Yeung *et al.*, 2010; Mclaughlin-Drubin *et al.*, 2011; Hsu *et al.*, 2012). Au Yeung e colaboradores demonstraram, por exemplo, que os níveis de expressão da DNMT1, em linhagens de câncer cervical infectadas com HPV16 (SiHa e CaSki), podem ser regulados pela proteína E6 de HPV16 através da proteína p53. Nesse estudo, a presença de E6 esteve associada a baixos níveis de p53 e aumento da atividade transcricional da DNMT1 (Au Yeung *et al.*, 2010). Isso pode resultar em alterações na regulação do genoma viral, mas também no genoma do hospedeiro, desempenhando um papel essencial tanto no ciclo de vida de HPV quanto no desenvolvimento das neoplasias cervicais (Johannsen e Lambert, 2013).

De forma geral, os estudos sugerem a metilação no DNA de HPV como um processo dinâmico, em que se observam mudanças no padrão de metilação do vírus, tanto durante o ciclo de vida de HPV nas diferentes camadas do epitélio quanto durante a progressão neoplásica (Kim *et al.*, 2003; Chaiwongkot *et al.*, 2013).

Dezenas de estudos têm investigado o padrão de metilação no genoma de HPV no contexto envolvendo ciclo de vida viral e associado à progressão neoplásica (revisado por Johannsen e Lambert, 2013). Na literatura a maioria dos trabalhos descrevem o padrão de metilação envolvendo os dois principais tipos de HPV envolvidos com o câncer cervical, HPV16 e HPV18, com poucos estudos observando outros HPVs de alto risco. Em ambos os genomas se observa um aumento progressivo de metilação a partir de lesões assintomáticas, lesões benignas, pré-neoplásicas e câncer (Bryant *et al.*, 2014; Turan *et al.*, 2007).

Contudo, observa-se uma heterogeneidade de metilação no genoma viral derivado de amostras clínicas (Chaiwongkot *et al.*, 2013), e o papel exato da metilação no DNA de HPV continua não esclarecido.

O foco inicial dos estudos foi a análise do padrão de metilação na LCR, provavelmente devido ao seu papel na regulação dos oncogenes, e em sítios adjacentes à LCR, nas porções proximais dos genes L1 e E6. Vinakurova e Doeberitz (2011) mostraram, em estudo utilizando microdissecção de biópsias cervicais infectadas com HPV16, incluindo três amostras de tecidos adjacentes normais, três lesões de baixo grau e cinco epitélios em transformação, que o padrão de metilação em CpG presentes na LCR de HPV16 pode variar tanto durante a transformação neoplásica quanto nos estágios de diferenciação do epitélio. Em infecções assintomáticas, foi observado um padrão de elevada metilação em todos os CpGs nos três segmentos da LCR (5', central e 3') e em todas as camadas do epitélio (superficial, intermediária e basal). Nas lesões de baixo grau, foi observada uma diminuição de metilação nos segmentos 5' e 3' da LCR das camadas basal e intermediária, já na camada superficial manteve-se um padrão de elevada metilação no segmento 3'. No câncer, houve uma inversão de padrão de metilação, com baixos níveis de metilação na extremidade 3' e aumento nos níveis nas regiões 5' e central da LCR.

Foi observada também em alguns estudos uma associação inversa entre os níveis de metilação na LCR de HPV16 e a progressão do câncer. Piyathilake *et al.* (2014), utilizando pirosequenciamento e 315 amostras cervicais, reportando altos níveis de metilação em sítios CpG na região central e promotora da LCR (nt 7862, 31, 37, 43, 52 e 58) associados ao menor risco para o diagnóstico de NIC2+ (Piyathilake *et al.*, 2014). Xi *et al.* (2011), utilizando método qualitativo para análise de metilação, observaram decréscimo estatisticamente significativo do número de sítios CpG metilados na 3'LCR de HPV16 em amostras cervicais de pacientes com lesão de alto grau (NIC2/3) em comparação à pacientes sem NIC2/3. Nesse estudo, foram ainda incluídas análises de risco e que pacientes com quatro ou mais de quatro sítios CpG metilados na região 3'LCR apresentaram efeito protetor para aquisição de NIC2/3, mesmo após ajustar por fatores associados ao desenvolvimento do câncer cervical, incluindo variante de HPV16, número de visitas HPV16 positivas (indicando persistência), consumo de cigarro e número de parceiros sexuais (Xi *et al.*, 2011).

Entretanto, os dados sobre metilação em sítios CpG localizados na LCR de HPV16 são conflitantes, com ampla variação entre amostras clínicas analisadas e um número amostral reduzido de carcinomas invasivos. Consequentemente, até o momento não se observa um padrão de metilação único que possa indicar um papel específico na regulação gênica viral ou um bom biomarcador de progressão.

De outro modo, observa-se um consenso, em que a presença de metilação em sítios CpG no segmento 3'LCR de HPV16 estaria associada, em amostras de câncer cervical, ao estado físico epissomal do DNA viral (Chaiwongkot *et al.*, 2013). Acredita-se que o aumento nos níveis de metilação nesses casos, levaria ao bloqueio da atividade repressora de E2. Isso porque quatro sítios CpG presentes nessa região estão localizados em sítios de ligação da proteína E2 (E2BS#3 e E2BS#4) e a metilação nesses sítios bloquearia fisicamente a ligação de E2, mantendo a transcrição a partir do promotor P97 (Figura 2.2 A e B) (Hong *et al.*, 2008; Kalantari *et al.*, 2008; Chaiwongkot *et al.*, 2013; Jacquin *et al.*, 2013; Marongiu *et al.*, 2014).

Por outro lado, estudos têm apontando a metilação em sítios CpG em L1 e L2 de HPV18 e HPV16 como possíveis marcadores biológicos, indicando infecções com maiores chances de persistência e progressão neoplásica (Turan *et al.*, 2006; Turan *et al.*, 2007; Lorincz *et al.*, 2013). Nesses casos, a hipótese é que a metilação silenciaria a transcrição dos genes tardios, uma vez que os produtos de L1 e L2, proteínas do capsídeo viral, não são necessários para o processo neoplásico (Turan *et al.*, 2006).

Turan e colaboradores (2006) observaram o padrão de metilação em 11 sítios CpG presentes em L1 de HPV18 em infecções cervicais assintomáticas, lesões pré-neoplásicas, de baixo e alto grau, e em carcinomas cervicais. Nesse estudo, foi observada progressão nos níveis de metilação, com aumento de metilação nas lesões de alto grau e metilação total em carcinomas (Turan *et al.*, 2006).

Mirabelo *et al.* (2013) analisaram 66 CpGs ao longo do genoma de HPV16 por pirosequenciamento, incluindo LCR, E6, E7, E2/E4, L2 e L1. Foram analisadas 273 amostras de esfregaços cervicais diagnosticadas como controles (DNA de HPV16 em infecções transitórias), e casos (DNA de HPV16 em infecções persistentes sem NIC2+ e com NIC2+) em coorte de mulheres de Guanacaste, Costa Rica. Nesse estudo, foram observadas altas taxas de metilação, estatisticamente significantes, em mulheres com CIN2+ em comparação aos controles para vários sítios em L2, L1,

E2-E4, E6 e E7. Foram descritas boas taxas de sensibilidade e especificidade (90% e 60%, respectivamente) para detecção de mulheres com CIN2+, e para distinguir entre casos de CIN2+ e infecções transitórias, apresentando melhor performance principalmente em mulheres mais velhas. Dos 66 sítios CpG analisados no estudo, 24 foram descritos como fortes candidatos a marcadores ($AUC \geq 0,75$): para L1, 5611, 5726, 5927, 6367, 6389, 6457, 6581, 6650, 6790, 7136 e 7145; para L2, 4240, 4249, 4261, 4270, 4277, 4437, 4441, 5173 e 5179; para E6, 218; para E7, 790; e para E2 3433 e 3436 (Mirabello *et al.*, 2013).

Como observado para alguns sítios CpG em L2 e L1 de HPV16 e HPV18, foi reportada também associação positiva entre os níveis de metilação e a progressão do câncer cervical para outros tipos virais, HPV 45, 31, 33, 52 e 58 (Wentzensen *et al.*, 2012; Murakami *et al.*, 2013; Kalantari *et al.*, 2014; Vasiljevic *et al.*, 2014).

Considerando o que foi abordado, em que a metilação no DNA é um processo dinâmico, multifatorial e associada à progressão do câncer cervical, nossa hipótese é que os níveis de metilação no DNA de HPV, em amostras de câncer cervical, podem variar conforme a influência de fatores associados à agressividade do tumor, ao prognóstico da doença, e características clínicas da paciente. Aliado a isso, os dados na literatura são escassos no que diz respeito à influência de fatores clínicos/epidemiológicos nos níveis de metilação em DNA de HPV, com ênfase em DNA de HPV16, e utilizando número limitado de amostras clínicas de câncer cervical.

3. OBJETIVOS

3.1. OBJETIVO GERAL

Caracterizar, em amostras de tumores cervicais invasivos, a diversidade dos níveis de metilação em sítios CpG presentes na LCR, L2 e L1 de HPV16 e HPV18.

3.2. OBJETIVOS ESPECÍFICOS

- Caracterizar os níveis de metilação em sítios CpG presentes na LCR de HPV16 e HPV18 por pirosequenciamento.
- Caracterizar os níveis de metilação em sítios CpG presentes na LCR, L2 e L1 de HPV16 e HPV18 por NGS.
- Associar os níveis de metilação encontrados às características clínicas e biológicas da doença: tipo histológico, linhagem de HPV, integração do DNA de HPV ao genoma da célula hospedeira, e a idade da paciente.
- Caracterização dos haplótipos de metilação de HPV16 e HPV18.

4. METODOLOGIA

4.1. CASUÍSTICA

Foram analisadas 229 amostras cervicais, sendo 165 infecções únicas com HPV16 e 64 com HPV18, selecionadas a partir de biópsias de tumores da cérvix uterina de mulheres atendidas no Hospital do Câncer II entre os anos de 2008 e 2012, diagnosticadas com câncer cervical invasivo (FIGO IB1 a IVB). Essas amostras fazem parte de um projeto maior aprovado previamente pelo Comitê de Ética em Pesquisa do INCA (CEP número: 156/10) e intitulado “Estudo dos genótipos de HPV presentes em tumores do colo do útero em mulheres matriculadas no INCA” (Anexo 1), em que foi determinada a prevalência de genótipos de HPV em 594 tumores cervicais (Anexo 2).

Informações sócio-demográficas e clínicas das pacientes, incluindo idade, estágio do tumor (baseado na classificação FIGO), grau de diferenciação e tipo histológico (adenocarcinomas e carcinomas de células escamosas) foram obtidos dos dados clínicos das pacientes na instituição. Informações sobre sorologia para HIV também foram consultados. Informações sobre as linhagens de HPV foram obtidas por Vidal *et al.* (2016). Foram incluídos ao estudo, como controles em várias etapas, DNA extraído de linhagens celulares tumorais infectadas com HPV16 (CaSki e SiHa) e HPV18 (HeLa).

4.1.1. CRITÉRIOS DE INCLUSÃO DAS AMOSTRAS

O fluxograma apresentando o número de amostras selecionadas em cada grupo utilizando duas diferentes técnicas para análise dos níveis de metilação, pirosequenciamento e NGS, está apresentado na Figura 4.1.

Para análise de metilação por pirosequenciamento foram selecionadas 106 amostras de câncer cervical invasivo com HPV16 (N=69) e HPV18 (N=37) (Figura 4.1). O critério de seleção foi incluir aleatoriamente carcinomas de células escamosas (CCE) e adenocarcinomas (ADC) na proporção aproximada de 2:1, respectivamente, devido à maior frequência de CCE na populacional do estudo original de amostras com HPV16 e HPV18 (Anexo 2). Além disso, devido à baixa prevalência de linhagens B e C de HPV16, decidiu-se incluir todas as amostras com essas linhagens (20 amostras) a fim de obter informações suficientes para analisar a influência de linhagens nos níveis de metilação.

Para análise de metilação por NGS foram selecionadas 175 amostras de câncer cervical invasivo com HPV16 (N=123) e HPV18 (N=52) (Figura 4.1). O único

critério de seleção foi incluir carcinomas de células escamosas (CCE) e adenocarcinomas (ADC) na proporção aproximada de 2:1. As amostras foram selecionadas aleatoriamente dentro de cada tipo histológico, utilizando ferramenta disponível online de sorteio (<https://www.random.org>). Além disso, foram adotados critérios de qualidade dos níveis de metilação gerados NGS a fim de validar amostras para as análises estatísticas. Sendo assim, serão analisadas 71 amostras HPV16 positivas e 49 HPV18 positivas (Figura 4.1). Esses critérios serão apresentados no tópico “4.4.4. Controle de Qualidade dos Níveis de Metilação” da metodologia.

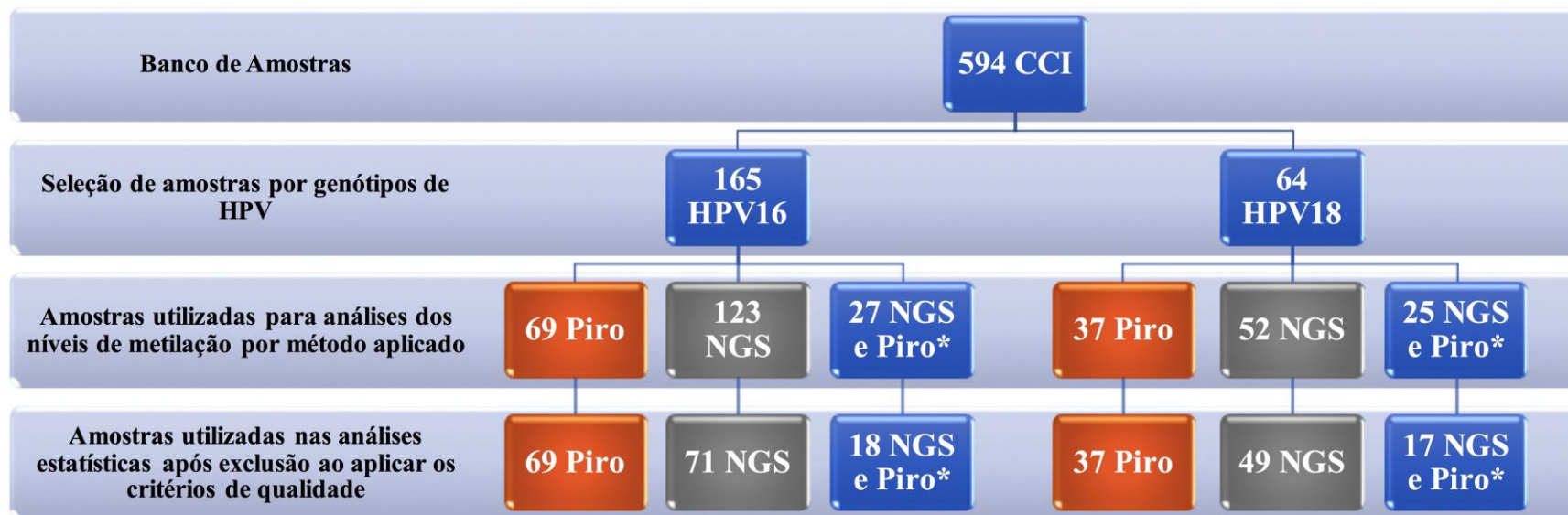


Figura 4.1. Fluxograma com o número de amostras selecionadas para as análises de metilação por pirosequenciamento e NGS. * NGS e Piro representam o número de amostras em comum testadas por ambas as técnicas.

4.2. TRATAMENTO COM BISSULFITO DE SÓDIO

A presença de metilação em citosinas de sítios CpG no DNA de HPV foi determinada tratando o DNA total extraído dos tumores e das linhagens celulares com bissulfito de sódio utilizando EpiTect Bissulfite Kit (cat. no. 59104, Qiagen, Germany), seguindo as orientações do fabricante. Posteriormente a região de interesse foi amplificada por PCR.

De forma resumida, 20 µL de solução contendo entre 500 a 1500 ng de DNA total extraído das amostras de cada paciente e linhagens foram submetidos a desnaturação térmica e conversão com bissulfito de sódio, seguindo etapa de ligação à coluna, lavagem e eluição em 40 µL de água fornecida pelo Kit e ao fim estocado a - 20^o C.

O tratamento com bissulfito de sódio atua convertendo quimicamente, através de etapas de sulfonação, deaminação e desulfonação, citosinas não metiladas em uracila enquanto as citosinas metiladas permanecem inalteradas. Como controle interno, alíquotas de água livre de nucleases foram incluídas no processo e convertidas com bissulfito de sódio para assegurar ausência de contaminantes. Além disso, DNA de linhagens de celulares provenientes de tumores cervicais infectados com HPV16 (CaSki e SiHa) e HPV18 (HeLa) foram também convertidos por bissulfito de sódio e utilizados como controles de metilação.

4.3. DETECÇÃO DE METILAÇÃO DE DNA POR PIROSEQUENCIAMENTO

4.3.1. REGIÃO DE INTERESSE E REAÇÃO DE AMPLIFICAÇÃO

Após a conversão com bissulfito de sódio, foi realizada etapa de amplificação por PCR cobrindo cinco sítios CpG presentes nas posições 31, 37, 43, 52 e 58 da região 3'LCR de HPV16 e nas posições 44, 50, 54, 60 e 66 de HPV18. Os ensaios de pirosequenciamento foram nomeados por amplicon: HPV16_LCR_Piro e HPV16_LCR_Piro. As posições dos nucleotídeos foram baseadas nos genomas de referência para HPV 16 (número de acesso: K02718.1) e HPV18 (número de acesso: X05015.1) disponíveis no banco de dados para papilomavírus, PaVE (do inglês Papilloma virus genome database, <https://pave.niaid.nih.gov>) e no GenBank. Para isso, foram utilizados pares de iniciadores sendo um deles biotilnado (senso ou anti-senso) para amplificar o DNA convertido por bissulfito de sódio (Tabela 4.1). Os

iniciadores biotilados foram purificados por HPLC pelo fornecedor Síntese Biotecnologia (Belo Horizonte, Brasil) para melhor performance.

Como controle inicial e padronização do tratamento por bissulfito de sódio, as regiões de interesse foram amplificadas por PCR a partir do DNA da linhagem celular CaSki tratado com bissulfito de sódio, que em seguida os produtos foram purificados e sequenciados pela metodologia de Sanger com a plataforma de sequenciamento automático ABI Prism 3130XL (Applied Biosystems), de acordo com o protocolo recomendado pelos fabricantes. Essa etapa foi importante para certificar que o tratamento estava sendo eficiente na conversão de citosinas não metiladas por uracila (Figura 4.2).

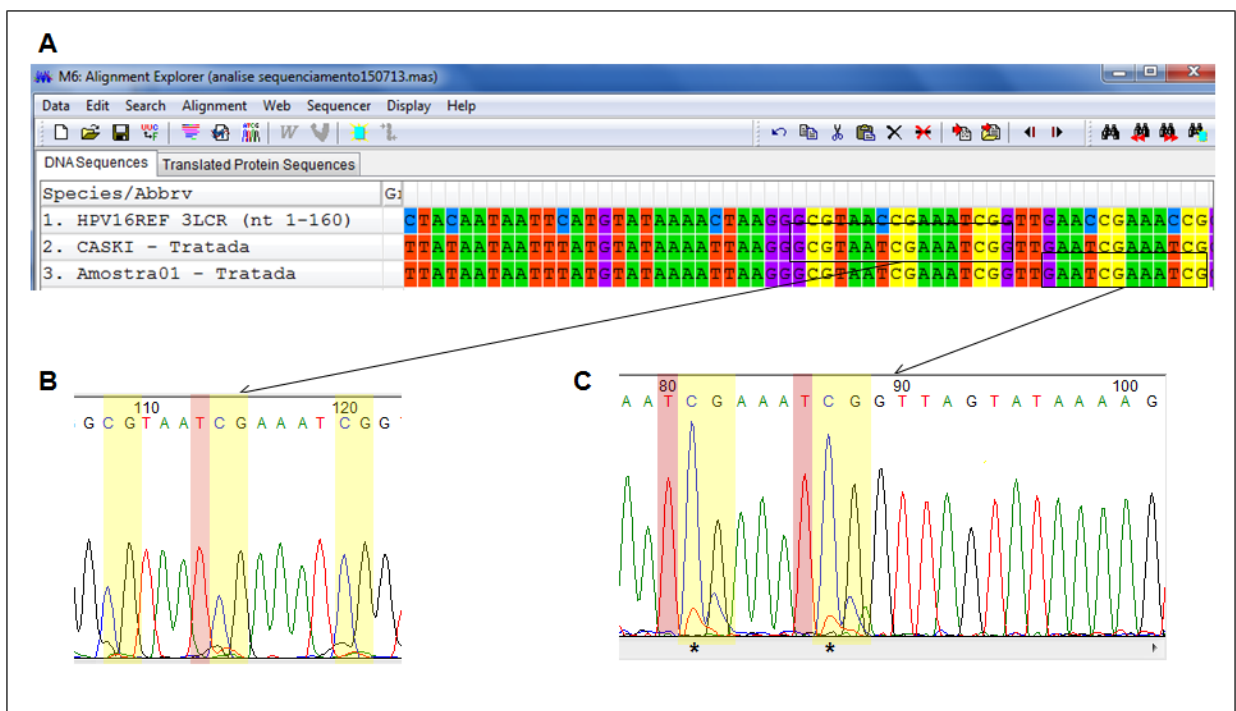


Figura 4.2. Etapa de checagem da conversão total por bissulfito de sódio. Em A, observa-se as seqüências obtidas da região 3'LCR de HPV16: referência sem conversão, HPV16REF (1), da linhagem CaSki tratada por bissulfito de sódio (2) e de uma das amostras também tratada (3); em amarelo estão os cinco sítios CpG (da esquerda para a direita posição nt 31, 37, 43, 52 e 58) e em vermelho as timinas (produtos da conversão). Em B está representado um segmento do eletroferograma da linhagem CaSki mostrando as citosinas (presentes nos sítios CpG) não convertidas para timina, e uma citosina (ao lado esquerdo do segundo CpG) convertida para timina. A mesma observação pode ser feita em C, entretanto, pode-se observar ainda a presença parcial de conversão das citosinas em CpG a partir da presença de picos menores de timinas sobrepostos por picos de citosina (*).

Para HPV16 foram utilizados conjuntos de iniciadores descritos na literatura (Rajeevan, Swan *et al.* 2006), produzindo amplicons de 178 pares de base. Para HPV18 iniciadores foram obtidos utilizando o programa PyroMark Assay Design versão 2.0 (Qiagen), utilizando o genoma de referência acima descrito e produzindo amplicons de 245 pares de base. Em todas as reações de amplificação preparou-se uma mistura com volume final de 30 μ L contendo 0,2 mM dNTPs (dATP, dCTP, dGTP e dTTP – Invitrogene), 6 pmol de cada primer, 1 U de Taq DNA polimerase Platinum (Life Technologies) e tampão 1 X com 67 mM Tris (pH 8,8), 6,7 mM MgSO₄, 16,6 mM (NH₄)₂SO₄, 10 mM 2-mercaptoethanol como previamente descrito (Kocher, Thomas *et al.* 1989) e água Milli-Q para completar o volume final da reação.

Os tubos contendo os componentes da reação foram posicionados em termociclador GeneAmp PCR System 9700 – Applied Biosystems, e submetidos ao seguinte esquema de amplificação: uma etapa inicial de 6 min a 95°C, seguido de 40 ciclos constituídos de 1 min a 95°C, 1 min a temperatura de anelamento (sendo: 61°C para HPV16 e 57°C para HPV18), e 1 min a 72°C. Em seguida, houve uma etapa final de 5 min a 72°C e os tubos foram armazenados a – 20°C após a amplificação. Para visualização das bandas foi utilizada a eletroforese em gel de agarose 2%, utilizando Gel Red (Uniscience) como corante fluorescente de ácidos nucléicos. O resultado foi observado em transiluminador de luz ultravioleta e fotografado utilizando Molecular Imagen Gel Doc XR Imaging System (Biorad, Hercules, EUA).

4.3.2. PIROSEQUENCIAMENTO

Os níveis de metilação da região de interesse determinado por pirosequenciamento consiste na análise por sequenciamento, durante a síntese da fita complementar, dos produtos biotinizados amplificados a partir de DNA tratado com bissulfito de sódio. No pirosequenciamento, quando um desoxirribonucleotídeo é inserido para síntese da nova fita de DNA, este libera um pirofosfato inorgânico (PPi) que, junto a uma ADP (adenosina difosfato), é convertido para ATP (adenosina trifosfato) pela ATP sulfúrilase, sendo este utilizado pela luciferase para oxidar a luciferina, produzindo um sinal de luz (Figura 4.3 B) capturado por uma câmera CCD (*charge-coupled device*) acoplada ao sistema. O sinal luminoso é detectado e apresentado como pico em um gráfico, com altura equivalente à intensidade de luz

emitida. Este pico de luz será proporcional à quantidade de nucleotídeos iguais inseridos em uma posição (Figura 4.3).

A reação de pirosequenciamento foi realizada através da plataforma *PyroMark Q24* (Qiagen, Germany) seguindo o protocolo descrito na literatura (Tost and Gut 2007, Tost and Gut 2007, Vasiljevic, Wu *et al.* 2011, Mirabello, Schiffman *et al.* 2013) e sob supervisão da Tecnologista Gisele Moledo, responsável pela Plataforma de Pirosequenciamento do Programa de Hematologia Oncologia Pediátrica. De forma resumida, 1 µL de *beads* de estreptavidina (GE Healthcare, Buckinghamshire, UK), 40 µL de tampão de ligação (Qiagen), 20 µL do produto da PCR e 19 µL de água foram misturados e incubados por 10 min em bloco de agitação a 1400 rpm. Na estação de trabalho *Biotage Q24 Vacuum Prep Workstation* os amplicons foram separados, desnaturados, lavados e adicionados à placa de microtitulação contendo 25 µL de tampão de pareamento com 0,3 µM de iniciadores de sequenciamento (Tabela 4.1). Os iniciadores de sequenciamento juntamente com os produtos biotinizados foram em seguida incubados a 80^o C por 2 min para que fossem hibridizados e resfriados à temperatura ambiente antes do pirosequenciamento. Os reagentes *PyroGold*, incluindo enzima, substrato e nucleotídeos, foram utilizados para a reação de pirosequenciamento. Os piogramas foram gerados e analisados com o programa *PyroMark Q24* (v.2.0.6.).

Para cada reação de pirosequenciamento foram incluídos amplicons produzidos a partir das linhagens de CaSki, SiHa e HeLa, utilizados como controles de hipermetilação e hipometilação para a região 3'LCR, respectivamente, como descritos na literatura (Badal, Chuang *et al.* 2003, Fernandez, Rosales *et al.* 2009).

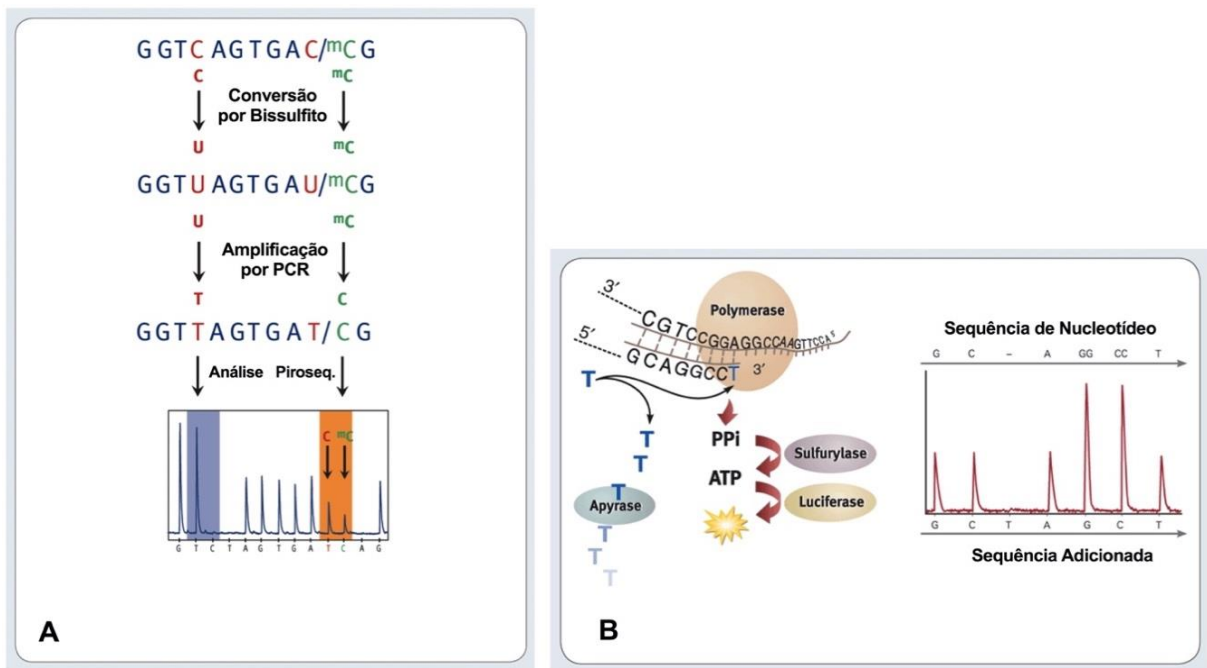


Figura 4.3. Princípio da análise de metilação e pirosequenciamento. A, citosinas não metiladas (C, vermelho) e metiladas (Cm, verde) são diferenciadas após conversão por bissulfito de sódio e PCR. A razão C/Cm em cada sítio CpG (picos em laranja na coluna) é medida no contexto da sequência. C não seguidos de G agem como controle para a etapa de conversão com bissulfito de sódio (coluna azul). Em B, enzimas envolvidas durante as etapas de pirosequenciamento. Picos apresentando o dobro do tamanho indicam incorporação de dois nucleotídeos na sequência analisada. Figura: www.nature.com/app_notes (doi:10.1038/nmeth800).

4.4. ENSAIOS DE METILAÇÃO DE DNA POR NGS

4.4.1. REGIÃO DE INTERESSE E REAÇÃO DE AMPLIFICAÇÃO

Duas regiões dentro dos genes tardios, L2 e L1 e uma região no segmento regulatório 3'LCR de HPV16 e HPV18 foram selecionadas para os ensaios de PCR e quantificação dos níveis de metilação por NGS. Foram utilizadas, para cada ensaio de amplificação, as seguintes nomenclaturas: HPV16_L2, HPV16_L1, HPV16_LCR, HPV18_L2, HPV18_L1 e HPV18_LCRE6 (Tabela 4.1). A localização dos sítios CpG correspondentes em cada um dos ensaios de HPV16 e HPV18 está representando na Figura 4.4.

Os iniciadores para amplificação de DNA convertido por bissulfito de sódio foram desenhados utilizando o programa MethPrimer (www.urogene.org/methprimer/), amplificando fragmentos contendo ao menos 4 sítios CpG nas ORFs L2 e L1, e 6 sítios CpG na região regulatória 3'LCR. Foram avaliados no total 33 sítios CpG no DNA de HPV16 e HPV18. Cada iniciador senso

foi desenvolvido contendo uma sequência específica de identificação (*barcodes*) com 12 nucleotídeos anexados de forma única na porção 5' do primer a fim de individualizar cada amostra a partir da amplificação por PCR. Os iniciadores reversos também foram marcados com 12 nucleotídeos para distinguir diferentes placas do mesmo ensaio, durante o texto será utilizado o termo em inglês *barcodes* para se referir a essas indexações. Em cada ensaio de PCR foram utilizados 96 iniciadores senso com *barcodes* e um primer anti-senso identificados com *barcodes*. Os Iniciadores foram sintetizados por Integrated DNA Technologies (IDT, Coralville, IA).

As reações de PCR foram realizadas utilizando o Kit Qiagen PyroMark (Qiagen, Valencia, CA), o qual permite o aumento na especificidade de anelamento dos iniciadores, e maior eficiência de amplificação em sequências de DNA convertido por bissulfato de sódio. Cada reação de PCR contve 1 μ L de DNA convertido, 2.5mM MgCl₂, 1X PyroMark Master Mix, 5 μ M de cada iniciador (senso e anti-senso) em um volume total de 25 μ L. As condições de PCR e temperaturas de anelamento para cada ensaio foram: uma etapa inicial de desnaturação de 15 min a 95°C, seguido de 50 ciclos constituídos de 30s a 94°C, 30s utilizando a melhor temperatura de anelamento de cada par de primer (Tabela 4.1), e 30s a 72°C. Em seguida, houve uma extensão final de 10 min a 72°C e as placas foram armazenados a – 20°C após a amplificação. Para visualização das amplificações foi utilizada a eletroforese em gel de agarose 3% (NuSieve e SeaKem Gold Agarose, 1:1), utilizando brometo de etídeo para identificação dos amplicons esperados. Controles positivos e negativos por placa foram incluídos para avaliar a performance da PCR. Além disso, foram incluídas duplicatas (5% a 10% do número total de amostras por PCR) a fim de verificar a reprodutibilidade do método. Cinco microlitros de cada reação de PCR foram reunidos formando um conjunto de amostras para cada placa de PCR produzida e cada ensaio. Com o objetivo de eliminar produtos de PCR inespecíficos, 100 μ L de todo o conjunto dos amplicons foi submetido a eletroforese em gel de agarose 3%. O conjunto de fragmentos do tamanho esperado (358pb) foi purificado através de excisão em gel utilizando Qiagen QIAquick Gel Extraction Kit (Qiagen, Valencia, CA) de acordo com o protocolo do fabricante.

Tabela 4.1. Informações dos iniciadores utilizados nos ensaios de análise de metilação por pirosequenciamento e NGS.

Ensaio	Metodologia		Iniciadores (5'-3') ^b	Posição CpG (nt)	Amplicon (pb) ^c	TA °C
HPV16_LCR_Piro ^a	Pirosequenciamento	senso	TTGTA AAAATTGTATATGGGTGTG	31, 37, 43, 52, 58	178	61
		anti-senso (biotinilado)	AAATCCTAAAACATTACAATTCTC			
		sequenciamento (senso)	AATTTATGTATAAAAATTAAGGG			
HPV18_LCR_Piro	Pirosequenciamento	senso	ATTTTAAATATGAATTATAATATGATTAAG	44, 50, 54, 60, 66	245	57
		anti-senso (biotinilado)	CACAAATCAAATAACTTATAAAAATC			
		sequenciamento (senso)	GTAGTATATAAAAAAGGGAGTAA			
HPV16L2_1	NGS - Illumina	senso	TTTTTGT TTTGT TTTGT TTTGT TTTT	4240, 4249, 4261, 4270	164	58
		senso_sublinhagem D2	TTTTTTTTTATTTTTGTGTTTTT			
		senso_sublinhagem D3	TTTTTTTTTATTTGTGTGTTTTT			
		anti-senso	ACATATACCTACCTATTTACATATTTTATA			
HPV16L1_1	NGS - Illumina	senso	TAATATATAATTATTGTTGATGTAGGTGAT	5602, 5608, 5611, 5617	148	58
		anti-senso	AACAACCAAAAAACATCTAAAAAA			
HPV16_LCR	NGS - Illumina	senso	TTGTA AAAATTGTATATGGGTGTG	7862, 31, 37, 43, 52, 58	226	59
		anti-senso	AAATCCTAAAACATTACAATTCTC			
HPV18_L2_4	NGS - Illumina	senso	TATATGTAGATGATATGGATTTTGTAGTGT	5331, 5333, 5367, 5406	183	58
		anti-senso	ATACAAACACATCCCAAAAAAAA			
HPV18_L1_2	NGS - Illumina	senso	TTTTAGATTTAGATTAATATTTTTTTGGA	7011, 7038, 7041, 7062,	241	58
		anti-senso	ACAAAACATACAAACACAATAAATATA	7068, 7090, 7110, 7116, 7122		
HPV18_LCRE6	NGS - Illumina	senso	ATTTTAAATATGAATTATAATATGATTAAG	7824, 44, 50, 54, 60, 66, 109,	279	59
		anti-senso	CACAAATCAAATAACTTATAAAAATC	111, 129 e 132		

^aOligonucleotídeo desenhado a partir da literatura (Rajeevan *et al.*, 2006).^bOs iniciadores de amplificação referente aos ensaios de LCR são os mesmos utilizados por pirosequenciamento e NGS, entretanto, por pirosequenciamento foram utilizados iniciadores anti-senso biotinilados. ^cO tamanho do amplicon gerado para fins da análise de NGS-Illumina inclui, além de Iniciadores e sequência amplificada, a adição de 12 pb referente às sequências únicas de identificação de cada amostra (*barcodes*). Os ensaios se referem aos amplicons gerados por tipos de HPV e regiões do genoma analisado. As posições dos sítios CpG foram baseadas nas referências disponíveis no PaVE (números de acesso: K02718.1 – HPV16, X05015.1 – HPV18).

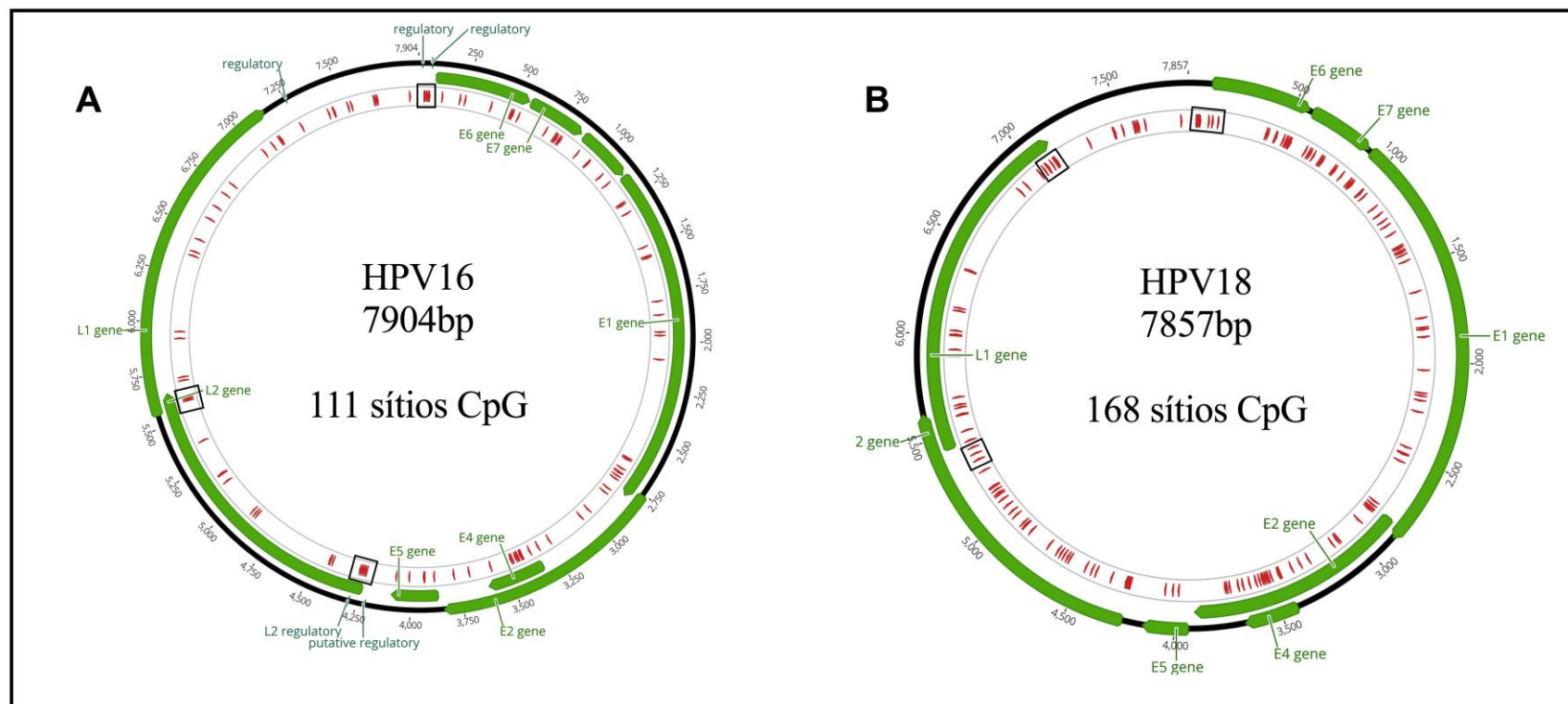


Figura 4.4. Identificação dos sítios CpG analisados nos genomas de HPV16 e HPV18. O posicionamento dos sítios CpG analisados em cada ensaio de análise de metilação nos genomas de HPV16 (A) e HPV18 (B) estão representados em caixas pretas. Desenho e contagem dos sítios através do programa Geneious 10.0.9. Número de acesso dos genomas de referência: K02718.1 (HPV16) e X05015.1 (HPV18).

4.4.2. PREPARO DA BIBLIOTECA DE DNA

Para o preparo das bibliotecas foram agrupados dois conjuntos de amplicons purificados com tamanhos próximos em uma mesma biblioteca. No primeiro conjunto foram incluídos os ensaios HPV16_L1 (148pb), HPV16_L2 (164pb) e HPV18_L2 (183pb), e no segundo conjunto, os ensaios HPV18_L1 (241pb), HPV18_LCRE6 (279pb) e HPV16_LCR (226pb).

As bibliotecas foram preparadas utilizando Kapa Biosystems KAPA HTP Library Preparation Kit Illumina platforms (Kapa Biosystems, Boston, MA) de acordo com as instruções dos fabricantes. De forma breve, 100-500ng de DNA fita-dupla foram submetidos às seguintes etapas enzimáticas durante a construção da biblioteca: (a) reparo das extremidades do DNA, para produzir extremidades cegas formando fragmentos de DNA dupla-fita fosforilados na extremidade 5', (b) adição de dAMP (deoxiadenosine monofosfato) na extremidade cega 3' (utilizaremos o termo do inglês *A-tailing*), (c) ligação, em que o adaptador Illumina TRUSeq INDEXED DNA com extremidades 3'-dTMP coesivas são ligados aos fragmentos com extremidades de 3' ligados à dAMP, (d) seguido de etapa de amplificação com 8 ciclos, permitindo um aumento no rendimento de DNA.

Precedendo a reação de ligação e após a etapa final de amplificação, os fragmentos foram purificados utilizando esperas magnéticas SPRI (Agencourt AMPure XP; Beckman Coulter, Brea, CA) e precipitação com etanol 80%, seguindo eluição em 25 uL. As bibliotecas foram sequenciadas utilizando reação de 300pb *paired-end* em Plataforma Illumina My-seq do Albert Einstein Epigenomics Core Facility. Por fim, a qualidade e quantificação das bibliotecas foram verificadas com Bioanalyzer utilizando Kit para análise de DNA de alta sensibilidade (Figura 4.5).

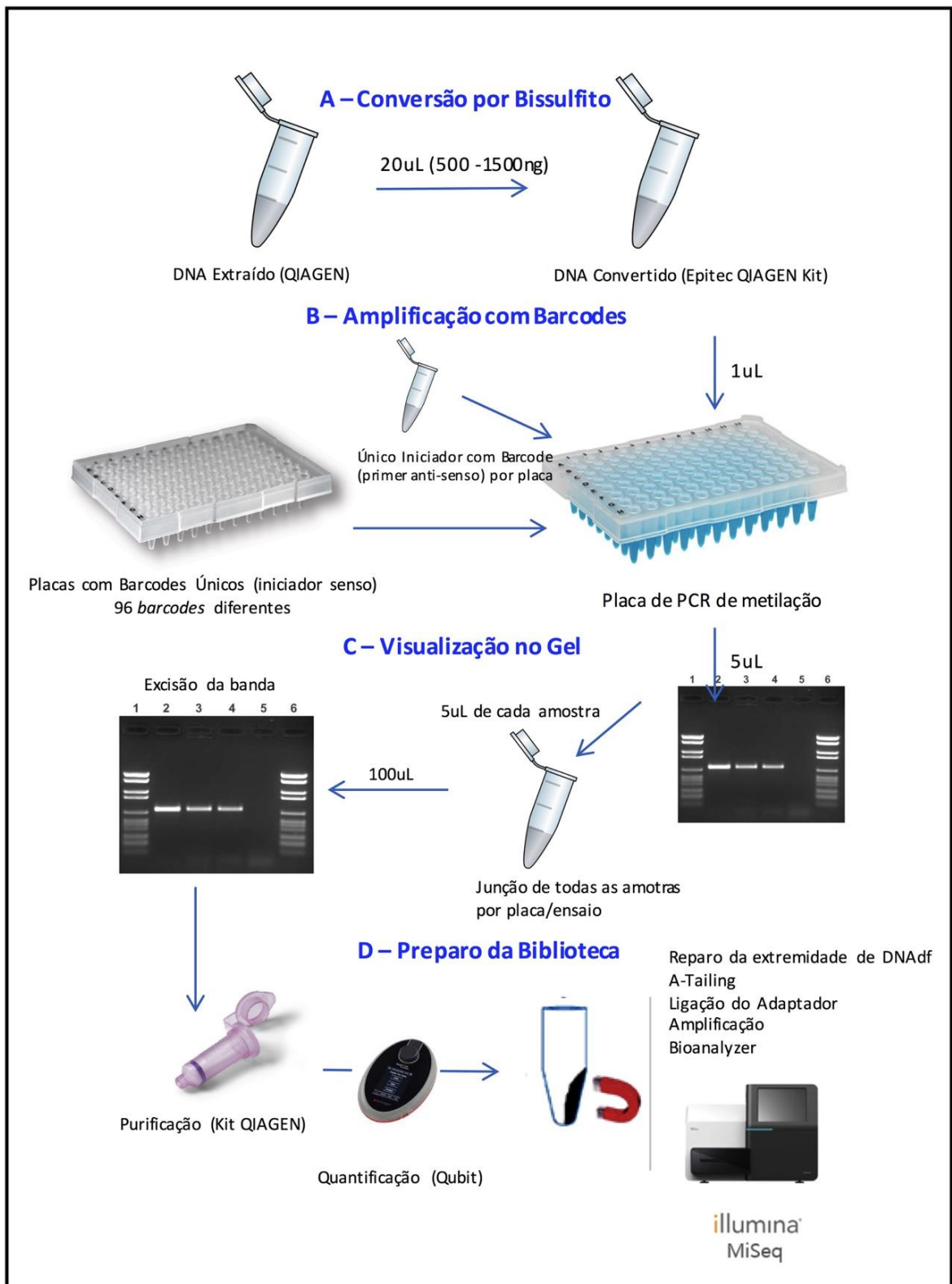


Figura 4.5. Princípio das análises de metilação por NGS. Os dados de metilação em sítios CpG por NGS foram gerados seguindo as seguintes etapas: conversão por bissulfito de sódio (A), amplificação por PCR utilizando iniciadores com sequências únicas (“barcodes”) de para identificação de cada amostras (B), visualização das amplificações em gel de agarose, seguido de formação de junção os produtos de PCR por placa e excisão em gel (C), preparo de biblioteca com adição de adaptadores com index de reconhecimento pela

plataforma Illumina, incluindo verificação da qualidade dos amplicons utilizando Bioanalyzer e sequenciamento (D).

4.4.3. BIOINFORMÁTICA

Arquivos com dados gerados por NGS através da plataforma Illumina Mi-Seq foram primeiramente filtrados, excluindo leituras de baixa qualidade, utilizando a média mínima de *PHRED score* de 25. O algoritmo PHRED calcula um grau de confiabilidade para cada base identificada, com base na qualidade do sinal que foi captada, sendo este valor denominado *PHRED score*. As etapas seguintes incluíram demultiplexação a partir das sequências únicas de 12pb de DNA (*barcodes*), para a identificação das leituras (*reads*) de cada amostra, e etapa de construção de biblioteca de referência convertida por bissulfito de sódio. Referências foram produzidas a partir das sequências de DNA genômico de HPV16 e HPV18, utilizando Bismark. Bismark é um programa para mapear leitura de sequências tratadas por bissulfito de sódio em um genoma de interesse e executar a previsão de metilação em um único passo (<http://www.bioinformatics.babraham.ac.uk/projects/bismark/>). Em seguida, foram mapeadas as leituras *paired-end* com o objetivo de determinar o status de metilação das citosinas em sítios CpG e não-CpG. Scripts utilizando o programa Python e desenvolvidos pelo grupo do Dr. Robert Burk, foram utilizados em um *pipeline* para processar os resultados gerados pelo Bismark através de: (i) produção de um padrão de metilação para cada molécula sequenciada, indicando se as citosinas nos ensaios foram metiladas (+), não metiladas (-), sem informação (o) ou em discordância com as leituras de cada direção (senso e anti-senso) de leitura (x), (ii) geração de contagem para cada padrão por amostra, (iii) montagem dos percentuais de metilação para cada citosina por amostra, através da comparação das razões de citosinas metiladas (C) pelo total do número de metiladas e não-metiladas (C+T) em cada fragmento de interesse analisado. A razão de C/(C+T) indica a proporção de citosinas metiladas em cada sítio CpG na amostra analisada. A combinação de sítio CpG metilados em uma única molécula, combinando com a quantificação de metilação em todos os sítios CpG, constituem os haplótipos de metilação. Isso está relacionado com a combinação do status de metilação do sítio CpG (+, C metilado; -, C não-metilado) determinado a partir de uma única leitura. Por exemplo, um padrão "+ - + -" com um número de leituras de 1400 representam 1400 moléculas únicas de DNA sequenciadas pelo Illumina em que o primeiro sítio CpG e o terceiro estão metilados, enquanto o segundo e quarto não estão metilados para cada leitura (Frimer *et al.*, 2015). Dados adicionais sobre os haplótipos de

metilação podem ainda prover informações adicionais dos padrões de metilação de cada CpG por fragmento além de inferir se alguns CpG são mais relevantes que outros.

4.4.4. CONTROLE DE QUALIDADE DOS NÍVEIS DE METILAÇÃO

Para cada tipo de HPV foram realizados 03 ensaios referentes a ampliações de segmentos específicos de L2, L1 e LCR. Em cada ensaio foram incluídos controles internos para validação dos dados por placa de PCR:

- a) Ensaios de HPV16: 06 controles negativos, 02 controles positivos (DNA da linhagem celular SiHa) e 13 duplicatas.
- b) Ensaios de HPV18: 03 controles negativos e 05 duplicatas.

Os critérios de qualidade dos níveis de metilação por NGS estabelecidos para validação de amostras para as análises estatísticas foram: (i) número de *reads*, amostras validadas apresentaram ao menos duas vezes o número de *reads* dos controles negativos, (ii) presença de amplificação por PCR, visualizada através de banda em gel de agarose, (iii) amostras de ensaios que apresentaram forte correlação entre as amostras em duplicatas, sendo assim, selecionadas amostras com alta carga viral (CT <22), apenas para ensaios de HPV16.

4.5. ANÁLISE DE INTEGRIDADE DAS ORFs E1/E2 DE HPV

A integridade do DNA viral foi determinada utilizando pares de iniciadores cobrindo as ORFs E1 e E2 de HPV16 e HPV18, descritas como sítios com maior frequência de ruptura para integração no genoma da célula hospedeira (Cullen *et al.*, 1991; Xu *et al.*, 2013; Liu *et al.*, 2016). Foram sintetizados, com base na literatura, cinco pares de iniciadores flanqueando a região E1 (E1a, E1b, E1c, E1d e E1e) e três para a região E2 (E2a, E2b e E2c) de HPV16, e cinco pares de iniciadores cobrindo a região E2 (E2P1, E2P2, E2P3 e E2P4) de HPV18 (Vernon *et al.*, 1997; Collins *et al.*, 2009; Cricca *et al.*, 2009). A região E1 de HPV18 foi amplificada utilizando cinco pares de iniciadores (E1P1, E1P2, E1P3, E1P4 e E1P5) desenvolvidos pelo nosso grupo através da ferramenta do NCBI (Primer-Blast, <http://www.ncbi.nlm.nih.gov/tools/primer-blast/>), como apresentado na Tabela 4.1. Amostras com ausência de pelo menos um amplicon foram denominadas com ruptura e amostras com presença de todos os amplicons intactos (Figura 4.6).

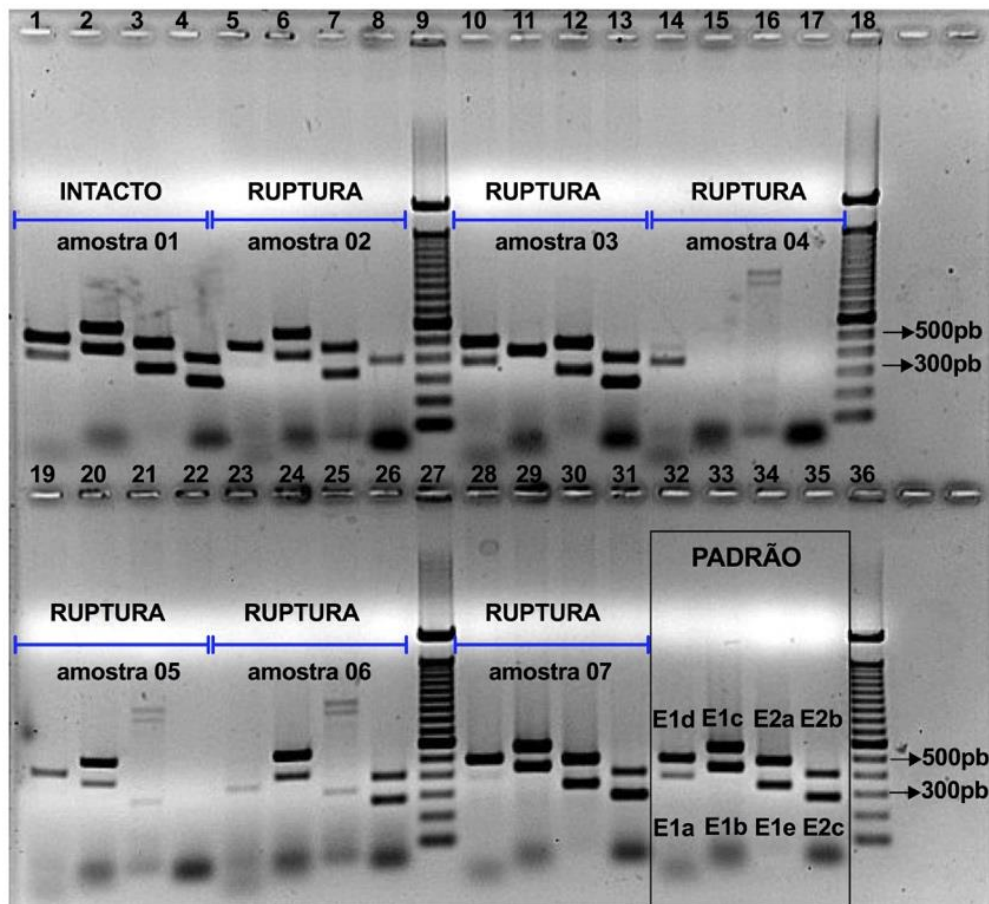


Figura 4.6. Representação da análise de integridade de E1 e E2 de HPV16. A presença das ORFs E1/E2 com ruptura ou intactas foi avaliada por PCR seguindo eletroforese em gel de agarose. Na figura, o gel superior e inferior apresenta exemplos de amostras com E1/E2 intacto ou com ruptura. A amostra 1 (linhas 1-4) apresenta presença de todos os amplicons, E1a (356pb, linha 1), E1b (409pb, linha 2), E1c (536pb, linha 2), E1d (455pb, linha 1), E1e (316pb, linha1), E2a (451pb, linha 3), E2b (379pb, linha 4) e E2c (257pb, linha 4). De outro modo, as amostras 02-07 (linhas 5-8, 10-13, 14-17, 19-22, 23-26, 28-31) retrata a ausência de pelo menos um amplicon. O tamanho em pb de todos os amplicons está representada na caixa em destaque no canto direito inferior do gel. Peso molecular de 100pb (*ladder*) está apresentado nas linhas 9, 18, 27 e 36. Foram incluídas duas reações de PCR, com amplicons de tamanhos diferentes, em uma mesma linha.

Tabela 4.2. Informações sobre os iniciadores utilizados nos ensaios de análise da integridade das ORFs E1 e E2 de HPV16 e HPV18.

Oligonucleóides			Sequências (5'-3')	Posição (nt)	Amplicon (pb)
HPV16	E1a	senso	CCATGGCTGATCCTGCAG	863-1219	356
		anti-senso	TCTCCTTTTTGCAGCTCT		
	E1b	senso	GACAGCGGGTATGGCAAT	1254-1663	409
		anti-senso	CATTCCCCATGAACATGC		
	E1c	senso	AATAAATCAACGTGTTGCGATTGG	1548-2084	536
		anti-senso	GTTTATAATGTCTACACATTGTTG		
	E1d	senso	GGATTGTGCAACAATGTG	2072-2527	455
		anti-senso	TGGAGGGCATTTTAGTTG		
	E1e	senso	CAACTAAAATGCCCTCCA	2529-2845	316
		anti-senso	CGCATGTGTTTCCAATAG		
	E2a	senso	CGAGGACAAGGAAAACGA	2738-3189	451
		anti-senso	CTTGACCCTCTACCACAG		
	E2b	senso	GGTTTATATTATGTTTATGAAGG	3220-3599	379
		anti-senso	TATGGGTGTAGTGTACTATTACA		
E2c	senso	GTAATAGTAACACTACACCCATA	3596-3853	257	
	anti-senso	GGATGCAGTATCAAGATTTG			
HPV18	E1P1	senso	GGTGTGCATCCCAGCAGTAA	888-1403	515
		anti-senso	GCCGCCACTACATACATTGC		
	E1P2	senso	GCGGCAATGTATGTAGTGGC	1400-1908	508
		anti-senso	GCTGCAACACTACTTCGCAA		
	E1P3	senso	TCAACCACCAAATGCGAAGT	1877-2211	334
		anti-senso	TCGTTTTTGGGCTCGCCTAT		
	E1P4	senso	GCAAACATTATAGGCGAGCCC	2181-2546	365
		anti-senso	TGTCCAACACGTGGTCGTT		
	E1P5	senso	GGTGGCCATGTTAGATGATGC	2506-2895	389
		anti-senso	GATTTTGTCCCTGCAACGCACT		
	E2P1	senso	TCCAGATTAGATTTGCACGA	2786-3192	407
		anti-senso	CAATTGTCTTTGTTGCCATC		
	E2P2	senso	ATACAAAACCGAGGATTGGA	3086-3388	303
		anti-senso	ACTTCCCACGTACCTGTGTT		
	E2P3	senso	AACACAGGTACGTGGGAAGT	3369-3739	371
		anti-senso	TTTCGCAATCTGTACCGTAA		
E2P4	senso	GACCTGTCAACCCACTTCT	3598-3994	397	
	anti-senso	ACATGGCAGCACACATACAT			

pb, pares de base; iniciadores cobrindo E1 e E2 de HPV16 derivaram de Vermon *et al.*(2007) e os iniciadores flanqueadores de E2 de HPV18 foram baseados do trabalho de Collins Constandinou-Williams *et al.*(2009).

4.6. CARGA VIRAL DE HPV16

A carga viral de HPV16 foi avaliada por PCR em tempo real (qPCR) utilizando o sistema TaqMan. As reações de qPCR foram desenhadas para amplificar fragmentos de 64 pb em E6 de HPV16 utilizando iniciadores desenvolvidos através do programa Primer Express Software Version 3.0 (Applied Biosystems, Life Technologies, Grand Island, NY). As seqüências dos iniciadores utilizadas foram: senso GGTGGACCGGTCGATGTATG e anti-senso: TACAGCTGGGTTTCTCTACGTGTT; e as sondas: HPV16E6 MGB: 6FAM-CTTGTTGCAGATCATC-NFQ. O gene RNase P humano foi utilizado como controle constitutivo. Para isso, foi utilizado o Kit O Applied Biosystems® TaqMan® RNase P Detection Reagents kit (PN 4316831) que fornece componentes necessários para detectar e quantificar cópias de DNA genômico humano do gene RNase P utilizando o sistema TaqMan, incluindo alíquota de DNA controle humano (10ng/uL).

As reações de qPCR foram realizadas em duplicatas por amostra e por gene de interesse (E6 e RNase P). DNA de SiHa com número de cópias conhecido de HPV16 foi utilizada para validar cada ensaio de qPCR. Além disso, foram incluídas curvas de seis diluições seriadas com DNA de SiHa e DNA genômico humano disponibilizado pelo Kit (3, 3x10, 3x10², 3x10³, 3x10⁴, 3x10⁵), a fim de validar e determinar a quantificação de cada gene.

Cada reação de qPCR conteve ~100ng de DNA, 5uL de 2X Fast Advanced TaqMan Master Mix (Applied Biosystems, PN 4444557), 90uM de cada iniciador (senso e anti-senso), 25uM da sonda MGB de HPV16 e 1X RNase P Primer-Probe Mix (P/N 4316831) em um volume total de 10uL. As reações de qPCR foram corridas em 7500FAST *instrument* (ABI) conforme as condições determinadas pelo fabricante: 95°C por 10 min, seguido de 40 ciclos de 95°C por 15 sec e 60°C por 1 min. Dados foram analisados utilizando o programa 7500 ABI, V2.0.6.

A carga viral foi determinada por milhões de células através da seguinte fórmula: (quantidade de cópias de HPV16 ÷ (quantidade de cópias de RNase P ÷ 2)) X 10⁶.

4.7. ANÁLISE ESTATÍSTICA

O nível de metilação de cada sítio CpG por amostra foi estimado como a percentual de citosinas metiladas, variando em proporções de 0 (sem metilação) a 100

(totalmente metiladas). Análise estatística exploratória foi realizada comparando medianas através do teste de Mann-Whitney U Test, a fim de descrever os níveis de metilação em sítios CpG por amostra e possíveis associações por variável de estudo, incluindo idade, integridade de E1/E2, grau de diferenciação, histologia, estadiamento e linhagem de HPV.

Foram realizadas análises bivariadas e multivariadas com o uso de modelos estatísticos de Quase-Verossimilhança, utilizando função de ligação logit, para avaliar a influência das variáveis analisadas nos níveis de metilação.

O modelo de Quasi-Verossimilhança, que pertence à classe dos modelos lineares generalizados, foi utilizada pois não é necessário pressupor normalidade dos dados ou qualquer distribuição de probabilidade para a variável dependente. Esse tipo de modelo é indicado para dados de proporção no desfecho (metilação), que, no presente estudo, é o nível de metilação nos sítios CpG.

Tendo como desfecho os níveis de metilação, esse modelo estatístico avaliou, para os dados gerados por pirosequenciamento, o efeito da integridade de E1/E2 e da idade, como variáveis de explicação, nos níveis médios de metilação dos cinco sítios CpG em 3'LCR de HPV16. Em modelos multivariados, foram consideradas variáveis de controle o tipo de tumor (CCE/ADC) e linhagem de HPV porque as amostras não foram selecionadas de forma aleatória. O mesmo modelo foi aplicado para os dados de metilação gerados por NGS. Análise bivariada foi utilizada a fim de selecionar, para análise multivariada, os sítios CpG com possíveis diferenças nos níveis de metilação para cada variável de estudo, selecionando sítios com $p\text{-valor} \leq 0,21$. As análises foram realizadas individualmente para cada sítio CpG estudado em L2, L1 e LCR de HPV16. As variáveis testadas foram: idade, integridade de E1/E2, grau de diferenciação, histologia, estadiamento e linhagem de HPV. As análises utilizando o modelo estatístico não foi realizada para HPV18 devido à limitação do tamanho amostral em relação ao número de casos com E1/E2 intacto, o que reduziria o poder para se encontrar uma diferença.

O nível de significância utilizado nas análises foi de 0,05, exceto para a seleção das variáveis para os modelos multivariados, que considerou $p=0,20$. As análises

exploratórias e gráficos foram conduzidos no GraphPad Prism 7 e os modelos foram gerados no software R v.3.2.4.

5. RESULTADOS

Os resultados serão apresentados conforme os dados gerados por métodos de análise de metilação, pirosequenciamento e sequenciamento de nova geração (NGS). Nos resultados será caracterizada a população de estudo específica de cada grupo analisado, além dos critérios de seleção das amostras, a descrição dos níveis de metilação em DNA de HPV16 e HPV18 e a caracterização das diferenças nos níveis de metilação em relação a variáveis descritas como características possivelmente associadas à agressividade do tumor, ao prognóstico da doença, e a idade da paciente ao diagnóstico. Os dados produzidos sobre os níveis de metilação e características clínicas e biológicas por amostra estão apresentados nos anexos 5 a 12.

5.1. ANÁLISES DE METILAÇÃO POR PIROSEQUENCIAMENTO

5.1.1. CARACTERIZAÇÃO DA POPULAÇÃO DE ESTUDO

A idade média das pacientes com amostras de tumores com HPV16 e HPV18 foi de 44 ($\pm 13,4$) e 48 ($\pm 13,3$) anos, respectivamente. Ambos com maior predominância de tumores em estágio mais avançado (\geq II), 68% (47/69) e 89% (33/37), respectivamente. Para HPV16, 38% (26/69) das amostras apresentaram HPV16 da linhagem A, 27% (19/69) linhagens B e C (Africana) e 35% linhagem D (24/69). Para HPV18, 70% (26/37) foram da linhagem A (Asiática Ameríndia, Asiática Americana e Européia, denominadas linhagens Europeias nas análises de associação) sendo o restante, 22% (8/37), linhagem B (Africanas, denominadas linhagens Não-Europeias nas análises de associação) (Tabela 5.1).

5.1.2. INTEGRIDADE DE E1/E2 DE HPV

Dos 69 tumores cervicais com HPV16 selecionados para a análise de metilação por pirosequenciamento, 46 (69%) foram amostras com os genes E1/E2 intactos de HPV16. Em 21 tumores (31%) foram observados sítios de ruptura nos genes E1 e/ou E2. Em duas amostras os resultados foram inconclusivos e foram excluídos da análise de associação entre os níveis de metilação e integridade de E1/E2. Das 21 amostras com ruptura, 13 (62%) ocorreram apenas em E1, enquanto 02 amostras (9%) com

ruptura em E2 exclusivamente, e 06 amostras (29%) com ruptura em ambos os genes. Diferentemente, a análise de integridade de HPV18 mostrou predominância de amostras com ruptura em E1 ou E2 (86% - 32/37), sendo apenas 05 (14%) com os genes intactos. A ruptura ocorreu principalmente em E2 (66% - 21/32), seguido da ruptura em ambos os genes (31% - 10/32) e apenas um caso com ruptura exclusiva em E1 (3%).

5.1.3. DESCRIÇÃO DOS NÍVEIS DE METILAÇÃO EM 3'LCR DE HPV16 E HPV18

A análise de metilação por pirosequenciamento foi determinada em sítios CpG presentes na região 3'LCR de HPV16 e HPV18, segmento em que se localiza o promotor dos oncogenes virais E6 e E7. Foi observado uma maior mediana de metilação em HPV16 em relação ao HPV18 para todos os sítios CpG, com diferença estatisticamente significativa entre os sítios CpG homólogos, associados a sítios ligação da proteína E2, nt 37 vs 44 ($p=0,001$), 43 vs 50 ($p=0,049$) e 58 vs 66 ($p=0,010$) (Figura 5.2). Além disso, os níveis de metilação na LCR de HPV16 apresentaram maior amplitude entre as amostras, 0% a 90%, enquanto para HPV18 variou de 0% a 72%. Não se observou diferença dos níveis de metilação entre os sítios CpG dentro de cada tipo de HPV. As linhagens celulares, utilizadas como controles de validação durante as análises de metilação, apresentaram níveis de metilação conforme descrito na literatura para este segmento da região LCR (Badal *et al.*, 2003; Fernandez *et al.*, 2009). CaSki apresentou um padrão de hipermetilação com níveis de metilação entre 90% a 94%, SiHa e HeLa foram observados níveis reduzidos de metilação, de 0 a 4%, sendo consideradas hipometiladas.

Tabela 5.1. Características das amostras selecionadas da população de estudo agrupadas por técnica de análise de metilação e tipo de HPV.

Características	Pirosequenciamento		NGS		NGS (após seleção)*	
	HPV16	HPV18	HPV16	HPV18	HPV16	HPV18
Total	69	37	123	52	71	49
Idade (anos)						
Média (DP)	44 (13,4)	47 (12,8)	48 (13,3)	47 (11,7)	48 (14)	46 (11)
Mediana	43	45	48	45	49	44
Tipo de Tumor (N)						
CCE	44	25	79	34	43	31
ADN	25	10	41	18	25	18
Sem informação	-	2	3	-	3	-
Linhagem de HPV (N)						
A	26	26	62	40	37	38
B	9	8	4	6	3	6
C	10	NA	3	NA	-	NA
D	24	NA	35	NA	27	NA
Sem informação	-	3	19	6	4	5
Estágio FIGO (N)						
I	22	4	38	11	23	11
II	23	15	43	16	26	15
III	21	17	33	24	16	22
IV	3	1	9	1	6	1
Sem informação	-	-	-	-	-	-
Grau de Diferenciação (N)						
G1	6	3	6	7	4	7
G2	45	18	77	28	45	25
G3	11	10	22	13	14	13
Sem informação	7	6	18	4	8	4

ADN: adenocarcinomas, CCE carcinoma de células escamosas; N=número de amostras; NA=não se aplica

Número de amostras utilizadas em ambas técnicas: 27 (HPV16) e 25 (HPV18)

*Distribuição das amostras selecionadas após exclusão utilizando os critérios de qualidade e que foram utilizada nas análises estatísticas.

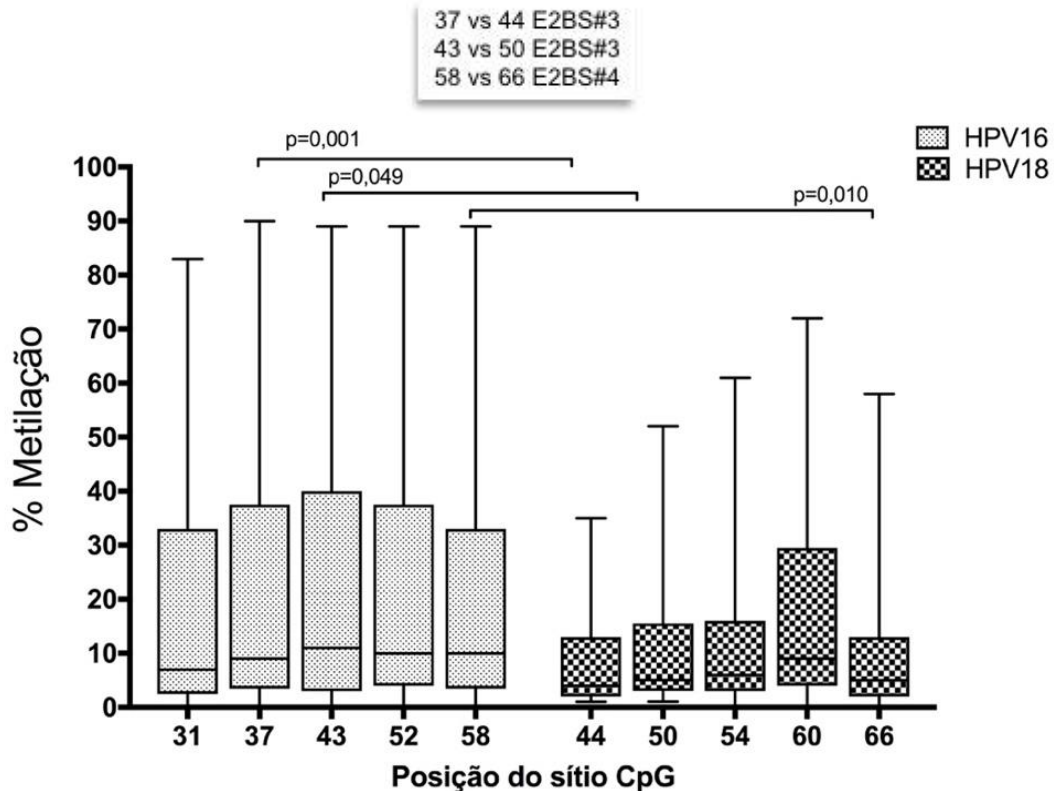


Figura 5.1. Distribuição dos níveis de metilação analisados por pirosequenciamento em 5 sítios CpG presentes na LCR de HPV16 e HPV18 em amostras tumorais. A caixa representa os quartis superiores e inferiores separados pela mediana. Nas extremidades se encontram os valores máximos e mínimos. As análises estatísticas foram realizadas com o teste de *Mann-Whitney*, comparando sítios CpG de HPV16 vs HPV18 com funções homólogas (CpGs presentes em sítios de ligação de E2). Foi observada diferença estatisticamente significativa, $p < 0,05$, dos níveis de metilação entre 37 vs 44, 43 vs 50 e 58 vs 66. E2BS: sítio de ligação da proteína E2.

5.1.4. ASSOCIAÇÃO DOS NÍVEIS DE METILAÇÃO NA LCR DE HPV16 E HPV18 COM O TIPO DE TUMOR, O ESTÁGIO TUMORAL, A VARIANTE DE HPV E INTEGRIDADE DE E1/E2

DNA de HPV16 em adenocarcinomas apresentaram maiores medianas de metilação do que carcinomas de células escamosas, em todos os sítios CpG da região LCR. Para HPV18 não foi observado o mesmo padrão (Figura 5.2 A). Amostras de pacientes com tumores em estágio inicial (I) apresentaram medianas de metilação maiores quando comparados aos estágios avançados (II-IV) na LCR de HPV16 e, mais sutilmente, para HPV18 (Figura 5.2 B). Linhagens europeias (A) de HPV16 apresentaram maiores níveis de metilação em relação às não-europeias (B/C/D).

Para HPV18, o efeito foi inverso com as linhagens não-européias (B) apresentando maiores níveis de metilação (Figura 5.2). Entretanto, não foi observada significância estatística entre os níveis de metilação e as variáveis citadas acima (linhagem de HPV, tipo histológico e estágio do tumor).

Diferentemente, amostras com ORFs E1/E2 intactas apresentaram níveis de metilação maiores. Essa diferença foi observada entre todos os sítios CpG de HPV16 ($p < 0,001$) e para os sítios nt 50 ($p = 0,049$) e 54 ($p = 0,040$) de HPV18 (Figura 5.2 D).

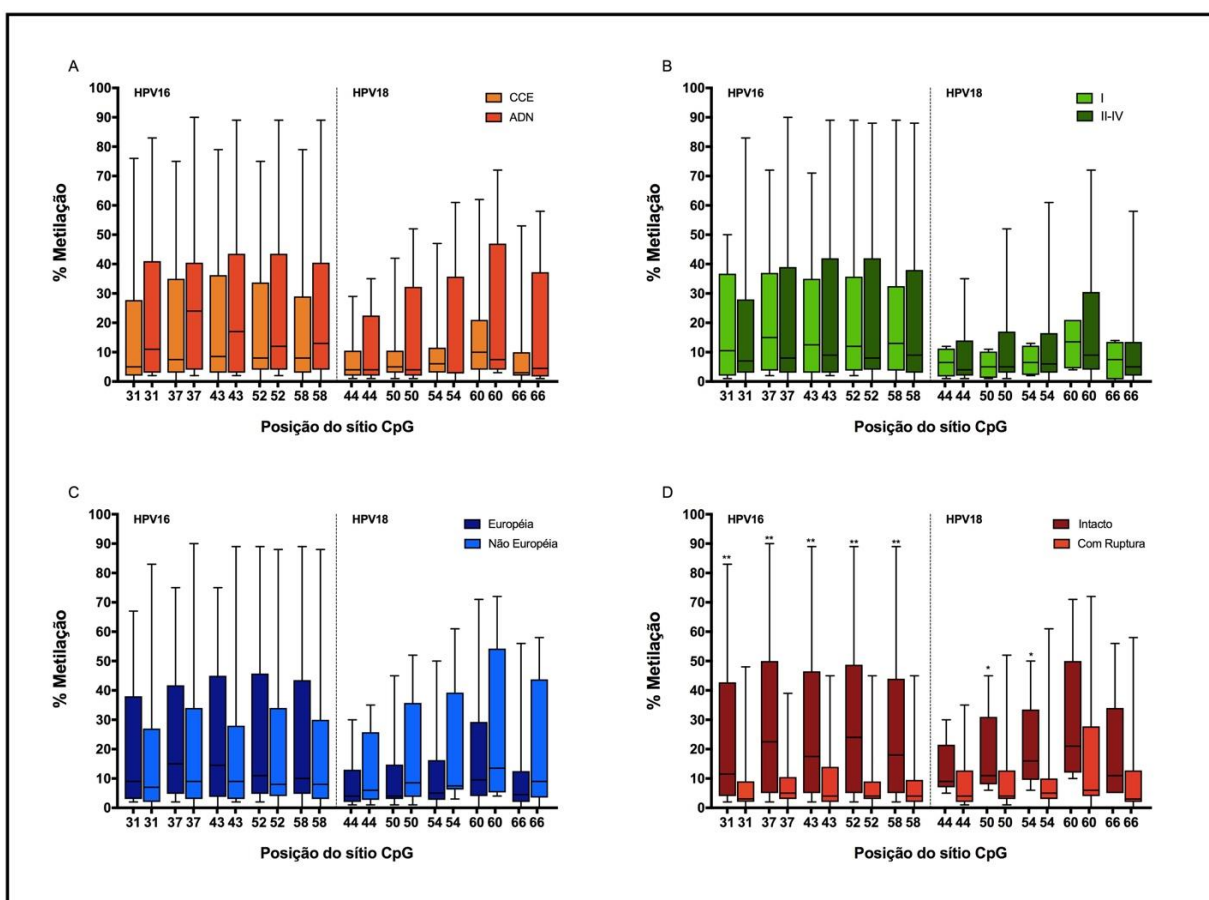


Figura 5.2. Associação dos níveis de metilação analisados por pirosequenciamento em 10 sítios CpG presentes na LCR de HPV16 e HPV18 por variáveis de estudo. A caixa representa os quartis superiores e inferiores separados pela mediana. Nas extremidades se encontram os valores máximos e mínimos. Não foi observada diferença estatisticamente significativa, $p > 0,05$, utilizando o teste de *Mann-Whitney*, entre os níveis de metilação por tipo histológico, estágio tumoral e variante de HPV (A, B e C). Amostras com DNA de HPV intacto, sem ruptura de E1/E2, apresentaram maiores níveis de metilação, diferença estatisticamente significativa, $p < 0,05$ (*) (D). ADC=adenocarcinomas e CCE=carcinomas de células escamosas.

5.1.5. CORRELAÇÃO DOS NÍVEIS DE METILAÇÃO NA LCR DE HPV16 E HPV18 COM A IDADE DA PACIENTE

A mediana de idade das pacientes em tumores com HPV16 e HPV18 foi de 43 e 45 anos, respectivamente (p -valor = 0,306) (Tabela 5.1). Foi observada uma correlação positiva entre os níveis médios de metilação dos cinco sítios analisados individualmente na LCR de HPV16 e a idade da paciente, com os sítios nt 37 e 58 apresentando maior correlação ($r = 0,326$ e p -valor = 0,006). Essa correlação não foi observada para HPV18 (Figura 5.3).

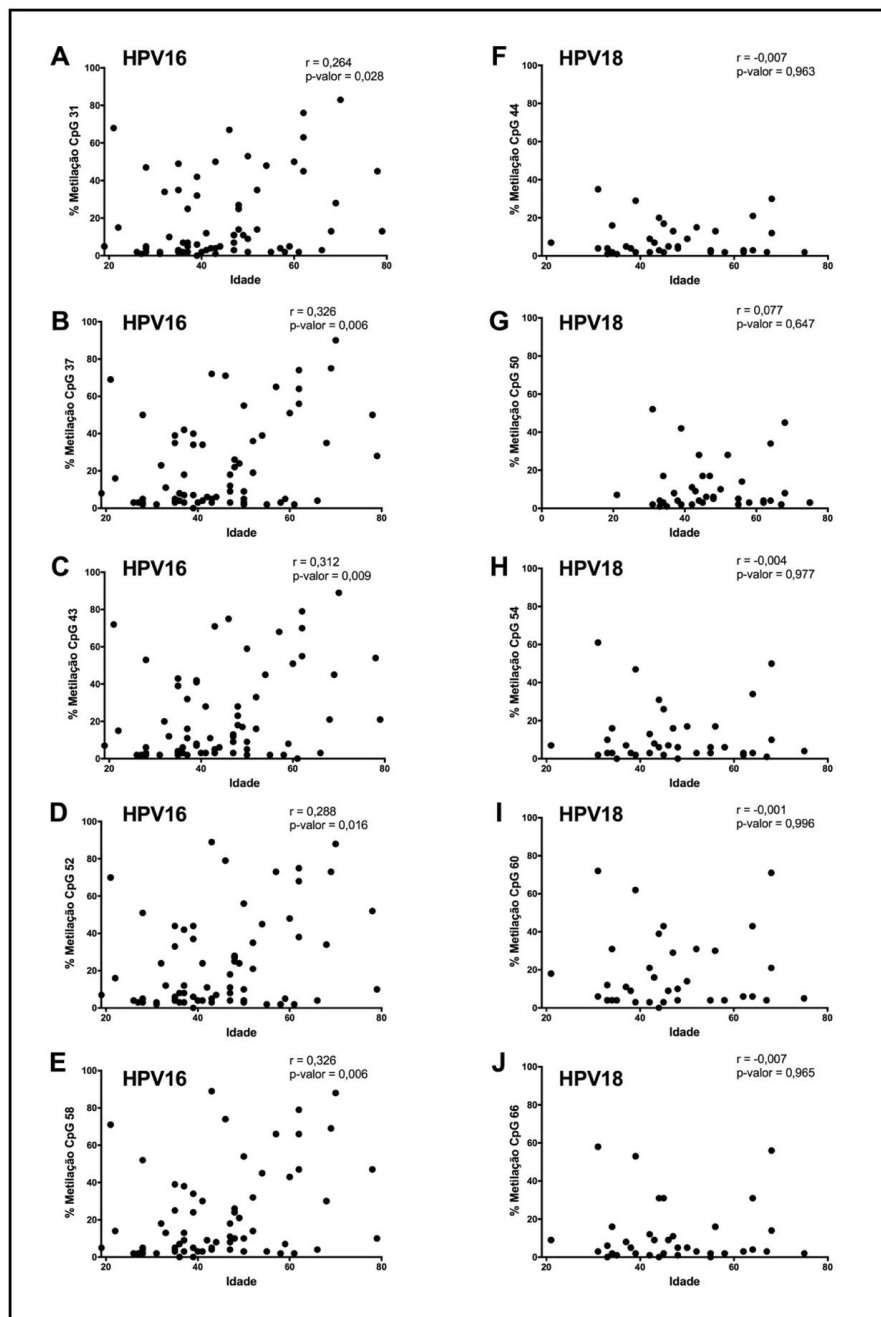


Figura 5.3. Distribuição e correlação dos níveis de metilação, analisados por pirosequenciamento, e a idade da paciente no momento do diagnóstico. Os níveis de

metilação de todos sítios CpG analisados na LCR de HPV16 (A) apresentaram uma correlação de Spearman positiva com a idade. Essa correlação não foi observada para HPV18 (B).

5.1.6. ANÁLISE BIVARIADA E MULTIVARIADA UTILIZANDO MODELO DE QUASI-VEROSSIMILHANÇA

Considerando que os critérios de seleção das amostras podem camuflar possíveis associações não observadas e que as variáveis analisadas no estudo podem contribuir diferentemente para os níveis de metilação, uma análise bivariada e multivariada, utilizando modelo de Quasi-Verossimilhança, foi realizada a fim de confirmar e compreender melhor a importância da integridade de E1/E2 e da idade da paciente nos padrões de metilação da região 3'LCR de HPV16. Essa análise não foi realizada para HPV18 devido à limitação do tamanho amostral, em que apenas 14% (5/37) apresentaram os genes E1/E2 intactos. Para HPV16, as análises foram aplicadas utilizando **as médias de metilação dos 5 sítios CpG** por amostra, já que apresentam comportamento similar nos níveis de metilação entre eles, sem diferença estatisticamente significativa (Figura 5.2). A análise bivariada mostrou que a presença dos genes E1/E2 intactos e idade da paciente estiveram positivamente associados com os níveis médios de metilação (p-valor = 0,003 e 0,006, respectivamente) (Tabela 5.2). Com a análise multivariada, após o controle por tipo de tumor e linhagem de HPV16, ambas variáveis foram também positivamente associadas com os níveis médios de metilação, de forma independente (p-valor = 0,002 e 0,021, respectivamente) (Tabela 5.2). Esses achados mostraram que presença de E1/E2 íntegros está associada aos altos níveis de metilação, e que pacientes mais velhos também apresentaram maiores níveis de metilação do que os mais novos, independente do fator de integridade.

Tabela 5.2. Análises bivariadas e multivariadas para associação dos níveis de metilação por variáveis de estudo utilizando pirosequenciamento.

	Análise Bivariada			Análise Multivariada*		
	Coefficiente	Erro Padrão	p-valor	Coefficiente	Erro Padrão	p-valor
Integridade de E1 e E2						
Com Ruptura	1	-		1	-	
Intacto	1,373	0,438	0,003	1,422	0,443	0,002
Idade	0,034	0,012	0,006	0,028	0,012	0,021

*Ajustado para o tipo de tumor (CCE e ADC) e linhagem de HPV16 (A vs B/C).

5.2. ANÁLISES DE METILAÇÃO POR NGS

5.2.1. CARACTERIZAÇÃO DA POPULAÇÃO DE ESTUDO

A idade média das pacientes foi de 48 ($\pm 13,3$) para amostras com HPV16 e 47 ($\pm 11,7$) para amostras com HPV18. Ambos com maior predominância de tumores em estágios avançados (\geq II), 69% (85/123) e 79% (41/52), respectivamente para HPV16 e HPV18. Para HPV16, 50% (62/123) das amostras selecionadas eram da linhagem A, 43% da linhagem D (54/123) e 7% (7/123) linhagens B e C (linhagens Africanas). Para HPV18, 77% (40/52) foram da linhagem A e 23% da linhagem B (8/52) (Tabela 5.1).

5.2.2. INTEGRIDADE DE E1/E2 DE HPV

Para o grupo de amostras selecionadas para análise de metilação por NGS das 123 amostras HPV16 positivas, a integridade das ORFs E1 e E2 foi avaliada em 78 delas (ou seja, as com dados de NGS dentro dos critérios de qualidade estabelecidos), sendo predominantemente DNA com E1/E2 intactas (62% - 48/78). Das amostras com ruptura (38% - 30/78), mais da metade foram com ruptura exclusiva em E1 (53% - 16/30). Em 13 amostras (43% - 13/40) foi observada ruptura em ambas ORFs e apenas um caso (3%) com ruptura exclusivamente em E2. Das 52 amostras HPV18 positivas selecionadas, 41 foram analisadas a integridade de E1/E2, sendo observada predominância de DNA com ruptura (80% - 33/41). Em 22 amostras a ruptura aconteceu exclusivamente em E2 (67% - 22/33), 9 em ambas as ORFs (27% - 9/33) e apenas 2 amostras com ruptura exclusivamente em E1 (6% - 2/33).

5.2.3. REPRODUTIBILIDADE DOS DADOS DE NGS

Das 123 amostras HPV16 positivas utilizadas para a análise por NGS, após exclusão pelo critério de qualidade (descrito na metodologia), baseado no número de *reads*, amplificação por PCR, carga viral e correlação entre as duplicatas, foram selecionadas para as análises: 66 amostras referentes ao ensaio HPV16_L2; 70 amostras ao ensaio HPV16_L1; e 67 amostras do ensaio HPV16_LCR. Para HPV18, das 52 amostras originais, foram analisadas: 47 amostras referentes ao ensaio HPV18_L2; 45 amostras ao ensaio HPV18_L1; e 42 amostras ao ensaio HPV18_LCR. Para cada ensaio, foi observada forte correlação entre as duplicatas mostrando uma excelente reprodutibilidade do experimento (Figura 5.4).

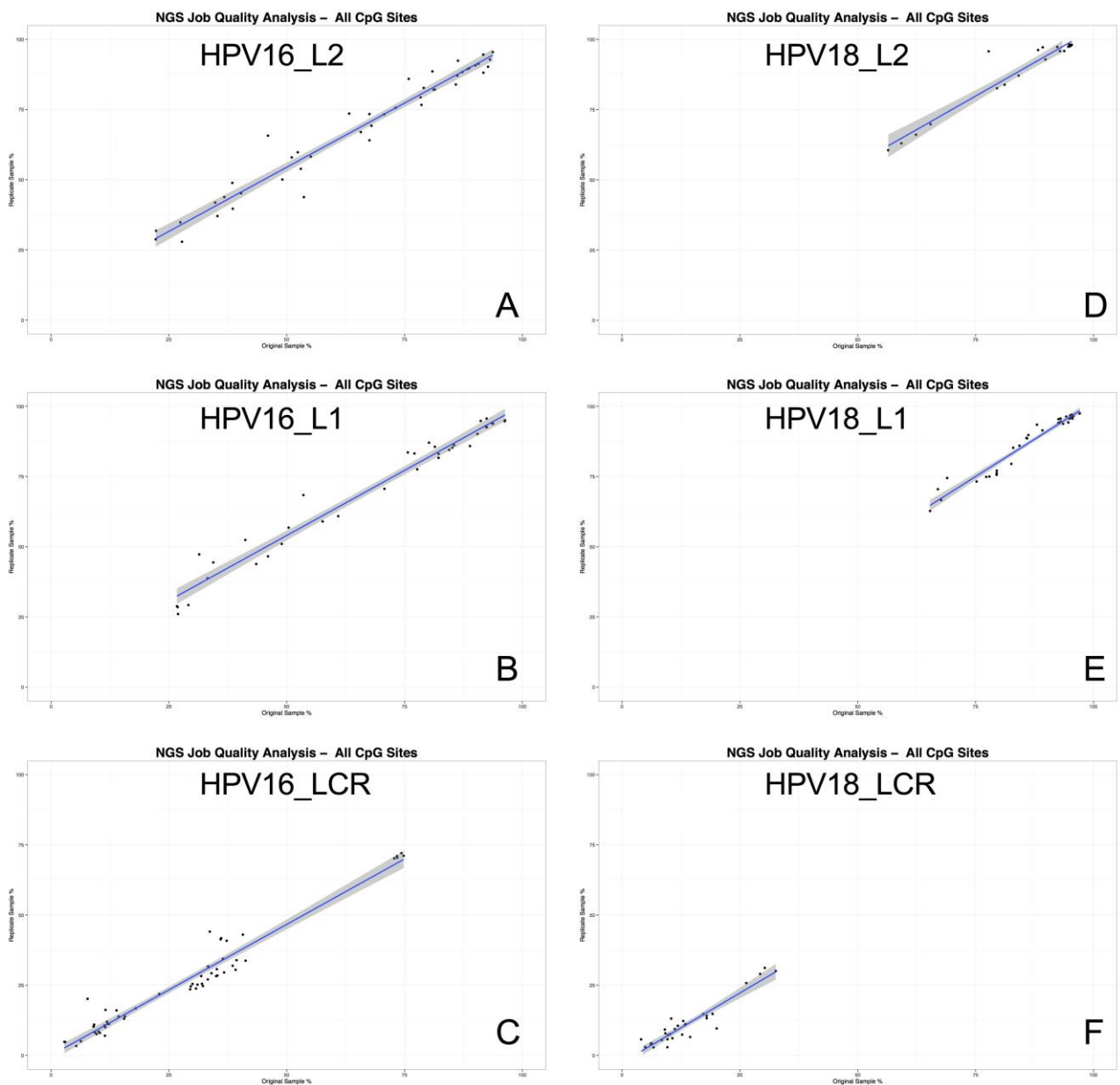


Figura 5.4. Controle de reprodutibilidade: correlação entre os níveis de metilação entre as duplicatas de amostras para os diferentes ensaios. Foi observada forte correlação entre as duplicatas incluídas por ensaio de PCR, HPV16_L2 (A, $r^2 = 0,956$, $p < 0,001$), HPV16_L1 (B, $r^2 = 0,968$, $p < 0,001$), HPV16_LCR (C, $r^2 = 0,948$, $p = 0,003$), HPV18_L2 (D, $r^2 = 0,927$, $p = 2,5 \times 10^{-05}$), HPV18_L1 (E, $r^2 = 0,951$, $p = 0,071$), HPV18_LCR (F, $r^2 = 0,883$, $p = 1,1 \times 10^{-05}$).

5.2.4. DESCRIÇÃO DOS NÍVEIS DE METILAÇÃO EM L1, L2 E 3'LCR DE HPV16 E HPV18

Foram analisados 15 sítios CpG nas regiões L2, L1 e LCR de HPV16. Os maiores níveis de metilação ocorreram em sítios em L2 e L1, com 04 dos 09 sítios apresentando medianas de metilação $\geq 70\%$ (nt 4240, 4277, 5602 e 5611). O menor nível de metilação foi observado na LCR de HPV16 (mediana de 13%, nt 7862).

Entretanto, houve ampla variação nas medianas de metilação dos sítios CpG por região analisada (L2, 33% a 75%; L1, 46% a 75%; LCR, 13% a 32%) (Figura 5.5 A). Além disso, ampla variação nos níveis de metilação por amostra também foi observada (6% - 96%).

Para os ensaios com HPV18, foram analisados 19 sítios CpG em L2, L1 e LCR. Da mesma forma que HPV16, L2 e L1 de HPV18 apresentaram medianas de metilação superiores a LCR, entretanto, com menor variação entre cada região analisada (L2, 92% a 94%; L1, 66% a 91%; LCR, 6% a 10%) (Figura 5.5 B). Em L2 de HPV18, os quatro sítios analisados apresentaram medianas de metilação superiores a 90%. Em L1 de HPV18, sete dos nove sítios apresentaram medianas maiores que 80%. Na LCR de HPV18, as medianas de metilação variaram de 6% a 10%. Foi observada também menor amplitude nos níveis de metilação na LCR por amostra, variando de 2% a 64%, do que em HPV16.

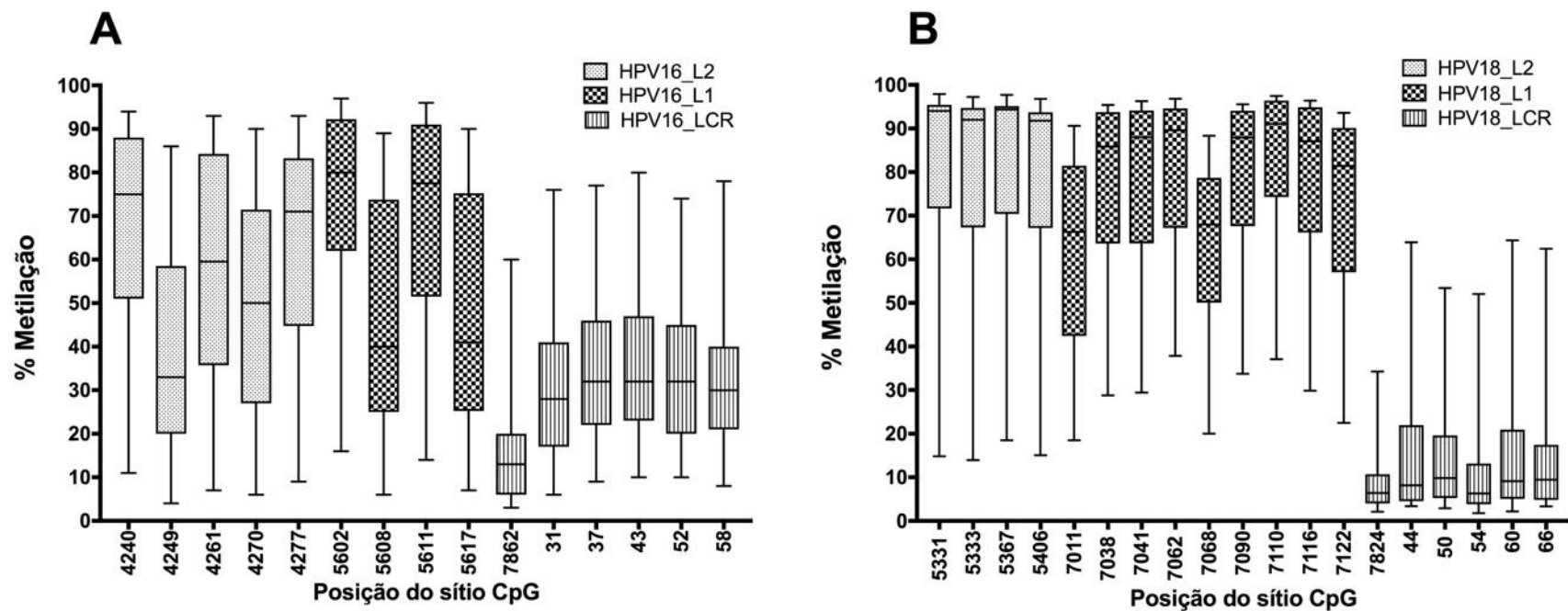


Figura 5.5. Distribuição dos níveis de metilação analisados por NGS em sítios CpG presentes em L2, L1 e LCR de HPV16 e HPV18. A caixa representa os quartis superiores e inferiores separados pela mediana. Nas extremidades se encontram os valores máximos e mínimos. Para ambos os genótipos virais, L2 e L1 apresentaram maiores níveis de metilação quando comparados a LCR. Entretanto, L2 e L1 de HPV16 apresentou menor uniformidade, com medianas de metilação atingindo de 33% a 75% (A). Para HPV18, L2 e L1 apresentaram medianas de metilação variando de 68% a 98% (B).

5.2.5. DESCRIÇÃO DOS HAPLÓTIPOS DE METILAÇÃO EM L2, L1 E LCR DE HPV16 E HPV18

Os valores representam as médias de frequências dos haplótipos de metilação para as amostras analisadas em cada ensaio. Para L2 e L1 de HPV16 o haplótipo mais frequente nas amostras foi o que apresentou todos os sítios metilados, 27% e 35%, seguido dos haplótipos sem metilação, 20% e 18%, respectivamente. Em L2 de HPV16, o haplótipo + - + + + foi o terceiro tipo mais frequente com média de 11% entre as amostras. O sítio CpG não metilado correspondente é o da posição 4249. Metilação no sítio CpG 4240 esteve presente nos haplótipos mais representados, em mais de 50% da média das frequências dos haplótipos. Em L1 de HPV16, o terceiro haplótipo mais frequente (14%) foi + - + - , sendo os sítios 5608 e 5617 não metilados, o primeiro sítio CpG 5602 esteve metilado na maior parte dos haplótipos (~67%) (Figura 5.6 A e B).

Na LCR de HPV16, o haplótipo mais frequente (47%) foi o sem metilação em todos os sítios seguido de ausência de metilação no primeiro sítio (13%), (- + + + + +). O sítio corresponde à posição 7862, localizado no sítio E2BS#2, não estava metilado em ~80% dos haplótipos. Já os demais sítios, localizados em E2BS#3, E2BS#4 e SP1, apresentaram frequência equivalente entre os haplótipos com ~28% (Figura 5.6 C).

Para L2 e L1 de HPV18 o haplótipo mais frequente nas amostras apresentou todos os sítios metilados, 68% e 39%, seguido dos haplótipos sem metilação em todos os sítios CpGs de L2 (9%) e com ausência de metilação no sítio 7011 (- + + + + + +) (Figura 5.7 A). Foi observada também reduzida frequência de haplótipos (12%) sem metilação em CpG 5406 em L2 (Figura 5.7 A).

Na LCR de HPV18, como na LCR de HPV16, o haplótipos mais frequente foi o sem metilação em todos os sítios, seguido do haplótipo com metilação apenas no terceiro sítio CpG, - - + - - - - - -, posição 50 (3%) (Figura 5.7).

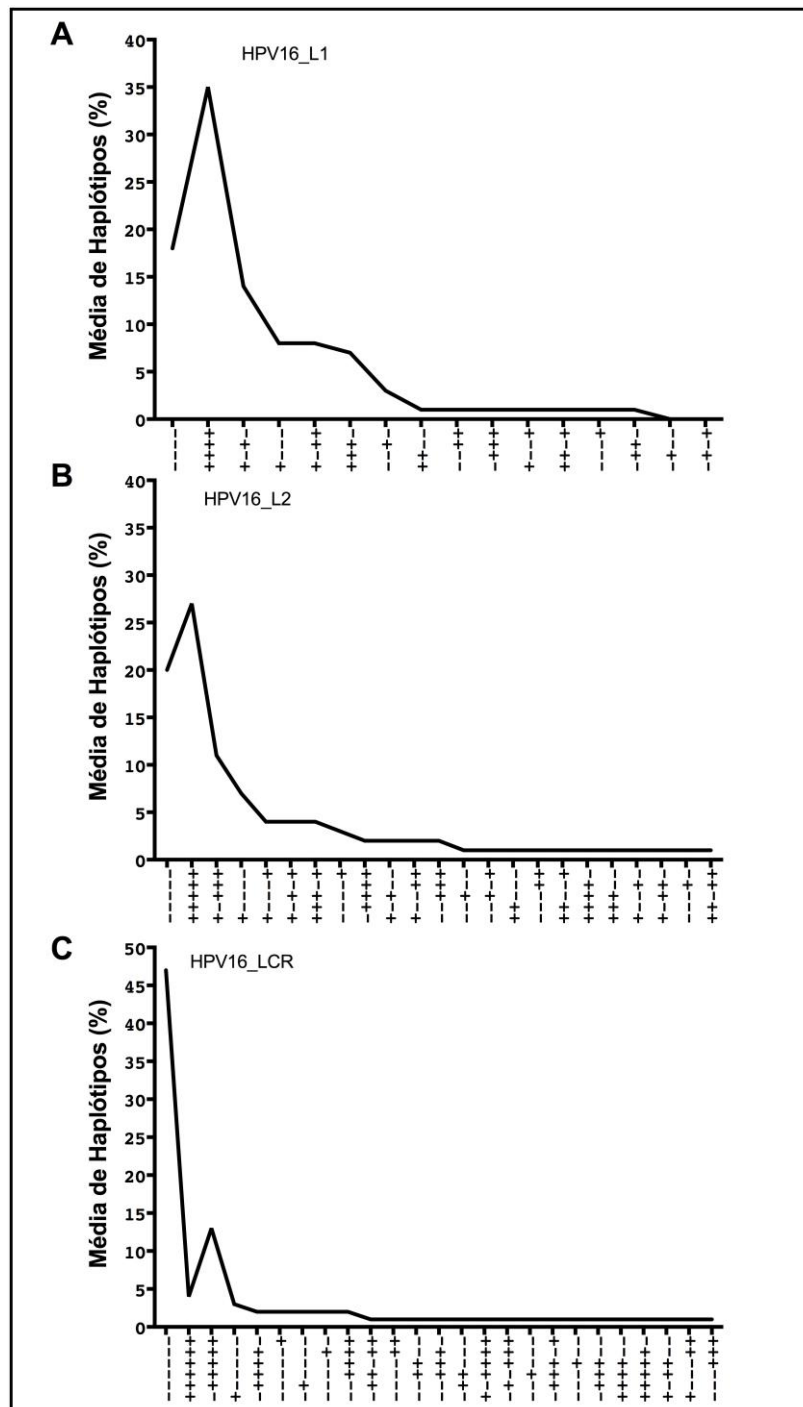


Figura 5.6. Frequência dos principais haplótipos de metilação presentes em ensaios de HPV16 por NGS. A, frequência dos haplótipos de metilação em 04 sítios CpG em L1 de HPV16 (CpG 5602, 5608, 5611 e 5617). B, frequência dos haplótipos de metilação em 05 sítios CpG em L2 de HPV16 (CpG 4240, 4249, 4261, 4270 e 4277). C, frequência dos haplótipos de metilação em 06 sítios CpG na LCR de HPV16 (CpG 7862, 31, 37, 43, 52 e 58). Estão representados haplótipos com frequência maior ou igual a 1% de frequência, exceto para o ensaio HPV16_L1. Em relação à ordem dos haplótipos, o sítio distal ao eixo X corresponde ao primeiro sítio CpG de cada fragmento analisado.

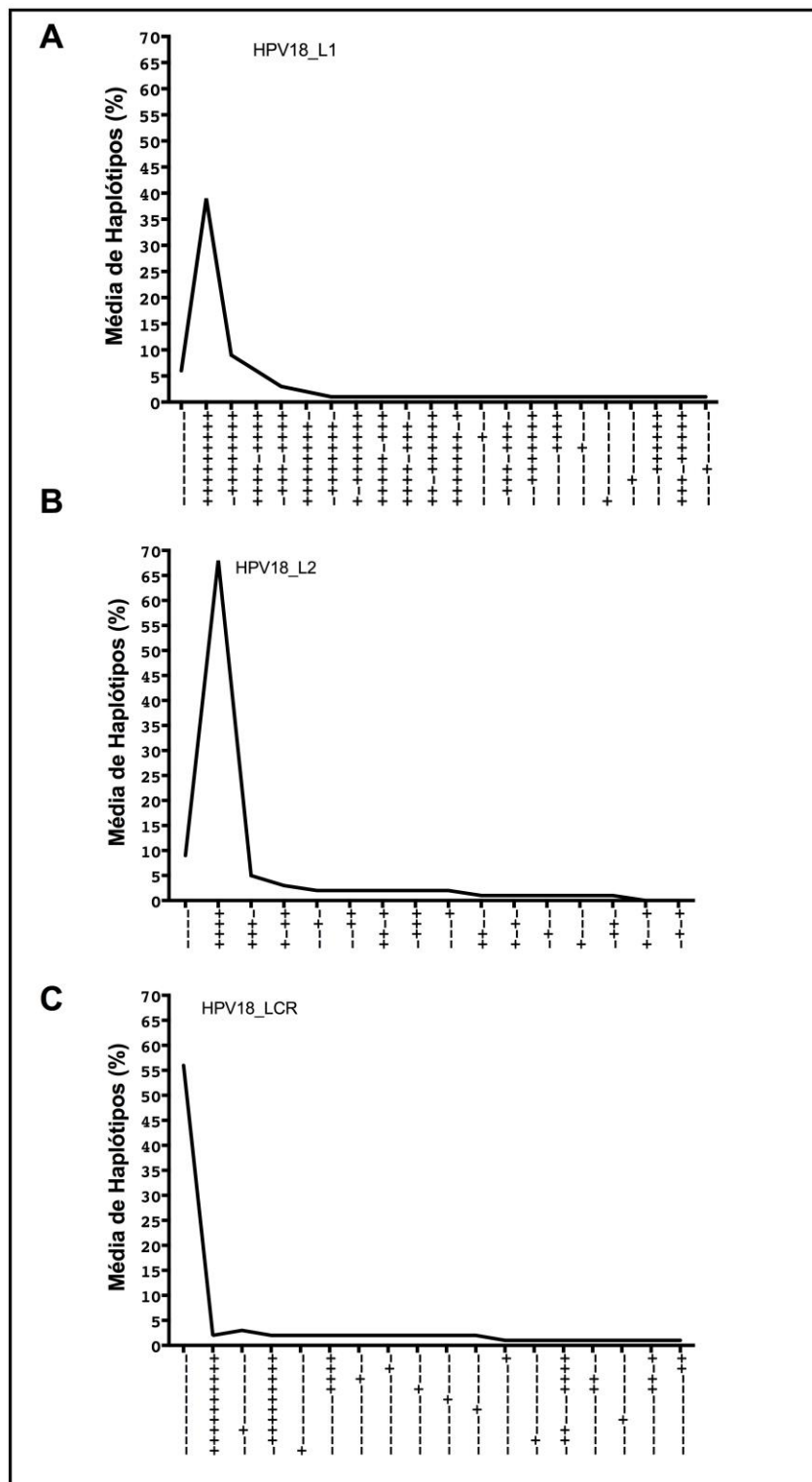


Figura 5.7. Frequência dos principais haplótipos de metilação presentes em ensaios de HPV18 por NGS. A, frequência dos haplótipos de metilação em 09 sítios CpG em L1 de HPV18 (CpG 7011, 7038, 7041, 7062, 7068, 7090, 7110, 7116 e 7122). B, frequência dos haplótipos de metilação em 04 sítios CpG em L2 de HPV18 (CpG 5331, 5333, 5367 e 5406). C, frequência dos haplótipos de metilação em 10 sítios CpG na LCR e E6 de HPV18 (7824, 44, 50, 60, 66, 109, 111, 129 e 132). Para análises de associação e caracterização do padrão de metilação em HPV18, foram incluídos apenas sítios CpG na LCR, entretanto, a

análise se estendeu à porção 5' do gene E6. Estão representados haplótipos com frequência maior ou igual a 1% de frequência, exceto para o ensaio HPV18_L2. Em relação à ordem dos haplótipos, o sítio distal ao eixo X corresponde ao primeiro sítio CpG de cada fragmento analisado.

5.2.6. ASSOCIAÇÃO DOS NÍVEIS DE METILAÇÃO EM L2, L1 E LCR DE HPV16 E HPV18 COM O TIPO DE TUMOR, A VARIANTE DE HPV E O ESTÁGIO TUMORAL

Sítios CpG analisados em L2 e L1 de HPV16 de CCE apresentaram maiores medianas de metilação do que adenocarcinomas, exceto nt 4277 (Figura 5.8 A). De forma inversa, medianas de metilação dos sítios CpG na LCR de HPV16 estiveram mais aumentadas em adenocarcinomas, seguindo o padrão observado por pirosequenciamento para esses sítios. Entretanto, utilizando NGS essa diferença foi mais acentuada, com diferença estatisticamente significativa entre tumores no sítio 43 ($p=0.047$) e 58 ($p=0.038$). Na análise de metilação por NGS foi incluído um sexto sítio CpG na posição 7862 na LCR, também associado à ligação da proteína E2. Nesse sítio, foi observada maior diferença nos níveis de metilação entre os tumores, $p=0,003$ (Figura 5.8 A).

Linhagens europeias de HPV16 apresentaram maiores níveis de metilação em relação às não-europeias em todos os sítios presentes na região L1 e L2, com diferença estatisticamente significativa para os sítios 4270 ($p=0,027$), 5608 ($p=0,040$), 5617 ($p=0,045$). Essa diferença não foi notada entre os sítios presentes na LCR de HPV16 (Figura 5.8 B). Amostras de pacientes com tumores em estágio inicial (I) apresentaram maiores medianas de metilação do que em estágios avançados (II-IV) em 13 dos 15 sítios CpG analisados (Figura 5.8 C). Entretanto, não foi observada diferença estatisticamente significativa.

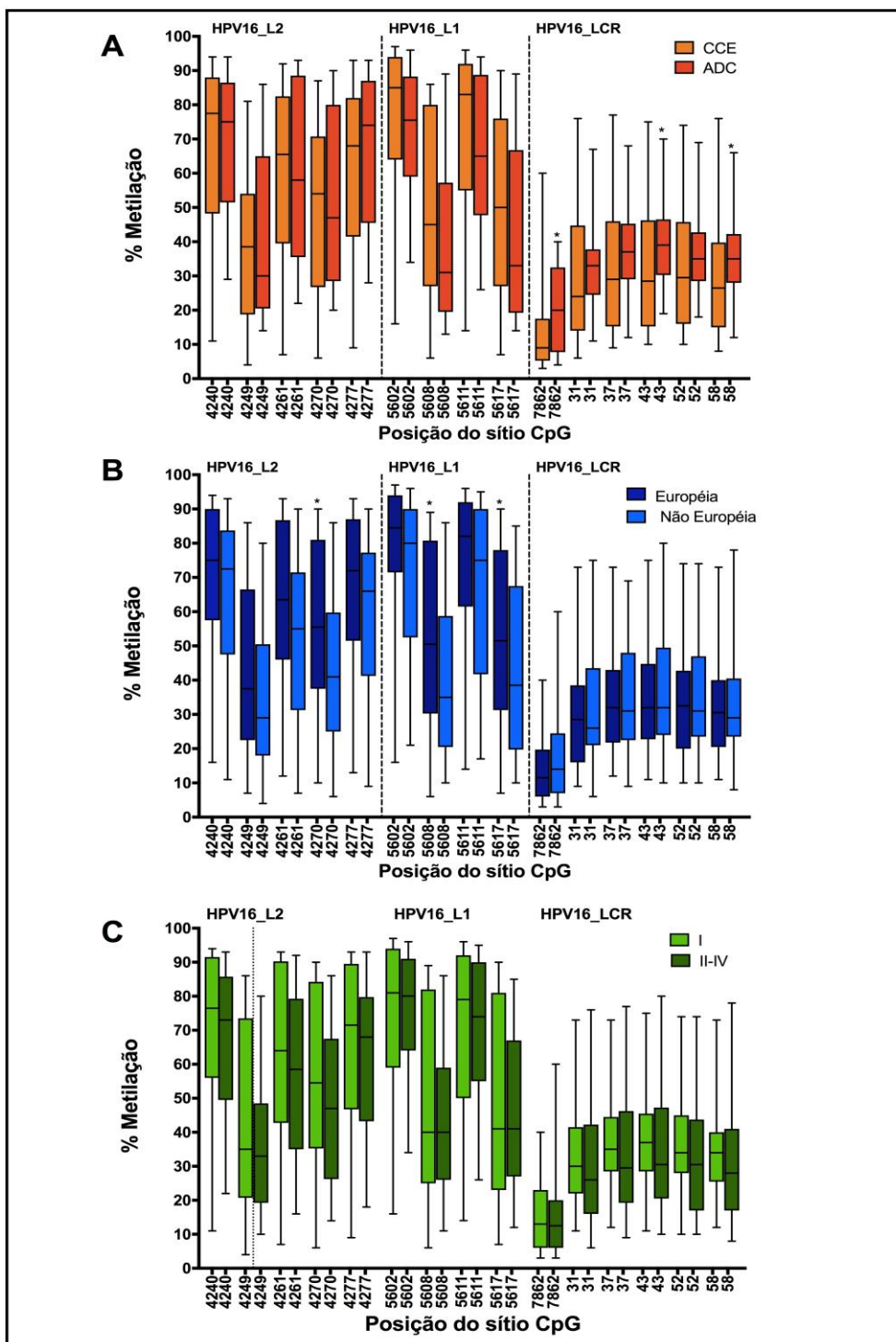


Figura 5.8. Distribuição dos níveis de metilação analisados por NGS em sítios CpG presentes em L2, L1 e LCR de HPV16 por tipo histológico, variante de HPV e estágio tumoral. A caixa representa os quartis superiores e inferiores separados pela mediana. Nas extremidades se encontram os valores máximos e mínimos. Foram observadas diferenças estatisticamente significantes, $p < 0,05$ (*), entre os níveis de metilação na LCR de HPV16 por tipo histológico (A) e em L2 e L1 por variante de HPV (B). As análises foram feitas utilizando *Man-Whitney Test*. ADC= adenocarcinomas e CCE=carcinomas de células escamosas.

Para HPV18, não foi observado um padrão nos níveis de metilação em relação ao tipo de tumor (Figura 5.9 A), com medianas de metilação muito próximas.

O mesmo pôde ser inferido entre as linhagens europeias e não-europeias (Figura 5.9 B). Já amostras de tumores em estágio precoce (I) apresentaram, para todos os sítios CpG, maiores medianas de metilação, com diferença estatisticamente significativa na posição 7824 ($p=0,035$) (Figura 5.9 C).

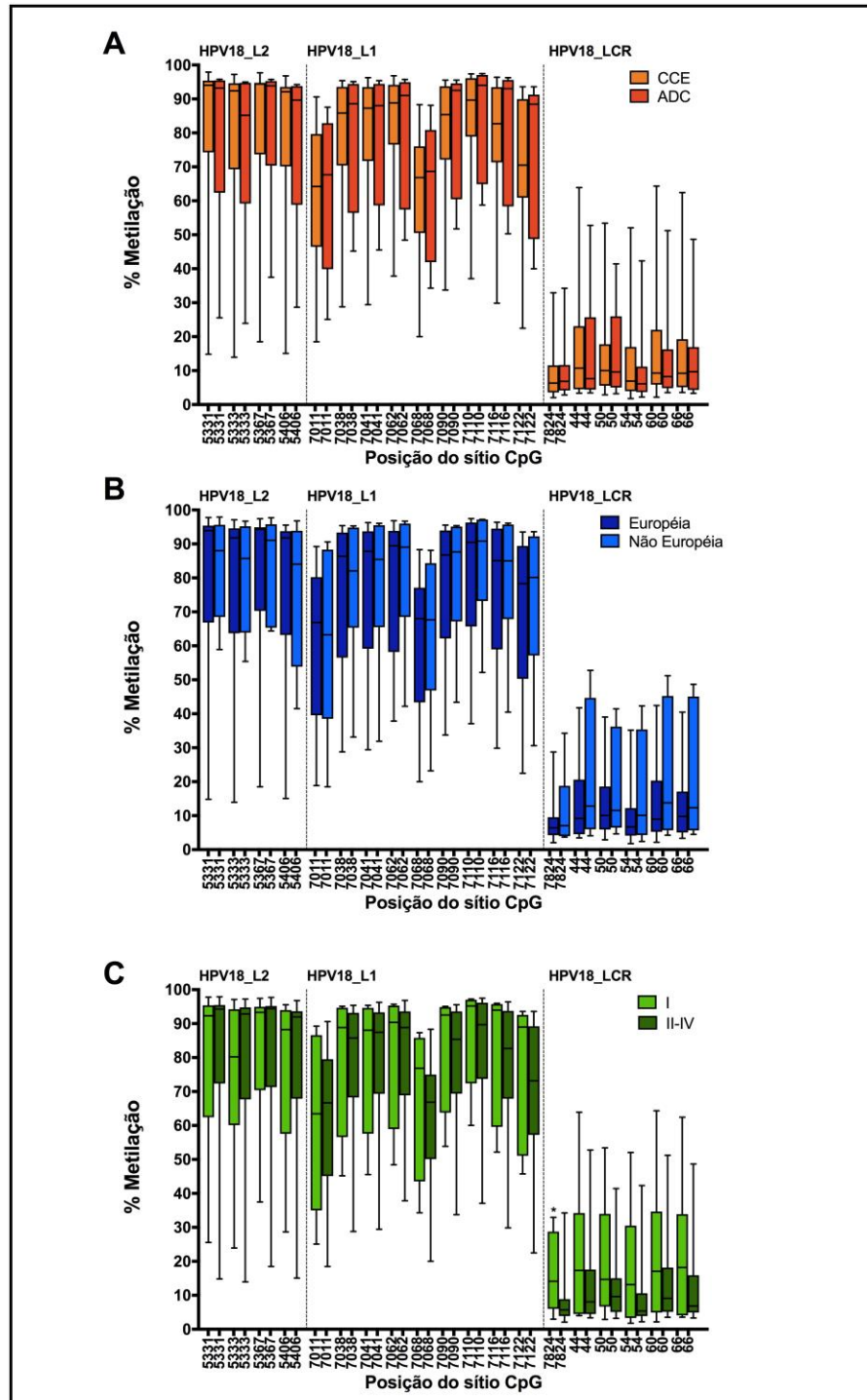


Figura 5.9. Distribuição dos níveis de metilação analisados por NGS em sítios CpG presentes em L2, L1 e LCR de HPV18 por tipo histológico, variante de HPV e estágio tumoral. A caixa representa os quartis superiores e inferiores separados pela mediana. Nas extremidades se encontram os valores máximos e mínimos. Para HPV18, apenas o nt 7824 na LCR apresentou diferença estatisticamente significativa, $p < 0,05$ (*), entre os níveis de

metilação e o estágio tumoral (C). As análises foram feitas utilizando *Man-Whitney Test*. ADC= adenocarcinomas e CCE=carcinomas de células escamosas.

5.2.7. ASSOCIAÇÃO DOS NÍVEIS DE METILAÇÃO EM L2, L1 E LCR DE HPV16 E HPV18 COM A INTEGRIDADE DE E1/E2

Amostras com ORFs E1/E2 intactas apresentaram aumento, estatisticamente significativo, dos níveis de metilação em todos os sítios CpG localizados em regiões de ligação da proteína E2 na 3'LCR de HPV16. Não foi observada diferença estatisticamente significativa para o sítio de ligação da proteína SP1 (nt 31, $p=0,089$) na 3'LCR, nem em sítios CpG presente em L2 e L1 de HPV16 (Figura 5.10 A). Em relação à HPV18, não foi observada diferença estatisticamente significativa entre os níveis de metilação e a integridade de E1 e E2. Apesar dos sítios CpG na 3'LCR de HPV18 apresentarem maiores níveis de metilação no grupo de amostras com DNA intacto, essa diferença não foi estatisticamente significativa (nt 50, $p=0,090$; nt 60, $p=0,096$) (Figura 5.10 B).

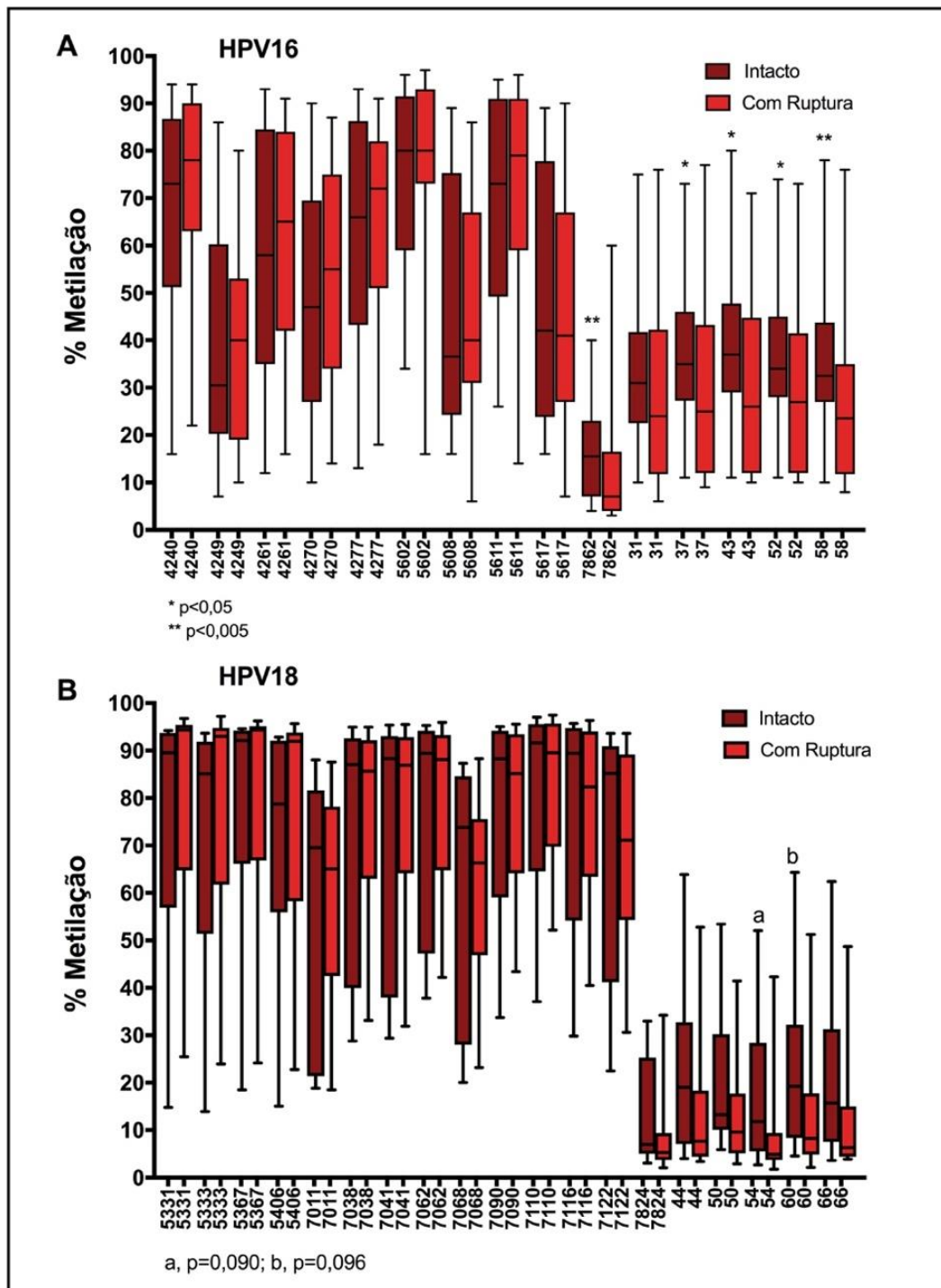


Figura 5.10. Associação entre os níveis de metilação e a integridade das ORFs E1 e E2 de HPV16 e HPV18, analisados por NGS. A caixa representa os quartis superiores e inferiores separados pela mediana. Nas extremidades se encontram os valores máximos e mínimos. Amostras com DNA de HPV16 intacto, sem ruptura de E1/E2, apresentaram maiores níveis de metilação, com diferença estatisticamente significativa ($p < 0,05$) utilizando o teste de *Mann-Whitney*, em todos os sítios CpG localizados na 3'LCR HPV16, exceto o nt 31 (A). Não foi observada diferença estatisticamente significativa quando analisados sítios em HPV18 (B).

5.2.8. ANÁLISE BIVARIADA E MULTIVARIADA UTILIZANDO MODELO DE QUASI-VEROSSIMILHANÇA

Análise bivariada foi realizada, através de modelo de Quasi-Verossimilhança, a fim de selecionar, para análise multivariada, os sítios CpG com possíveis diferenças nos níveis de metilação em cada variável de estudo. Na análise bivariada foi aplicado um critério menos estrigente, com p-valor $\leq 0,20$. Durante essa análise foram incluídos o grau de diferenciação tumoral (G1/G2 vs G3) e a carga viral como variáveis explicativas para o desfecho (nível de metilação). Segundo os critérios adotados, as variáveis que apresentaram diferença estatisticamente significativa entre os níveis de metilação em pelo menos um sítio CpG foram: estadiamento do tumor (CpG 4270), grau de diferenciação (CpG 5602, 5611 e 5617), integridade de E1/E2 (CpG 7862, 37, 43, 52 e 58), tipo histológico (CpG 4277, 5608, 5611, 5617, 7862 e 43), linhagem de HPV (CpG 4249, 4261, 4270, 4277, 5602, 5608, 5611 e 5617) e carga viral (CpG 4240, 4249, 4261, 4270, 4277, 5602, 5608, 5611, 5617 e 7862) (Tabela 5.3).

Interessante notar que a carga viral está associada de forma inversa com os níveis de metilação em sítios CpG em L2 e L1 de HPV16, porém de forma positiva em LCR, no sítio 7862 (Tabelas 5.3 e 5.4).

Na tabela 5.4 constam apenas os modelos multivariados que foram significativos ou marginalmente significativos. Na análise multivariada o grau de diferenciação do tumor (G3) e a linhagem de HPV (Européia) explicam, de forma independente, o aumento dos níveis de metilação nos sítios CpG 5602 ($p=0,05$ e $p=0,04$, respectivamente), 5608 ($p=0,06$ e $p=0,04$, respectivamente) e 5611 ($p=0,05$ e $p=0,04$, respectivamente) em L1 de HPV16. O aumento dos níveis de metilação nos sítios 5602 e 5611 também estiveram associados à carga viral independente do grau de diferenciação e da linhagem de HPV. Os níveis de metilação no sítio 7862 na LCR apresentaram associações ao tipo de tumor (aumento dos níveis de metilação em adenocarcinomas) e associação à carga viral ($p=0,002$ e $p=0,04$, respectivamente), de forma independente.

Da mesma forma que justificado para as análises multivariadas com os dados de pirosequenciamento, essa análise não foi realizada para HPV18 devido à limitação do tamanho amostral, em que 80% (31/41) das amostras apresentaram ruptura de E1/E2.

Tabela 5.3. Modelos bivariados para os níveis de metilação por sítio CpG utilizando modelo de quasi-verossimilhança.

Regiões	Sítio CpG	Idade			Estadiamento FIGO			Grau			Integridade E1/E2			Tumor			Linhagem de HPV			Carga Viral		
		<i>em anos</i>			<i>estádios II-IV (Ref: estadio I)</i>			<i>G3 (Ref: G1-G2)</i>			<i>Com ruptura (Ref: Intacto)</i>			<i>CC (Ref: ADN)</i>			<i>Não europeia (Ref: Europeia)</i>			<i>Log, por 10⁶ de células</i>		
		Coef,	Erro-padrão	p-valor	Coef,	Erro-padrão	p-valor	Coef,	Erro-padrão	p-valor	Coef,	Erro-padrão	p-valor	Coef,	Erro-padrão	p-valor	Coef,	Erro-padrão	p-valor	Coef,	Erro-padrão	p-valor
L2	4240	0,00	0,01	0,89	-0,18	0,28	0,52	-0,03	0,33	0,94	0,23	0,28	0,43	-0,13	0,27	0,63	-0,22	0,26	0,40	-0,11	0,04	0,02
	4249	0,00	0,01	0,71	-0,31	0,25	0,22	0,14	0,30	0,64	-0,03	0,27	0,92	-0,10	0,25	0,68	-0,36	0,24	0,14	-0,09	0,05	0,08
	4261	0,00	0,01	0,80	-0,24	0,28	0,40	0,23	0,33	0,49	0,17	0,28	0,55	-0,09	0,27	0,75	-0,46	0,26	0,07	-0,12	0,05	0,02
	4270	0,00	0,01	0,88	-0,34	0,25	0,19	0,09	0,30	0,76	0,17	0,26	0,51	-0,12	0,25	0,65	-0,52	0,23	0,03	-0,11	0,05	0,02
	4277	0,00	0,01	0,74	-0,19	0,27	0,48	0,05	0,32	0,89	0,17	0,28	0,54	-0,34	0,27	0,20	-0,33	0,25	0,20	-0,12	0,05	0,02
L1	5602	0,00	0,01	0,75	0,12	0,27	0,64	0,66	0,36	0,07	0,09	0,27	0,73	0,24	0,27	0,37	-0,47	0,26	0,08	-0,09	0,04	0,04
	5608	-0,01	0,01	0,49	-0,26	0,26	0,32	0,40	0,31	0,21	0,04	0,26	0,89	0,46	0,26	0,09	-0,47	0,25	0,07	-0,08	0,05	0,12
	5611	0,00	0,01	0,85	0,04	0,28	0,89	0,65	0,37	0,08	0,12	0,28	0,67	0,36	0,28	0,20	-0,45	0,27	0,09	-0,11	0,05	0,03
	5617	-0,01	0,01	0,52	-0,26	0,26	0,32	0,47	0,32	0,14	-0,11	0,26	0,68	0,34	0,26	0,20	-0,44	0,25	0,08	-0,09	0,05	0,08
LCR	7862	0,01	0,01	0,41	-0,06	0,24	0,81	-0,26	0,30	0,38	-0,46	0,25	0,08	-0,64	0,22	0,00	0,25	0,22	0,27	0,07	0,04	0,10
	31	0,00	0,01	0,56	0,02	0,22	0,94	-0,27	0,27	0,31	-0,23	0,22	0,30	-0,14	0,21	0,51	0,15	0,20	0,44	-0,02	0,03	0,49
	37	0,00	0,01	0,68	-0,09	0,20	0,66	-0,24	0,24	0,34	-0,32	0,20	0,12	-0,22	0,19	0,27	0,08	0,18	0,65	0,01	0,03	0,80
	43	0,00	0,01	0,71	-0,05	0,20	0,80	-0,15	0,24	0,54	-0,36	0,20	0,08	-0,27	0,19	0,17	0,13	0,19	0,48	0,01	0,03	0,71
	52	0,00	0,01	0,78	-0,06	0,20	0,78	-0,13	0,24	0,60	-0,33	0,20	0,11	-0,20	0,19	0,30	0,11	0,19	0,56	0,00	0,03	0,92
	58	0,00	0,01	0,77	-0,04	0,21	0,84	-0,11	0,25	0,67	-0,39	0,21	0,07	-0,22	0,20	0,26	0,10	0,19	0,59	0,00	0,03	0,96

Em vermelho, sítios CpG com diferença estatisticamente significantes ($p \leq 0,20$), seleccionados para análise multivariada.

Tabela 5.4. Modelos multivariados para os níveis de metilação por sítio CpG, utilizando modelo de Quasi-Verossimilhança.

Regiões	Sítio	Grau			Tumor			Linhagem de HPV			Carga Viral		
		<i>G3</i> (Ref: <i>G1-G2</i>)			<i>SCC</i> (Ref: <i>ADN</i>)			<i>Não europeia</i> (Ref: <i>Europeia</i>)			<i>Log10⁶ células</i>		
		Coef.	Erro-padrão	p-valor	Coef.	Erro-padrão	p-valor	Coef.	Erro-padrão	p-valor	Coef.	Erro-padrão	p-valor
L1	5602	0,72	0,36	0,05	-	-	-	-0,58	0,28	0,04	-	-	-
		0,77	0,41	0,07	-	-	-	-	-	-	-0,09	0,04	0,04
	5608	-	-	-	0,50	0,26	0,06	-0,47	0,25	0,06	-	-	-
		0,70	0,36	0,06	-	-	-	-0,60	0,29	0,04	-	-	-
	5611	-	-	-	0,58	0,32	0,07	-	-	-	-0,11	0,04	0,02
		0,94	0,45	0,04	-	-	-	-0,65	0,35	0,07	-0,09	0,05	0,06
		0,70	0,36	0,05	-	-	-	-0,60	0,28	0,04	-	-	-
LCR	7862	-	-	-	-0,79	0,24	0,002	-	-	-	0,07	0,03	0,04

Os valores de *p* em negrito indicam o p-valor marginal (*borderline*) e em vermelho indicam os considerados estatisticamente significantes ($p \leq 0,05$).

5.2.9. CORRELAÇÃO ENTRE OS NÍVEIS DE METILAÇÃO POR PIROSEQUENCIAMENTO E NGS

Das 67 amostras selecionadas pelos critérios de qualidades adotados para a análise de metilação por NGS para o ensaio HPV16_LCR, 18 (27%) foram também analisadas por pirosequenciamento. Foi observada forte correlação de Spearman entre os níveis de metilação por ambas as metodologias para a média de todos os sítios CpG ($r=0,73$, $p<0,001$) (Tabela 5.5), com o sítio CpG posição 31 apresentando melhor correlação individual ($r=0,82$, $p<0,001$). Para o ensaio HPV18_LCR, das 42 amostras selecionadas, 17 (40%) foram também analisadas por pirosequenciamento, observando uma correlação moderada entre as duas análises para todos os sítios CpG ($r=0,56$, $p<0,001$), com melhor correlação no sítio CpG 54 ($r=0,61$, $p=0,01$) (Tabela 5.5).

Tabela 5.5. Correlação entre os dois métodos de análise de metilação, pirosequenciamento e NGS, para os ensaios HPV16_LCR e HPV18_LCR.

Ensaio	Sítio CpG	Correlação de Spearman	
		r	p-valor
HPV16_LCR	Todos	0,73	<0,001
	31	0,82	<0,001
	37	0,69	0,002
	43	0,69	0,002
	52	0,69	0,002
	58	0,67	0,004
HPV18_LCR	Todos	0,56	<0,001
	44	0,55	0,022
	50	0,49	0,043
	54	0,61	0,010
	60	0,48	0,046
	66	0,57	0,015

6. DISCUSSÃO

Os níveis de metilação analisados no presente estudo foram avaliados separadamente em dados gerados por pirosequenciamento e por NGS a fim de não adicionar viés devido às diferenças entre essas metodologias. De todo modo, a correlação entre os achados pelas duas metodologias para a região LCR (que foi analisada por ambos os métodos em um número limitado de amostras) de HPV16 e HPV18, mostraram uma boa correlação de Spearman ($r=0,73$, $p<0,001$ e $r=0,56$, $p<0,001$, respectivamente). Frimer *et al.* (2015) correlacionaram os resultados das duas técnicas analisando sítios em E6, E2, L2 e L1, e observaram um índice de correlação de Pearson maiores (0,91-0,99). Os índices de correlação obtidos nessa tese podem se dever a características da região LCR analisada por ambas as metodologias, que apresentou uma intensidade de metilação menor que aquelas analisadas por Frimer *et al.* (2015). No presente estudo, as amostras utilizadas para análise de NGS foram tratadas por bissulfito de sódio utilizando o mesmo método, entretanto em momentos e locais diferentes, o que poderia implicar em uma menor concordância dos resultados.

No presente estudo, por pirosequenciamento e NGS, tumores cervicais apresentaram grande variabilidade nos níveis metilação em sítios CpG localizados na 3'LCR de HPV16 e HPV18, variando de 0% a 90% e 1% a 79%, respectivamente (Figuras 5.1 e 5.5). Valores similares foram relatados por outro estudo (2% a 84%) em 17 carcinomas de células escamosas analisados por pirosequenciamento (Jacquin *et al.*, 2013). Resultados conflitantes obtidos por outros autores podem decorrer das diferentes metodologias adotadas, e do número limitado de casos de câncer analisados, reduzindo a chance de identificar variações dos níveis de metilação influenciadas por fatores clínicos e biológicos intrínsecos da população estudada. Metilação na LCR de HPV16 já foi investigada com métodos qualitativos como análise de restrição combinada com bissulfito de sódio (COBRA, do inglês *Combined Bisulfite Restriction Analysis*) e PCR específico para metilação (MS-PCR) (Ding *et al.*, 2009; Vinokurova e Von Knebel Doeberitz, 2011; Leung *et al.*, 2015). Em outro estudo os níveis de metilação gerados por pirosequenciamento são variáveis categóricas (Cheung *et al.*, 2013), motivo pelo qual não poderiam ser comparados os dados apresentados aqui.

Alguns estudos analisaram metilação na LCR de HPV16 a sua importância no controle da expressão gênica dos oncogenes E6 e E7. (Leung *et al.*, 2015; Jacquin

et al., 2013; Chaiwongkot *et al.*, 2013). Quatro sítios de ligação da proteína viral E2 (E2BS) na LCR foram descritos, juntamente com o efeito da metilação nesses sítios, que é capaz de impedir a ligação da proteína E2. No presente estudo, os sítios CpG na 3'LCR (CpG 37, 43, 52 e 58), localizados em dois sítios de ligação de E2 (E2BS#3 e #4), apesar da grande amplitude nos níveis de metilação por amostra, apresentaram medianas de metilação similares (Figuras 5.1 e 5.5), provavelmente refletindo a funcionalidade similar. A ligação de E2 na região 3'LCR apresenta papel repressor na atividade do promotor de E6/E7, e também um papel ativador da replicação do DNA mediada por E1 (O'Connor *et al.*, 1995). Sendo assim, a metilação nesses sítios favoreceria o desenvolvimento do câncer cervical em casos com presença de altas concentrações da proteína E2. Já nos casos com perda da expressão de E2 (casos com integração do DNA viral ao genoma hospedeiro) a metilação não apresentaria relevância para o desenvolvimento do câncer (revisado por Johannsen e Lambert, 2013).

Por outro lado, o sítio CpG 31 de HPV16, um sítio de ligação do fator de transcrição Sp1, está associado ao aumento da atividade do promotor de E6/E7 (Gloss *et al.*, 1990). Pôde-se notar, no atual estudo, uma sutil redução dos níveis de metilação no sítio CpG 31 em relação aos CpGs localizados em E2BS#3 e E2BS#4. Da mesma forma, o sítio CpG 7862, analisado apenas por NGS e importante também para ligação da proteína E2 (E2BS#2), apresentou um nível menor de metilação em relação aos demais sítios na LCR. Isso provavelmente porque, diferentemente de E2BS#3 e #4, acredita-se que a atividade de E2BS#2 e Sp1 estejam associadas à ativação do promotor precoce (McBride, 2013; O'Connor *et al.*, 1995). Neste caso, ausência de metilação favoreceria o desenvolvimento do câncer cervical em casos onde o DNA viral se mantém na forma epissomal ou em tandem nos quais a proteína E2 pode estar presente.

Observou-se uma ampla variação dos níveis de metilação nos sítios CpG avaliados por NGS entre as diferentes regiões do genoma (L2, L1 e LCR) de HPV16 e HPV18 (Figura 5.6 A). Essa variação foi maior para HPV18 do que HPV16. Frimer *et al.* (2015) analisaram os níveis de metilação, utilizando a mesma metodologia por NGS em 13 amostras de câncer cervical, em sítios CpG localizados em E6, E2, L2 e L1 de HPV16 (Frimer *et al.*, 2015). Nesse estudo as medianas de metilação também variaram entre os genes, com menor mediana de metilação para E6 (1,1%) e maior mediana para L2 (34,6%). Em comparação ao presente estudo, Frimer *et al.* (2015) analisando os mesmos sítios CpG em L2 (nt 4240, 4249, 4261, 4270 e 4277)

observaram uma menor metilação comparado aos dados aqui apresentados, 35%, 5%, 17%, 6% e 13% vs 75%, 33%, 60%, 50% e 71%, respectivamente. Essa divergência entre os valores de metilação entre os estudos pode ser explicada por variação interna da técnica, pelo tamanho amostral de tumores cervicais utilizados por Frimer *et al.* (2015) (13 vs 70), ou características particulares das populações estudadas.

Além disso, foram observados, no presente estudo, níveis elevados de metilação em *L2* e *L1* de HPV16 e HPV18, atingindo medianas de metilação de até 94%. *L2* e *L1* de HPV18 apresentaram maior uniformidade nos níveis de metilação enquanto HPV16 uma maior dispersão. Os sítios com maiores medianas de metilação em HPV16 foram: 4240, 4277 e 5611 ($\geq 70\%$). Estes mesmos sítios CpG em HPV16 foram descritos (Mirabelo *et al.*, 2014), como fortes candidatos a marcadores ($AUC \geq 0,75$) para detecção de mulheres com CIN2+, sendo capazes de distinguir casos de CIN2+ de infecções transientes (Mirabello *et al.*, 2013). Assim, o achado corrobora a hipótese de possível biomarcador, já que a literatura (Bryant *et al.*, 2014; Lorincz, *et al.*, 2013) tem descrito um aumento progressivo de metilação em *L2* e *L1* associado ao aumento do grau das lesões cervicais. Não se sabe ao certo a função da metilação progressiva intragênica no corpo do gene em *L2* e *L1*. Acredita-se, entretanto, que os níveis elevados de metilação em amostras de câncer cervical estejam associados à diminuição nos níveis de proteínas do capsídeo *L2* e *L1* (Turan *et al.*, 2006). Pouco se sabe sobre o papel da metilação do DNA em regiões intragênicas, entretanto, existem evidências que esse tipo de metilação inativa promotores alternativos e impede o início espúrio de transcrição em regiões intragênicas (Maunakea *et al.*, 2010; Neri *et al.*, 2017).

Pela análise de NGS foi possível estabelecer haplótipos de metilação para as regiões analisadas e uma informação mais detalhada sobre o padrão de metilação na população de estudo. Em relação à LCR, o haplótipo sem metilação nos sítios CpG analisados foi o mais frequente, tanto para HPV16 (47%) como para HPV18 (56%). Os dados dos haplótipos também mostraram que em HPV16 o sítio CpG 7862, que apresentou a menor mediana de metilação (Figura 5.6 A), estava presente demetilado no segundo haplótipo mais frequente (com a frequência de 13%), onde apenas esse sítio estava demetilado (13%) (Figura 5.7 C). Já para *L1* e *L2*, o haplótipo com todos os sítios CpG metilados foi o mais frequente em HPV16 e HPV18. Entretanto, HPV16 apresentou mais moléculas (haplótipos) sem metilação

em todos os sítios CpG em L2 e L1 (20% e 18%, respectivamente) (Figura 5.7 A e B) quando comparado à HPV18 (9% e 6%, respectivamente) (Figura 5.8 A e B), refletindo mais uma diferença entre esses tipos virais.

Os haplótipos de metilação também podem ser utilizados para determinar associação com o desfecho clínico da doença. Segundo Frimer *et al.* (2015), os haplótipos de metilação podem ser mais informativos que os níveis de metilação em uma região específica ou em sítios CpG específicos. Nesse estudo, elevado risco relativo (OR – *odds ratio*), com o padrão nos sítios CpG 4240, 4249, 4261, 4270 e 4277 em L2 de HPV16 “- - - - +”, esteve associado com NIC3 quando comparado a pacientes que eliminaram o HPV. No presente estudo esse mesmo padrão foi o sexto mais frequente entre os haplótipos, representando 3% do total (Figura 5.7 B).

A partir dos dados de NGS foram feitas análises exploratórias de associação e um modelo estatístico com análises bivariadas e multivariadas para verificar se características clínicas e biológicas da doença poderiam influenciar nos padrões de metilação observados. A seguir esses dados são comentados:

i. Tipo Histológico

Observou-se que outros fatores, além da integração do DNA viral ao genoma da célula hospedeira, podem influenciar os níveis de metilação na *LCR*, *L2* e *L1* de HPV16 e HPV18. Sítios CpG na *LCR* de HPV16 apresentaram maiores níveis de metilação em adenocarcinomas de que carcinomas de células escamosas. Essa diferença foi estatisticamente significativa para os sítios de ligação da proteína E2 (CpG 7862, 43 e 58; Figura 5.9 A) por NGS. Nas análises bivariadas e multivariadas, através do modelo de Quasi-Verossimilhança, o sítio CpG 7862 manteve-se estatisticamente significativo, apresentando maior metilação em adenocarcinomas (Tabelas 5.3 e 5.4). Associações similares não foram observadas para a *LCR* de HPV18. Como observado em diferentes tecidos e células humanas (Ehrlich *et al.*, 1982), nossos dados indicam que o tipo histológico do tumor afeta os níveis e padrões de metilação do DNA viral, podendo este fator estar refletido na expressão dos genes virais *E6* e *E7*.

ii. Linhagens Virais

Outra associação encontrada nas análises univariadas para HPV16 foi o aumento, estatisticamente significativo, dos níveis de metilação em sítios CpG em L2

(nt 4270) e L1 (nt 5608 e 5617) e sua associação com linhagens europeias (Figura 5.9 B). Já o modelo estatístico multivariado apontou os sítios 5602, 5608 e 5611 de HPV16, como sítios em que os níveis de metilação são influenciados pelas linhagens europeias independente do grau de diferenciação, do tipo de tumor e da carga viral (Tabela 5.4). Essa associação entre linhagens e metilação não foi observada na LCR de HPV16, nem para sítios CpG analisados de HPV18 (Figuras 5.9 B e 5.10 B).

Existem dados indicando que linhagens de um mesmo tipo de HPV podem interferir biologicamente e etiologicamente no desenvolvimento do câncer cervical, no desenvolvimento de lesões precursoras e no processo invasivo das células contendo HPV (Burk *et al.*, 2003; Xi *et al.*, 2007; Schiffman *et al.*, 2010), com as linhagens não-europeias de HPV16 e HPV18 apresentando um potencial mais agressivo (Sichero *et al.*, 2007). É necessário verificar se essas características podem estar associadas as diferenças nos níveis de metilação nos sítios CpG observadas entre as linhagens de HPV16. Neste caso, a hipótese seria que a redução de metilação estaria associada à ativação de transcrição gênica, favorecendo o processo carcinogênico associado a linhagens não-europeias.

iii. Estadiamento

Em todos os sítios CpG na LCR de HPV18 amostras com tumores em estágio precoce (estágio I) apresentaram maiores níveis de metilação do que tumores avançados (estágios II-IV), entretanto, apenas o sítio CpG 7824 apresentou uma diferença significativa ($p=0,035$) (Figura 5.10 C). Esse resultado talvez reflita a importância desse sítio CpG para a progressão do câncer. O sítio CpG 7824 de HPV18 é homólogo ao sítio CpG 7862 de HPV16, ambos localizados em região de ligação da proteína E2 (E2BS#2) e está associado à ativação do promotor de E6 e E7 (Sanders e Maitland, 1994; Steger e Corbach, 1997; McBride, 2013). Uma explicação seria que a metilação nesse sítio estaria associada a um melhor prognóstico da doença, para variantes de HPV18.

iv. Idade ao Diagnóstico

Observou-se também uma correlação positiva entre os níveis de metilação na LCR de HPV16 por pirosequenciamento e a idade da paciente. No DNA humano, o envelhecimento tem sido associado com um padrão global de hipometilação e hipermetilação de sítios específicos, principalmente em ilhas CpG (Christensen *et al.*, 2009). Estudos anteriores que analisaram a metilação em sítios CpG na LCR e outras regiões genômicas de HPV e a idade das pacientes apresentaram dados

conflitantes. Leung *et al.* (2015), reportaram associação entre a idade da paciente e aumento nos níveis de metilação em E2BS#4 de HPV18 (5'LCR), mas não para HPV16 em carcinoma de células escamosas (Leung *et al.*, 2015). Mirabelo *et al.* (2013), reportaram aumento dos níveis de metilação em sítios CpG ao longo do genoma de HPV16 associado com a idade de pacientes com NIC2-3 (neoplasias intraepiteliais cervicais de alto grau) mas não nos controles (mulheres com infecções por HPV16 transientes no período <2 anos) (Mirabello *et al.*, 2013).

A correlação entre metilação e idade está provavelmente associada com o processo neoplásico e a superexpressão de DNMTs. A correlação positiva entre idade e metilação na LCR de HPV16 indica que o envelhecimento pode afetar a epigenética do tumor. Em tecidos normais, com o aumento da idade, foi observado um declínio global de metilação em sítios CpG (Zampieri *et al.*, 2015). Em câncer cervical e NIC1-3 foi observada alta expressão da proteína DNMT1 em comparação com o epitélio normal escamoso (Sawada *et al.*, 2007); e, em linhagens celulares de câncer cervical, foi observada associação entre o aumento de *E6* e *E7*, aumento da expressão de DNA metiltransferases e a metilação de genes supressores de tumorais (Au Yeung *et al.*, 2010; Li *et al.*, 2015). Dados gerados no presente estudo e os de Mirabello *et al.* (para NIC2-3 e mulheres com HPV16 transiente) indicam que essa associação pode ocorrer em lesões de alto grau (NIC2-3) e câncer, mas não em mulheres com infecção transiente por HPV16.

Curiosamente não foi observada associação da idade com os níveis de metilação na LCR de HPV16 utilizando pela análise de metilação por NGS. Esse fato pode estar associado aos critérios utilizados para a seleção das amostras em cada análise. Na análise por pirosequenciamento foram incluídas propositalmente amostras das linhagens B e C de HPV16 compreendendo 27% (19/69) das amostras. Já no grupo de amostras analisadas NGS, com seleção das linhagens se deu de forma aleatória, e amostras das linhagens B e C corresponderam a 7% (7/123) do total. Apesar da linhagem de HPV16 ter sido considerada variável de controle durante as análises multivariadas, devido limitações do número amostral, elas foram agrupadas em linhagens européias (A) e não-européias (B, C e D), e o peso individual de cada linhagem nos níveis de metilação pode ter sido influenciado.

v. Integridade de E1 e E2

Observou-se, no presente estudo, forte associação dos níveis de metilação por pirosequenciamento com a integridade das ORFs *E1/E2* de HPV16 e HPV18.

Amostras com as ORFs *E1* e *E2* intactos, exibiram maior percentual de metilação quando comparadas àquelas com ruptura em uma dessas ORFs. Essa associação foi observada utilizando análise exploratória inicial para todos os sítios CpG presentes em locais de ligação da proteína *E2* de HPV16, tanto para as análises de metilação por pirosequenciamento quanto por NGS (Figuras 5.2 D e 5.10 A). Sítios CpG em HPV18 também apresentaram essa tendência, entretanto apenas dois sítios, por pirosequenciamento, apresentaram significância estatística, provavelmente devido ao número reduzido de amostras com *E1/E2* intactos. A análise multivariada dos dados de pirosequenciamento, ajustada por variante de HPV e tipo de tumor, reforçou esse achado para HPV16, mostrando a integração como fator que influencia de forma independente nos níveis de metilação (Tabela 5.2).

Esses resultados corroboram os dados na literatura em que tumores cervicais com a presença da proteína *E2*, e conseqüentemente *E1* e *E2* intactos, a metilação teria papel fundamental bloqueando a atividade regulatória repressora de *E2* na região promotora dos oncogenes *E6* e *E7*. Chaiwongkot *et al.* (2013) mostraram que altos níveis de metilação em *E2BS#3* e *#4* em carcinomas de células escamosas com HPV16 contribuem com a superexpressão dos transcritos de *E6* e *E7* por bloquearem a ligação de *E2*.

Diversos estudos reforçam esses achados tanto em tumores cervicais quanto de orofaringe (Das Ghosh *et al.*, 2012; Huang *et al.*, 2014), entretanto, utilizando diferentes técnicas para análise de integração do DNA de HPV16 ao genoma da célula hospedeira. Parte-se do pressuposto que a integração ocorre principalmente com ruptura em *E2*. Dentre os métodos para análise de integração mais utilizados estão a PCR em tempo real, avaliando relação entre *E2:E6* ou *E2:E7*, e APOT (do inglês *Amplification of Papillomavirus Oncogene Transcripts*) (Mazumder Indra *et al.*, 2011; Chaiwongkot *et al.*, 2013; Cheung *et al.*, 2013; Lillsunde Larsson *et al.*, 2014; Marongiu *et al.*, 2014). Em todas essas abordagens existem limitações relacionadas às metodologias para detectar integração do DNA viral ao genoma da célula hospedeira e perda da atividade de *E2*.

No presente estudo, ruptura em *E1/E2*, mensurada por PCR utilizando iniciadores franqueando essas ORFs, indica, com confiança, integração do genoma viral ao genoma do hospedeiro. Sabe-se que integrações podem ocorrer com maior frequência com a ruptura em *E1/E2* e raramente são observadas rupturas na *LCR*,

E6 e *E7* (Xu *et al.*, 2013; Liu *et al.*, 2016). Por outro lado, *E1/E2* intactos podem indicar presença de DNA em estado epissomal, integrado em *tandem*, ou presença das duas formas (integrada e epissomal), em todos os casos tornando possível a expressão de *E2*.

Em relação ao HPV18, Turan e colaboradores identificaram baixos níveis de metilação na linhagem celular HeLa, que são células imortalizadas com HPV18 e com DNA viral integrado (Turan *et al.*, 2006). Diferentemente de HPV16, HPV18 apresentou quase que a totalidade das amostras com ruptura em *E1/E2* (~80%), o que, pelo mesmo racional, justificaria o menor nível de metilação na *LCR*.

Alguns estudos têm associado os níveis de metilação na *LCR* de HPV16, especificamente em sítios CpG de ligação da proteína *E2*, e o estado físico do DNA em tumores cervicais, com o prognóstico da doença. Mazumder *et al.* (2011), observaram em tumores cervicais melhor prognóstico em pacientes com HPV16 epissomal e com metilação na *LCR* em comparação às outras formas (Mazumder Indra *et al.*, 2011). Poulami Das *et al.* (2015), mostraram que tumores cervicais de pacientes com DNA de HPV16 em estado epissomal possuíam melhor sobrevida livre de doenças quando comparado com o vírus em seu estado integrado, incluindo uma menor recorrência, entretanto, o nível de metilação não foi avaliado (Das *et al.*, 2015).

Já Ding *et al.* (2009) observaram em 15 carcinomas de células escamosas, utilizando PCR específico para análise de metilação, uma maior sobrevida em 5 anos de pacientes sem metilação na *LCR* de HPV16 do que aqueles com metilação. Essa associação foi observada também em carcinomas de células escamosas de orofaringe, em que Barbieri *et al.* (2014) observaram tendência de melhor prognóstico para pacientes com DNA de HPV16 em estado epissomal e pior prognóstico para aqueles com alta frequência de metilação em 3'*LCR* (Barbieri *et al.*, 2014). Leung *et al.* (2015) analisaram os níveis de metilação por pirosequenciamento em 52 tumores cervicais e observaram que amostras com maiores níveis de metilação em *E2BS#3* e *#4* (3'*LCR*) em comparação com *E2BS#1* (5'*LCR*) apresentaram significativamente melhor sobrevida global e sobrevida livre de doença quando comparados aos casos com alta metilação em *E2BS#1* em relação a *E2BS#3* e *E2BS#4* (Leung *et al.*, 2015), reforçando a hipótese em que sítios de ligação da proteína *E2* na *LCR* podem apresentar funções diferentes.

vi. Carga viral

A carga viral de HPV16 foi outra variável de estudo utilizada nas análises de associação com o modelo estatístico. Inicialmente, a carga viral de HPV16 foi utilizada como um dos critérios de qualidade para seleção de amostras analisadas por NGS a fim de evitar erros devido ao fenômeno de esgotamento de DNA viral, já que as amostras duplicadas com resultados divergentes foram aquelas com baixa carga viral (<0,0016 cópias/célula). O esgotamento de DNA é um fenômeno que pode ocorrer quando existem baixas concentrações de determinada molécula de DNA. Isso poderia explicar ausência de produtos de PCR visíveis em géis de agarose e as divergências nos níveis de metilação entre as duplicatas de amostras específicas, já que alíquotas com diferentes moléculas de DNA apresentariam níveis de metilação diferentes.

Alguns estudos têm mostrado aumento gradual da carga viral de HPV16 conforme a presença de lesões de grau mais elevado e câncer cervical, associando também a integração do DNA viral com o maior risco em desenvolvimento de NIC3 (Shukla *et al.*, 2014; Xi, *et al.*, 2011). Em muitas biópsias cervicais ou em linhagens celulares, são encontradas múltiplas cópias de DNA de HPV apresentando integração em concatêmeros (na orientação cabeça cauda) (Baker *et al.*, 1987). Nesses casos, sugere-se que apenas uma cópia do genoma viral está ativa e as demais silenciadas através da metilação do DNA (Van *et al.*, 2004), já que o silenciamento favoreceria a sobrevivência das células, modulando o nível de expressão dos oncogenes virais.

No presente estudo, observou-se uma ampla variação na carga viral de HPV16 em amostras de câncer cervical invasivo (informação disponível nos Anexo 7-9). Esses dados foram utilizados como possíveis variáveis explicativas que afetariam os níveis de metilação nas análises bivariadas e multivariadas utilizando o modelo de Quasi-Verossimilhança. Foi observada uma associação inversa entre a carga viral e os níveis de metilação em L1 de HPV16, independente do grau de diferenciação, do tipo de tumor e da linhagem de HPV. Já na LCR, foi observada uma associação positiva para sítio 7862 de HPV16, independente do tipo de tumor (Tabela 5.4). Chaiwongkot *et al.* (2013) também observaram uma correlação positiva entre o número de cópias do genoma de HPV16, entretanto analisando apenas o número de cópias integradas em lesões cervicais, e o nível de metilação da LCR. Nesse estudo, mostrou-se que todas as amostras de lesões cervicais com DNA integrado (N=12), determinado pela técnica APOT, possuíam apenas uma cópia com a transcrição ativa. Na presente tese, a carga viral e metilação foram avaliadas no

grupo total de amostras de câncer cervical invasivo, podendo apresentar DNA de HPV no seu estado epissomal ou integrado em tandem, o que poderia justificar a falta de associação com os demais sítios CpG na LCR.

Na presente tese foi observado que os níveis de metilação no DNA de HPV16 e HPV18 presentes em tumores cervicais invasivos podem sofrer influência de múltiplos fatores clínicos e biológicos da doença. Os dados reforçam a hipótese de que o DNA de HPV16 em sua forma epissomal ou em tandem apresentam maiores níveis de metilação na 3'LCR, entretanto isso não afeta os níveis de metilação em L2 e L1 tanto para HPV16 como para HPV18. Além disso, as discussões apresentadas sugerem grande heterogeneidade em HPV16, em que tumores com DNA de HPV16 em forma epissomal, com *E2* intacta, provavelmente apresentam comportamento biológico diferente daqueles com formas integradas, podendo envolver um processo carcinogênico variado. Mostrou-se também que a idade pode apresentar um papel nos níveis de metilação na 3'LCR de HPV16 em amostras de câncer quando os níveis de metilação foram analisados por pirosequenciamento, entretanto, o método de análise de metilação ou os critérios de inclusão precisam ser melhor explorados para entender a falta de concordância entre as duas análises. Além disso, fatores como linhagem de HPV16, tipo de tumor, grau de diferenciação e carga viral, podem afetar sítios CpG específicos em L2 e L1 de HPV16, sugerindo possível potencial regulatório em tais regiões. Os achados sugerem que a metilação no DNA de HPV, principalmente HPV16, como um processo complexo, em que múltiplos fatores afetam os níveis de metilação desse vírus e estudos prospectivos, com maior número amostras, devem ser desenvolvidos a fim de explorar a importância dos achados para o desfecho clínico.

7. CONCLUSÕES

- Os níveis de metilação no DNA de HPV16 e HPV18 apresentam padrões de metilação similares, com elevados níveis de metilação em *L2* e *L1* em comparação à *LCR*, e com maior heterogeneidade nos níveis de metilação para HPV16.
- Ambas as técnicas de análise de metilação, pirosequenciamento e NGS, apresentaram boa concordância entre os resultados.
- ORFs *E1/E2* intactas, sugerem presença da forma epissomal de HPV, e foi associada ao aumento nos níveis de metilação na *LCR* de HPV16 e HPV18 por pirosequenciamento. Essa associação só se manteve para HPV16 na análise por NGS e não foi significativa para HPV18.
- Mulheres mais velhas apresentaram maiores níveis de metilação em 3'LCR de HPV16 pela análise de pirosequenciamento, essa correlação foi observada independente da integridade de *E1/E2*. Essa associação não foi encontrada pela análise por NGS em um outro grupo amostral.
- A análise de haplótipos de metilação indica a presença de um haplótipo predominante (variando de 28% a 68%) para as regiões analisadas de HPV16 e HPV18. Em *L2* e *L1* de HPV16 e HPV18 o haplótipo de metilação mais frequente nas amostras foi o que apresentou todos os sítios metilados. Entretanto, com maior frequência desse haplótipo para HPV18. Em *LCR* de HPV16 e HPV18 predominou o haplótipo sem metilação em todos os sítios CpG.
- A análise multivariada dos dados de NGS indica que vários fatores biológicos (linhagem de HPV, o tipo de tumor, o grau de diferenciação e carga viral) podem afetar sítios CpG específicos em *L2* e *L1* de HPV16.
- Estudos com um número amostral maior, a fim de realizar análises multivariadas com maior poder estatístico, devem ser desenvolvidos.

8. REFERÊNCIAS

ABBATE, E. A.; BERGER, J. M.; BOTCHAN, M. R. **The X-ray structure of the papillomavirus helicase in complex with its molecular matchmaker E2.** v. 18, n. 16, *Genes Dev*, Aug 15 2004, p. 1981-96.

ALFARO, A. *et al.* Different association of human papillomavirus 16 variants with early and late presentation of cervical cancer. v. 11, n. 12, *Plos One*, Dec 30 2016, e0169315.

ANTEQUERA, F. **Structure, function and evolution of CpG island promoters.** v. 60, n. 8, *Cell Mol Life Sci*, Aug 2003, p. 1647-58.

APGAR, B. S.; ZOSCHNICK, L.; WRIGHT, T. C., JR. **The 2001 Bethesda System terminology.** v. 68, n. 10, *Am Fam Physician*, Nov 15 2003, p. 1992-8.

AU YEUNG, C. L. *et al.* **HPV-16 E6 upregulation of DNMT1 through repression of tumor suppressor p53.** v. 24, n. 6, *Oncol Rep*, Dec 2010, p. 1599-604.

BACKES, D. M. *et al.* **Systematic review of human papillomavirus prevalence in invasive penile cancer.** v. 20, n. 4, *Cancer Causes Control*, May 2009, p. 449-57.

BADAL, S. *et al.* **The human papillomavirus-18 genome is efficiently targeted by cellular DNA methylation.** v. 324, n. 2, *Virology*, Jul 1 2004, p. 483-92.

BADAL, V. *et al.* **CpG methylation of human papillomavirus type 16 DNA in cervical cancer cell lines and in clinical specimens: genomic hypomethylation correlates with carcinogenic progression.** v. 77, n. 11, *J Virol*, Jun 2003, p. 6227-34.

BARBIERI, D. *et al.* **Detection of HPV16 African variants and quantitative analysis of viral DNA methylation in oropharyngeal squamous cell carcinomas.** v. 60, n. 3, *J Clin Virol*, Jul 2014, p. 243-9.

BAKER, C.C. *et al.*, **Structural and transcriptional analysis of human papillomavirus type 16 sequences in cervical carcinoma cell lines.** V. 61, J Virol, 1987, p. 962–71.

BAYLIN, S. B. **DNA methylation and gene silencing in cancer.** v. 2 Suppl 1, Nat Clin Pract Oncol, Dec 2005, p. S4-11.

BAYO, S. *et al.* **Risk factors of invasive cervical cancer in Mali.** v. 31, n. 1, Int J Epidemiol, Feb 2002, p. 202-9.

BERNARD, H. U. **Regulatory elements in the viral genome.** v. 445, n. 1-2, Virology, Oct 2013, p. 197-204.

BERUMEN, J. *et al.* **Genome amplification of human papillomavirus types 16 and 18 in cervical carcinomas is related to the retention of E1/E2 genes.** v. 56, n. 5, Int J Cancer, Mar 1 1994, p. 640-5.

BIRD, A. **DNA methylation patterns and epigenetic memory.** v. 16, n. 1, Genes Dev, Jan 1 2002, p. 6-21.

BIRD, A. P. **CpG-rich islands and the function of DNA methylation.** v. 321, n. 6067, Nature, May 15-21 1986, p. 209-13.

BRYANT, D. *et al.* **Quantitative measurement of Human Papillomavirus type 16 L1/L2 DNA methylation correlates with cervical disease grade.** v. 59, n. 1, J Clin Virol, Jan 2014, p. 24-9.

BURK, R. D. *et al.* **Distribution of human papillomavirus types 16 and 18 variants in squamous cell carcinomas and adenocarcinomas of the cervix.** v. 63, n. 21, Cancer Res, Nov 01 2003, p. 7215-20.

CHAIWONGKOT, A. *et al.* **Differential methylation of E2 binding sites in episomal and integrated HPV 16 genomes in preinvasive and invasive cervical lesions.** v. 132, n. 9, Int J Cancer, May 1 2013, p. 2087-94.

CHANDLER, L. A.; JONES, P. A. **Hypomethylation of DNA in the regulation of gene expression.** v. 5, Dev Biol (N Y 1985), 1988, p. 335-49.

CHEUNG, J. L. *et al.* **Virological characteristics of cervical cancers carrying pure episomal form of HPV16 genome.** v. 131, n. 2, Gynecol Oncol, Nov 2013, p. 374-9.

CHRISTENSEN, B. C. *et al.* **Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context.** v. 5, n. 8, PLoS Genet, Aug 2009, p. e1000602.

CICCARONE, F. *et al.* **Age-dependent expression of DNMT1 and DNMT3B in PBMCs from a large European population enrolled in the MARK-AGE study.** v. 15, n. 4, Aging Cell, Aug 2016, p. 755-65.

CLIFFORD, G. *et al.* **Chapter 3: HPV type-distribution in women with and without cervical neoplastic diseases.** v. 24 Suppl 3, Vaccine, Aug 31 2006, p. S3/26-34.

CLIFFORD, G. M. *et al.* **Human papillomavirus genotype distribution in low-grade cervical lesions: comparison by geographic region and with cervical cancer.** v. 14, n. 5, Cancer Epidemiol Biomarkers Prev, May 2005, p. 1157-64.

CLIFFORD, G. M. *et al.* **Comparison of HPV type distribution in high-grade cervical lesions and cervical cancer: a meta-analysis.** v. 89, n. 1, Br J Cancer, Jul 07 2003, p. 101-5.

COLLINS, S. I. *et al.* **Disruption of the E2 gene is a common and early event in the natural history of cervical human papillomavirus infection: a longitudinal cohort study.** v. 69, n. 9, Cancer Res, May 1 2009, p. 3828-32.

CRICCA, M. *et al.* **Disruption of HPV 16 E1 and E2 genes in precancerous cervical lesions.** v. 158, n. 1-2, J Virol Methods, Jun 2009, p. 180-3.

CULLEN, A. P. *et al.* **Analysis of the physical state of different human papillomavirus DNAs in intraepithelial and invasive cervical neoplasm.** v. 65, n. 2, J Virol, Feb 1991, p. 606-12.

DALING, J. R. *et al.* **The relationship of human papillomavirus-related cervical tumors to cigarette smoking, oral contraceptive use, and prior herpes simplex virus type 2 infection.** v. 5, n. 7, Cancer Epidemiol Biomarkers Prev, Jul 1996, p. 541-8.

DAS GHOSH, D. *et al.* **Some novel insights on HPV16 related cervical cancer pathogenesis based on analyses of LCR methylation, viral load, E7 and E2/E4 expressions.** v. 7, n. 9, PLoS One, 2012, p. e44678.

DAS, P. *et al.* **Human papillomavirus (HPV) genome status & cervical cancer outcome--A retrospective study.** v. 142, n. 5, Indian J Med Res, Nov 2015, p. 525-32.

DESAINTES, C.; DEMERET, C. **Control of papillomavirus DNA replication and transcription.** v. 7, n. 6, Semin Cancer Biol, Dec 1996, p. 339-47.

DING, D. C. *et al.* **Methylation of the long control region of HPV16 is related to the severity of cervical neoplasia.** v. 147, n. 2, Eur J Obstet Gynecol Reprod Biol, Dec 2009, p. 215-20.

DOORBAR, J. **Molecular biology of human papillomavirus infection and cervical cancer.** v. 110, n. 5, Clin Sci (Lond), May 2006, p. 525-41.

DUENSING, S.; MUNGER, K. **The human papillomavirus type 16 E6 and E7 oncoproteins independently induce numerical and structural chromosome instability.** v. 62, n. 23, Cancer Res, Dec 1 2002, p. 7075-82.

DURST, M. *et al.* **A papillomavirus DNA from a cervical carcinoma and its prevalence in cancer biopsy samples from different geographic regions.** v. 80, n. 12, Proc Natl Acad Sci U S A, Jun 1983, p. 3812-5.

EHRlich, M. *et al.* **Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells.** v. 10, n. 8, *Nucleic Acids Res*, Apr 24 1982, p. 2709-21.

FAVRE, M. **Structural polypeptides of rabbit, bovine, and human papillomaviruses.** v. 15, n. 5, *J Virol*, May 1975, p. 1239-47.

FAVRE, M. *et al.* **Chromatin-like structures obtained after alkaline disruption of bovine and human papillomaviruses.** v. 21, n. 3, *J Virol*, Mar 1977, p. 1205-9.

FERNANDEZ, A. F. *et al.* **The dynamic DNA methylomes of double-stranded DNA viruses associated with human cancer.** v. 19, n. 3, *Genome Res*, Mar 2009, p. 438-51.

FORMAN, D. *et al.* **Global burden of human papillomavirus and related diseases.** v. 30 Suppl 5, *Vaccine*, Nov 20 2012, p. F12-23.

FRIMER, M. *et al.* **HPV16 CpG methyl-haplotypes are associated with cervix precancer and cancer in the Guanacaste natural history study.** v. 138, n. 1, *Gynecol Oncol*, Jul 2015, p. 94-100.

GLOSS, B.; BERNARD, HU. **The E6/E7 human papillomavirus type 16 is activated in absence of E2 proteins by a sequence-aberrant SP1-distal element.** V. 64, n. 11, *Journal of Virology*, Nov 1990, p. 5577-5584.

GOODMAN, M. T. *et al.* **Prevalence, acquisition, and clearance of cervical human papillomavirus infection among women with normal cytology: Hawaii Human Papillomavirus Cohort Study.** v. 68, n. 21, *Cancer Res*, Nov 01 2008, p. 8813-24.

GRONBAEK, K.; HOTHER, C.; JONES, P. A. **Epigenetic changes in cancer.** v. 115, n. 10, *APMIS*, Oct 2007, p. 1039-59.

HERMAN, J. G.; BAYLIN, S. B. **Gene silencing in cancer in association with promoter hypermethylation.** v. 349, n. 21, *N Engl J Med*, Nov 20 2003, p. 2042-54.

HERRERA, L. A. *et al.* **The epigenetic origin of aneuploidy.** v. 9, n. 1, *Curr Genomics*, Mar 2008, p. 43-50.

HOFFMANN, M. J.; SCHULZ, W. A. **Causes and consequences of DNA hypomethylation in human cancer.** v. 83, n. 3, *Biochem Cell Biol*, Jun 2005, p. 296-321.

HONG, D. *et al.* **Methylation status of the long control region of HPV 16 in clinical cervical specimens.** v. 1, n. 4, *Mol Med Rep*, Jul-Aug 2008, p. 555-60.

HSU, C. H. *et al.* **The HPV E6 oncoprotein targets histone methyltransferases for modulating specific gene transcription.** v. 31, n. 18, *Oncogene*, May 3 2012, p. 2335-49.

HUANG, Y. T. *et al.* **Epigenetic patterns in successful weight loss maintainers: a pilot study.** *Int J Obes (Lond)*, Dec 18 2014.

IARC. **Working Group on the Evaluation of Carcinogenic Risks to Human. Biological agents. Volume 100B. A review of human carcinogens.** . IARC Monogr Eval Carcinog Risks Hum., p.1-441. 2012

INCA/MS. **Estimativa 2016: incidência de câncer no Brasil / Instituto Nacional de Câncer – Rio de Janeiro: INCA.** 2015.

JACQUIN, E. *et al.* **Methylation of human papillomavirus Type 16 CpG sites at E2-binding site 1 (E2BS1), E2BS2, and the Sp1-binding site in cervical cancer samples as determined by high-resolution melting analysis-PCR.** v. 51, n. 10, *J Clin Microbiol*, Oct 2013, p. 3207-15.

JEMAL, A. *et al.* **Global cancer statistics.** v. 61, n. 2, *CA Cancer J Clin*, Mar-Apr 2011, p. 69-90.

JOHANNSEN, E.; LAMBERT, P. F. **Epigenetics of human papillomaviruses.** v. 445, n. 1-2, *Virology*, Oct 2013, p. 205-12.

JOHANSSON, C.; SCHWARTZ, S. **Regulation of human papillomavirus gene expression by splicing and polyadenylation.** v. 11, n. 4, Nat Rev Microbiol, Apr 2013, p. 239-51.

KALANTARI, M. *et al.* **Conserved methylation patterns of human papillomavirus type 16 DNA in asymptomatic infection and cervical neoplasia.** v. 78, n. 23, J Virol, Dec 2004, p. 12762-72.

KALANTARI, M. *et al.* **Effects of cellular differentiation, chromosomal integration and 5-aza-2'-deoxycytidine treatment on human papillomavirus-16 DNA methylation in cultured cell lines.** v. 374, n. 2, Virology, May 10 2008, p. 292-303.

KALANTARI, M. *et al.* **Methylation of human papillomavirus 16, 18, 31, and 45 L2 and L1 genes and the cellular DAPK gene: Considerations for use as biomarkers of the progression of cervical neoplasia.** v. 448, Virology, Jan 05 2014, p. 314-21.

KIM, D. *et al.* **Psychological factors and DNA methylation of genes related to immune/inflammatory system markers: the VA Normative Aging Study.** v. 6, n. 1, BMJ Open, Jan 05 2016, p. e009790.

KIM, K. *et al.* **Methylation patterns of papillomavirus DNA, its influence on E2 function, and implications in viral infection.** v. 77, n. 23, J Virol, Dec 2003, p. 12450-9.

LEONHARDT, H.; CARDOSO, M. C. **DNA methylation, nuclear structure, gene expression and cancer.** v. Suppl 35, J Cell Biochem Suppl, 2000, p. 78-83.

LEUNG, T. W. *et al.* **HPV 16 E2 binding sites 1 and 2 become more methylated than E2 binding site 4 during cervical carcinogenesis.** v. 87, n. 6, J Med Virol, Jun 2015, p. 1022-33.

LI, L. *et al.* **E6 and E7 gene silencing results in decreased methylation of tumor suppressor genes and induces phenotype transformation of human cervical carcinoma cell lines.** v. 6, n. 27, *Oncotarget*, Sep 15 2015, p. 23930-43.

LILLSUNDE LARSSON, G. *et al.* **Viral load, integration and methylation of E2BS3 and 4 in human papilloma virus (HPV) 16-positive vaginal and vulvar carcinomas.** v. 9, n. 11, *PLoS One*, 2014, p. e112839.

LIM, U.; SONG, M. A. **Dietary and lifestyle factors of DNA methylation.** v. 863, *Methods Mol Biol*, 2012, p. 359-76.

LIU, L. *et al.* **Aging, cancer and nutrition: the DNA methylation connection.** v. 124, n. 10-12, *Mech Ageing Dev*, Dec 2003, p. 989-98.

LIU, Y. *et al.* **Comprehensive mapping of the human papillomavirus (HPV) DNA integration sites in cervical carcinomas by HPV capture technology.** v. 7, n. 5, *Oncotarget*, Feb 2 2016, p. 5852-64.

LORINCZ, A. T. *et al.* **HPV16 L1 and L2 DNA methylation predicts high-grade cervical intraepithelial neoplasia in women with mildly abnormal cervical cytology.** v. 133, n. 3, *Int J Cancer*, Aug 1 2013, p. 637-44.

MARONGIU, L. *et al.* **Human Papillomavirus 16, 18, 31 and 45 viral load, integration and methylation status stratified by cervical disease stage.** v. 14, *BMC Cancer*, May 30 2014, p. 384.

MAZUMDER INDRA, D. *et al.* **Genetic and epigenetic changes of HPV16 in cervical cancer differentially regulate E6/E7 expression and associate with disease progression.** v. 123, n. 3, *Gynecol Oncol*, Dec 2011, p. 597-604.

MAUNAKEA, A. K. *et al.* **Conserved role of intragenic DNA methylation in regulating alternative promoters.** v. 466, *Nature*, 2010, p. 253–257.

MCBRIDE, A. A. **The papillomavirus E2 proteins.** v. 445, n. 1-2, *Virology*, Oct 2013, p. 57-79.

MCKAIG, R. G.; BARIC, R. S.; OLSHAN, A. F. **Human papillomavirus and head and neck cancer: epidemiology and molecular biology.** v. 20, n. 3, Head Neck, May 1998, p. 250-65.

MCLAUGHLIN-DRUBIN, M. E.; CRUM, C. P.; MUNGER, K. **Human papillomavirus E7 oncoprotein induces KDM6A and KDM6B histone demethylase expression and causes epigenetic reprogramming.** v. 108, n. 5, Proc Natl Acad Sci U S A, Feb 1 2011, p. 2130-5.

MILDE-LANGOSCH, K.; RIETHDORF, S. **Role of cell-cycle regulatory proteins in gynecological cancer.** v. 196, n. 2, J Cell Physiol, Aug 2003, p. 224-44.

MIRABELLO, L. *et al.* **Elevated methylation of HPV16 DNA is associated with the development of high grade cervical intraepithelial neoplasia.** v. 132, n. 6, Int J Cancer, Mar 15 2013, p. 1412-22.

MUNGER, K. *et al.* **Mechanisms of human papillomavirus-induced oncogenesis.** v. 78, n. 21, J Virol, Nov 2004, p. 11451-60.

MUNOZ, N. *et al.* **Epidemiologic classification of human papillomavirus types associated with cervical cancer.** v. 348, n. 6, N Engl J Med, Feb 6 2003, p. 518-27.

MUNOZ, N. *et al.* **Chapter 1: HPV in the etiology of human cancer.** v. 24 Suppl 3, Vaccine, Aug 31 2006, p. S3/1-10.

MURAKAMI, I. *et al.* **Methylation of human papillomavirus-52 and -58 is a candidate biomarker in cervical neoplasia.** v. 58, n. 1, J Clin Virol, Sep 2013, p. 149-54.

NAMBARU, L. *et al.* **Prognostic significance of HPV physical status and integration sites in cervical cancer.** v. 10, n. 3, Asian Pac J Cancer Prev, Jul-Sep 2009, p. 355-60.

NERI, F. *et al.* **Intragenic DNA methylation prevents spurious transcription initiation.** v. 543, Nature, March 2017, p. 72-77.

O'CONNOR, M. S.-Y. C. H.-U. B. **Transcription factor binding sites in the long control region of genital HPVs.** In G. Myers, H. Delius, J. Icenogel, H.-U. Bernard, C. Baker, A. Halpern, and C. Wheeler (ed.), Human papillomaviruses 1995. Theoretical Biology and Biophysics, Los Alamos, N.Mex.,1995, p. 21-24. The HPV database, <https://pave.niaid.nih.gov/lanl-archives/compendium/95PDF/3/oconnor.pdf>.

OTTINGER, M. *et al.* **Cell-type specific transcriptional activities among different papillomavirus long control regions and their regulation by E2.** v. 395, n. 2, Virology, Dec 20 2009, p. 161-71.

PARK, J. S. *et al.* **Physical status and expression of HPV genes in cervical cancers.** v. 65, n. 1, Gynecol Oncol, Apr 1997, p. 121-9.

PINTO, A. P.; TULIO, S.; CRUZ, O. R. **[Hpv cofactors in cervical carcinogenesis].** v. 48, n. 1, Rev Assoc Med Bras, Jan-Mar 2002, p. 73-8.

PIYATHILAKE, C. J. *et al.* **Folate and vitamin B12 may play a critical role in lowering the HPV 16 methylation-associated risk of developing higher grades of CIN.** v. 7, n. 11, Cancer Prev Res (Phila), Nov 2014, p. 1128-37.

ROLLINS, R. A. *et al.* **Large-scale structure of genomic methylation patterns.** v. 16, n. 2, Genome Res, Feb 2006, p. 157-63.

ROMANCZUK, H.; THIERRY, F.; HOWLEY, P. M. **Mutational analysis of cis elements involved in E2 modulation of human papillomavirus type 16 P97 and type 18 P105 promoters.** v. 64, n. 6, J Virol, Jun 1990, p. 2849-59.

SANDERS, C. M.; MAITLAND, N. J. **Kinetic and equilibrium binding studies of the human papillomavirus type-16 transcription regulatory protein E2 interacting with core enhancer elements.** v. 22, n. 23, Nucleic Acids Res, Nov 25 1994, p. 4890-7.

SAWADA, M. *et al.* **Increased expression of DNA methyltransferase 1 (DNMT1) protein in uterine cervix squamous cell carcinoma and its precursor lesion.** v. 251, n. 2, *Cancer Lett*, Jun 28 2007, p. 211-9.

SCHIFFMAN, M. *et al.* **The carcinogenicity of human papillomavirus types reflects viral evolution.** v. 337, n. 1, *Virology*, Jun 20 2005, p. 76-84.

SCHIFFMAN, M. *et al.* **A population-based prospective study of carcinogenic human papillomavirus variant lineages, viral persistence, and cervical neoplasia.** v. 70, n. 8, *Cancer Res*, Apr 15 2010, p. 3159-69.

SCHWARZ, E. *et al.* **Structure and transcription of human papillomavirus sequences in cervical carcinoma cells.** v. 314, n. 6006, *Nature*, Mar 7-13 1985, p. 111-4.

SICHERO, L. *et al.* **High grade cervical lesions are caused preferentially by non-European variants of HPVs 16 and 18.** v. 120, n. 8, *Int J Cancer*, Apr 15 2007, p. 1763-8.

STEGER, G.; CORBACH, S. **Dose-dependent regulation of the early promoter of human papillomavirus type 18 by the viral E2 protein.** v. 71, n. 1, *J Virol*, Jan 1997, p. 50-8.

SHUKLA, S. *et al.* **Physical state & copy number of high risk human papillomavirus type 16 DNA in progression of cervical cancer.** v. 139, n. 4, *Indian J Med Res*, Apr 2014, p. 531-543.

TAN, S. H. *et al.* **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.** v. 68, n. 10, *J Virol*, Oct 1994, p. 6411-20.

THAIN, A. *et al.* **CpG methylation directly inhibits binding of the human papillomavirus type 16 E2 protein to specific DNA sequences.** v. 70, n. 10, *J Virol*, Oct 1996, p. 7233-5.

THIERRY, F.; HOWLEY, P. M. **Functional analysis of E2-mediated repression of the HPV18 P105 promoter.** v. 3, n. 1, *New Biol*, Jan 1991, p. 90-100.

THIERRY, F.; YANIV, M. **The BPV1-E2 trans-acting protein can be either an activator or a repressor of the HPV18 regulatory region.** v. 6, n. 11, *EMBO J*, Nov 1987, p. 3391-7.

TOMMASINO, M. *et al.* **The role of TP53 in Cervical carcinogenesis.** v. 21, n. 3, *Hum Mutat*, Mar 2003, p. 307-12.

TROTTIER, H.; FRANCO, E. L. **The epidemiology of genital human papillomavirus infection.** v. 24 Suppl 1, *Vaccine*, Mar 30 2006, p. S1-15.

TURAN, T. *et al.* **Methylation of the human papillomavirus-18 L1 gene: a biomarker of neoplastic progression?** v. 349, n. 1, *Virology*, May 25 2006, p. 175-83.

TURAN, T. *et al.* **High-throughput detection of human papillomavirus-18 L1 gene methylation, a candidate biomarker for the progression of cervical neoplasia.** v. 361, n. 1, *Virology*, Apr 25 2007, p. 185-93.

VACCARELLA, S. *et al.* **Smoking and human papillomavirus infection: pooled analysis of the International Agency for Research on Cancer HPV Prevalence Surveys.** v. 37, n. 3, *Int J Epidemiol*, Jun 2008, p. 536-46.

VASILJEVIC, N. *et al.* **A comparison of methylation levels in HPV18, HPV31 and HPV33 genomes reveals similar associations with cervical precancers.** v. 59, n. 3, *J Clin Virol*, Mar 2014, p. 161-6.

VAN TINE, B.A. *et al.* **Clonal selection for transcriptionally active viral oncogenes during progression to cancer.** v. 78, *J Virol*, 2004, p. 11172–86.

VERNON, S. D. *et al.* **Association of human papillomavirus type 16 integration in the E2 gene with poor disease-free survival from cervical cancer.** v. 74, n. 1, *Int J Cancer*, Feb 20 1997, p. 50-6.

VIDAL, J. P. *et al.* **Genetic diversity of HPV16 and HPV18 in Brazilian patients with invasive cervical cancer.** v. 88, n. 7, J Med Virol, Jul 2016, p. 1279-87.

VINOKUROVA, S.; VON KNEBEL DOEBERITZ, M. **Differential methylation of the HPV 16 upstream regulatory region during epithelial differentiation and neoplastic transformation.** v. 6, n. 9, PLoS One, 2011, p. e24451.

WALBOOMERS, J. M. *et al.* **Human papillomavirus is a necessary cause of invasive cervical cancer worldwide.** v. 189, n. 1, J Pathol, Sep 1999, p. 12-9.

WENTZENSEN, N. *et al.* **Methylation of HPV18, HPV31, and HPV45 genomes and cervical intraepithelial neoplasia grade 3.** v. 104, n. 22, J Natl Cancer Inst, Nov 21 2012, p. 1738-49.

XI, L. F. *et al.* **Inverse association between methylation of human papillomavirus type 16 DNA and risk of cervical intraepithelial neoplasia grades 2 or 3.** v. 6, n. 8, PLoS One, 2011, p. e23897.

XI, L. F. *et al.* **Viral load in the natural history of human papillomavirus type 16 infection: a nested case-control study.** v. 203, n. 10, J. Infect Dis, May 2011, p. 1425-1433.

XI, L. F. *et al.* **Risk for high-grade cervical intraepithelial neoplasia associated with variants of human papillomavirus types 16 and 18.** v. 16, n. 1, Cancer Epidemiol Biomarkers Prev, Jan 2007, p. 4-10.

XU, B. *et al.* **Multiplex Identification of Human Papillomavirus 16 DNA Integration Sites in Cervical Carcinomas.** v. 8, n. 6, PLoS One, 2013, p. e66693.

ZAMPIERI, M. *et al.* **Reconfiguration of DNA methylation in aging.** v. 151, Mech Ageing Dev, Nov 2015, p. 60-70.

9. ANEXOS

9.1. Anexo 1. Registro do comitê de ética em pesquisa (CEP).



A(o): Dra. Liz Maria Almeida
Pesquisador(a) Principal

Rio de Janeiro, 25 de fevereiro de 2011.

Registro CEP nº 156/10 (Este nº. deve ser citado nas correspondências referentes a este estudo)
Título do Estudo: Estudo dos genótipos de HPV presentes em tumores do colo do útero em mulheres matriculadas no INCA

Prezado(a) Pesquisador(a),

Informo a composição do Comitê de Ética em Pesquisa do Instituto Nacional de Câncer que analisou e aprovou o estudo acima especificado.

<i>Adriana Scheliga</i>	<i>Médica Oncologista Clínica- HC I – Coordenadora;</i>
<i>Alexandra Oliveira Martins</i>	<i>Enfermeira do Serv. de Pesq. Clínica - CGTC</i>
<i>Ana Lúcia Amaral Eisenberg</i>	<i>Médica Patologista – DNPAT</i>
<i>Ana Fátima Paranhos Mizuti</i>	<i>Psicóloga do Serviço de Psicologia - HC I</i>
<i>André Marcelo Machado Soares</i>	<i>Teólogo / Filósofo Extra INCA</i>
<i>Carlos Henrique D. Silva</i>	<i>Médico HC IV – Coordenador Substituto</i>
<i>Célia Helena Fernandes Costa</i>	<i>Enfermeira da Oncologia - HC I</i>
<i>Cláudio Gustavo Stefanello</i>	<i>Coordenador do Laboratório de BNT - CGTC</i>
<i>Cristiano Guedes Dupre</i>	<i>Médico Oncologista - HC I</i>
<i>Danielle Nogueira Ramos</i>	<i>Doutora em Epidemiologia e Biostatística - CONPREF</i>
<i>Eliana Teixeira Maranhão</i>	<i>Fisioterapeuta do Serviço de Fisioterapia - HC I</i>
<i>Hilton Spíndola Antunes</i>	<i>Dentista do Serviço de Pesquisa Clínica - CGTC</i>
<i>Ivan da Costa Marques</i>	<i>Matemático Extra INCA</i>
<i>Márcia de Fátima Batista de Menezes</i>	<i>Enfermeira da Divisão de Enfermagem - HC I</i>
<i>Michelle Quarli</i>	<i>Enfermeira - Extra INCA</i>
<i>Paulo Alexandre Mira</i>	<i>Médico Oncologista - HC II</i>
<i>Rosane Vianna Jorge</i>	<i>Doutora em Química Biológica da Divisão de Farmacologia - CGTC</i>
<i>Urquellê das Neves Pinheiro</i>	<i>Representante de Estudantes - CESRJ</i>
<i>Walter Esteves Pereira</i>	<i>Advogado Extra INCA</i>
<i>Wlana Barbosa Felippo</i>	<i>Enfermeira da Comissão de Infecção Hospitalar - HC III</i>

Atenciosamente,

Dra. Adriana Scheliga
Coordenadora do Comitê de Ética em Pesquisa
CEP-INCA

9.2. Anexo 2. Prevalência de HPV na população do estudo principal.

HPV	N Total	CCE (N)	%	ADC (N)	%	Outros (N)	%	ND	%
16	370	294	79.5	45	12.2	30	8.1	1	0.3
18	77	39	50.6	22	28.6	15	19.5	1	1.3
26	2	2	100.0	0	0.0	0	0.0	0	0.0
31	7	7	100.0	0	0.0	0	0.0	0	0.0
33	7	7	100.0	0	0.0	0	0.0	0	0.0
35	12	9	75.0	2	16.7	1	8.3	0	0.0
39	6	5	83.3	0	0.0	1	16.7	0	0.0
45	33	24	72.7	6	18.2	3	9.1	0	0.0
51	2	2	100.0	0	0.0	0	0.0	0	0.0
52	8	8	100.0	0	0.0	0	0.0	0	0.0
56	2	2	100.0	0	0.0	0	0.0	0	0.0
58	11	11	100.0	0	0.0	0	0.0	0	0.0
59	6	4	66.7	1	16.7	1	16.7	0	0.0
68	2	2	100.0	0	0.0	0	0.0	0	0.0
73	8	7	87.5	0	0.0	1	12.5	0	0.0
83	1	1	100.0	0	0.0	0	0.0	0	0.0
X	16	11	68.8	1	6.3	4	25.0	0	0.0
MI	20	16	80.0	2	10.0	2	10.0	0	0.0
NI	4	3	75.0	1	25.0	0	0.0	0	0.0
Total	594	454		80		58		2	

X: Identificada a presença de HPV por PCR utilizando *primers* consenso PGMY entretanto; MI: múltiplas infecções por HPV; NI: não identificada a presença de HPV por PCR utilizando *primers* consenso PGMY. CCE: carcinoma de células escamosas. ADC: adenocarcinomas. Outros: histologia diferentes de CCE e ADC. ND: Não determinado

9.3. Anexo 3. Artigo publicado sobre o projeto de doutorado

Virus Research 232 (2017) 48–53



Contents lists available at ScienceDirect

Virus Research

journal homepage: www.elsevier.com/locate/virusres



Methylation at 3'LCR of HPV16 can be affected by patient age and disruption of *E1* or *E2* genes



Sérgio Menezes Amaro Filho^a, Neilane Bertoni^c, Ayslan Castro Brant^{a,d}, João Paulo Castello Branco Vidal^a, Shayany Pinto Felix^a, Silvia Maria Baeta Cavalcanti^b, Fernanda N. Carestiatto^b, Luís Felipe Leite Martins^c, Liz Maria de Almeida^c, Miguel Angelo Martins Moreira^{a,*}

^a Genetics Program – Instituto Nacional de Câncer (INCA), Rio de Janeiro, Brazil

^b Department of Microbiology and Parasitology, Universidade Federal Fluminense, Niterói, Brazil

^c Population Research Department – Instituto Nacional de Câncer (INCA), Rio de Janeiro, Brazil

^d Genetics Department – Universidade Federal do Rio de Janeiro, Brazil

ARTICLE INFO

Article history:

Received 8 December 2016

Received in revised form 24 January 2017

Accepted 24 January 2017

Available online 29 January 2017

Keywords:

Human papillomavirus

HPV16

Invasive cervical cancer

Methylation

Episomal HPV genome

HPV integration

ABSTRACT

CpG methylation at early promoter of HPV16 DNA, in the 3' end of the Long Control Region (3'LCR), has been associated to the presence of episomal forms of viral genome and, consequently, intact *E1* and *E2* ORFs. The DNA methylation would block the access of *E2* viral protein to the *E2* binding sites at early-promoter. However, is still unclear if methylation at 3'LCR of HPV16 DNA can also vary depending of other tumor characteristics in addition to viral DNA physical state. In this study, we evaluate whether the methylation level at the five CpG located at 3'LCR of HPV16 is associated to patient age and *E1* and/or *E2* ORFs integrity. DNA pyrosequencing was used to measure the methylation level in 69 invasive cervical cancer samples obtained from biopsies of patients attended at Brazilian National Institute of Cancer (INCA). PCR amplifications were performed to assess disruption status of *E1* and *E2* genes of HPV16. The methylation average per sample ranged widely, from <1 to 88.00%. Presence of intact *E1/E2* genes and patient age were positively associated with average methylation in both bivariate analyses ($p=0.003$ and $p=0.006$, respectively), and multivariate analysis ($p=0.002$ and $p=0.021$, respectively), adjusted for tumor type (squamous cell carcinomas or adenocarcinomas) and HPV16 lineage. These findings showed that presence of intact *E1/E2* open reading frames was associated with high levels of DNA methylation, and older patients showed higher levels of methylation than younger ones independently of viral genome disruption.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Human papillomaviruses (HPV) are non-enveloped viruses (~55 nm) with a circular genome of double-stranded DNA of approximately 8 Kilobase-pairs (kb) inside an icosahedral capsid. More than 40 HPV genotypes have been found to infect the genital mucosa and 12 among them are well characterized as high-risk types (HR-HPV) for cervical cancer according to the International Agency for Research on Cancer IARC Working Group on the Evaluation of Carcinogenic Risks to Human (2009). The HPV genome contains three different regions: (1) the early region (E),

with genes coding for proteins predominantly expressed in early stages of infection (*E1*, *E2*, *E4*, *E5*, *E6* and *E7*), comprising 4 kb of the viral genome and associated with viral DNA replication, regulation of HPV gene expression, and oncogenesis; (2) the late region (L) of approximately 3 kb, with genes encoding structural capsid proteins (*L1* and *L2*) expressed in later stages of infection; and (3) the Long Control Region (LCR), a non-coding region of approximately 1 kb, functionally divided into 5', central and 3' portions, with sequence motifs required for gene expression and initiation of viral DNA replication (O'Connor et al., 1995).

As in eukaryotic organisms, the HPV virion presents the genome arranged in a chromatin-like structure (Favre et al., 1977). These structure may be modified by host enzymes like histone acetyltransferases (HATs) and histone deacetylases (HDACs), which may allow or block access to the transcriptional machinery (reviewed by Kalantari et al., 2004). Additionally, host DNA methyltransferases

* Corresponding author at: Genetics Program, Instituto Nacional de Câncer (INCA), Andre Canvalcanti 37, Rio de Janeiro, RJ, Brazil.
E-mail address: miguelm@inca.gov (M.A.M. Moreira).

(DNMTs) are capable of adding methyl groups to cytosines at CpG sites of the viral genome which may alter HPV expression (Badal et al., 2004).

Although methylation of CpG sites has been found to be a major epigenetic modification frequently associated with transcriptional repression, the effect of HPV DNA methylation in the regulation of viral transcription is still unclear (Bird, 2002). Methylation of the *L1* and/or *L2* genes of HPV16 and HPV18 has been proposed as a marker for cervical neoplastic progression. This event probably results in transcriptional silencing of these late genes whose products are not required for neoplastic transformation. On the other hand, methylation at the 3' segment of the LCR (3'LCR) in some cervical cancers could be associated with the episomal state of HPV DNA (Chaiwongkot et al., 2013).

The 3'LCR, where the *E6/E7* promoter is located, controls expression of *E6* and *E7* viral oncogenes of HPV16 and HPV18. This promoter has been well characterized for HPV16 and HPV18, named P97 and P105, respectively. This region is essential for immortalization of human keratinocytes by triggering massive production of *E6* and *E7* oncoproteins in high-risk HPV infections. The activity of P97 and P105 can be regulated by host transcription factors like AP-1, NF1, SP1, TFIID, TF1, Oct-1 and the viral factor *E2* (Desaintes and Demeret, 1996; Lung et al., 2012; Thierry and Howley, 1991).

The *E2* viral protein plays a key role during HPV infection and HPV-induced carcinogenesis, repressing transcription of *E6* and *E7* in a dose dependent process (McBride, 2013). Three conserved *E2* DNA binding sites (*E2BS*) at the 3'LCR, with the consensus palindromic motif ACCG(N)₄CGGT, have been described (Steger and Corbach, 1997a). *E2BS* at the 3'LCR are low affinity binding sites and, in presence of high concentrations of *E2* proteins, transcription of the *E6/E7* promoter is repressed. *E2* activity is associated to displacement of P97 and P105 promoter activators, Sp1 and TFIID, from their respective binding sites. Thus, the absence or low level of *E2* proteins results in overexpression of *E6* and *E7* oncogenes and cancer progression (Tan et al., 1994; Steger and Corbach, 1997b).

Two *E2BS* at the 3'LCR of HPV16 (*E2BS#3* and *E2BS#4*, distal and proximal, respectively) overlapping four CpG sites at genomic positions 37, 43, 52 and 58 (in respect to GenBank accession K02718.1), are flanked by a close CpG site, at position 31 of an SP1 transcription factor binding site. In invasive cervical cancer (ICC) with HPV16 infection, these sites have shown a wide spectrum of methylation levels between samples rather than a similar level of methylation (Chaiwongkot et al., 2013; Hong et al., 2008).

In malignant lesions, *E2* functions are frequently lost due to integration of the HPV genome to the host chromosome. Integration occurs by rupture of the viral genome downstream of *E6* and *E7*, preventing *E2* transcription and resulting in upregulation of *E6* and *E7*. Conversely, in approximately 40% of cervical tumors, presence of episomal HPV16 DNA has been reported (Cricca et al., 2007; Cullen et al., 1991; Liu et al., 2016), suggesting presence of a functional *E2* protein. Additionally, intact *E2* ORFs are also present in tumors with integrated, multiple viral genome tandem copies. Chaiwongkot et al. reported these findings in squamous cell carcinomas associated with HPV16, suggesting that a higher methylation level of *E2BS* at 3'LCR probably contributed to upregulate *E6* and *E7* transcription by blocking *E2* binding (Chaiwongkot et al., 2013). However, these studies did not evaluate the presence of *E2* protein.

Methylation at CpG sites is a complex event in the human genome, displaying a tissue-specific pattern, and influenced by aging and life style (Christensen et al., 2009; Gama-Sosa et al., 1983; Lim and Song, 2012). Moreover, the *E6* protein of HPV16 in cervical carcinoma cell lines SiHa and CaSki has been shown to increase the level of DNMT1, one of the enzymes involved in CpG methylation and epigenetic silencing of tumor suppressor genes (Au Yeung et al., 2010).

In addition to presence of a functional *E2* ORF, no data are available on the influence of other tumor characteristics on the methylation pattern of HPV DNA. In this study, we evaluate whether the methylation level of the five CpGs sites at the SP1 binding site, *E2BS#3*, and *E2BS#4*, at the 3'LCR of HPV16, may also be associated with patient age and *E1* and/or *E2* ORFs integrity.

2. Material and methods

2.1. Samples

The study material was selected from a pool of 334 biopsies of invasive cervical cancer associated to HPV16. Patients were referred to Instituto Nacional de Câncer (Rio de Janeiro, Brazil) for cancer treatment between June 2011 and March 2014. They were requested to provide epidemiological data in a questionnaire after sign a consent form. All procedures were approved by the Institutional Ethics Committee (CAAE: 53398416.0.0000.5274). Biopsies were collected and stored in RNA-Later at -80°C . DNA was isolated with QIAamp DNA mini Kit (Qiagen, Hilden, Germany). Presence of HPV DNA was detected by polymerase chain reaction (PCR) with PGMY 09/11 primers or by nested PCR with PGMY09/11 and GP5+/GP6+. HPV types were identified by DNA sequencing and BLASTn (<http://blast.ncbi.nlm.nih.gov/>). Samples with multiple infections were characterized by visual inspection of electropherograms and the detection of overlapping sequence peaks, these samples were not included in this work.

Following the nomenclature proposed by Burk et al. (2013), HPV16 lineages were identified in a set of 334 samples based on LCR and *E6* sequencing. These comprised 217 of lineage A, 10 of B, 10 of C, and 97 of D (see Vidal et al., 2016). For the current study, seventy samples were selected taking into account the following criteria: (1) inclusion of samples representing all the four HPV16 lineages, and (2) samples representing the histological types adenocarcinomas (ADN) or squamous cell carcinoma (SCC). Thus, the select samples comprised: the 10 samples infected by HPV16 of lineage B (all SCC), the 10 infected by HPV16 of lineage C (1 ADN and 9 SCC), 26 infected by lineage A (12 ADN and 14 SCC), and 24 infected by lineage D (12 ADN and 12 SCC). The different histological types were randomly selected among the ADNs and SCCs of lineages A and D. One sample of lineage B was excluded because PCR products could not be amplified following sodium bisulfite treatment. Cell lines CaSki and SiHa were used as controls for bisulfite treatment and pyrosequencing.

2.2. Sodium bisulfite treatment and PCR amplification

Sodium bisulfite treatment was carried out with EpiTect-Bisulfite Kit (Qiagen, Germany), with an input of 300–1500 ng of DNA for cytosine to uracil conversion. Following treatment, a region covering 178 bp, containing five CpG sites (with cytosine positions 31, 37, 43, 52, and 58) at the 3'LCR of HPV16 was PCR-amplified by nested PCR following Rajeevan et al. (2006).

PCR was carried out in 30 μL mixtures containing 0.2 mM of each dNTP, 6 pmol of each primer, 1U of Platinum Taq DNA Polymerase (Life Technologies) and 1X PCR buffer (67 mM Tris pH 8.8, 6.7 mM MgSO_4 , 16.6 mM $(\text{NH}_4)_2\text{SO}_4$, and 10 mM 2-mercaptoethanol (Kocher et al., 1989). PCR conditions were: 95°C for 6 min, followed by 45 cycles of 95°C for 1 min, 61°C for 1 min, 72°C for 1 min, and one step of final extension at 72°C for 5 min. Presence of PCR products was verified in agarose gels.

2.3. Quantitation of DNA methylation by pyrosequencing

PCR products were submitted to pyrosequencing in a PyroMark Q24 platform (Qiagen, Germany) following a standard

protocol. Briefly, 1 μ L of streptavidin beads (GE Healthcare, Buckinghamshire, UK), 40 μ L PyroMark binding buffer (Qiagen), 20 μ L PCR product and 19 μ L water were mixed and incubated for 10 min on a shaking platform at 1400 rpm. A Biotage Q24 Vacuum Prep Workstation was used for separating, denaturing and washing PCR products which were subsequently added to a microtiter plate containing 25 μ L of annealing buffer with 0.3 μ M of sequencing primer (5'-AATTTATGTATAAAATTAAGGG-3'). Primer annealing was carried out by incubation at 80 °C for 2 min and cooling to room temperature before pyrosequencing. PyroGold reagents, including enzyme, substrate and nucleotides, were used for the pyrosequencing reaction. Pyrograms were generated and analyzed with the PyroMark Q24, v.2.0.6 (Qiagen; Hilden; Germany). Dispensation order was 5'-YGTAATYGAAATYGTTGAATYGAAATYG-3'. Based on the findings of Badal et al. (2003), and Fernandez et al. (2009), the genomic DNA of the cell lines CaSki and SiHa cell lines were used as controls for hypermethylation or hypomethylation pattern of the CpG at 3' LCR, respectively.

2.4. Disruption of E1 or E2 gene

HPV integration to the host DNA genome frequently occurs with E1 or E2 disruption (Xu et al., 2013) resulting in suppression of E2 transcription. The presence of a non-disrupted E2 gene was identified by PCR amplification of overlapping fragments encompassing the E1 and E2 coding regions using the 8 primer pairs described by Collins et al. (2009) and Vernon et al. (1997).

PCR was carried out in 25 μ L mixtures containing 0.2 mM of each dNTP, 25 pmol of each primer (Supplemental Table 1), 1 U of Platinum Taq DNA Polymerase (Life Technologies) and 1X buffer, 2 mM MgCl₂. PCR conditions were: 95 °C for 5 min, followed by 35 cycles of 95 °C for 30 s, annealing temperature for 30 s (see Supplemental Table 1), extension at 72 °C for 40 s, and final extension at 72 °C for 5 min. Presence of PCR products (Supplemental Fig. 1) was verified in 1.5% ultrapure agarose gels (Life Technologies). Absence of an amplicon, which indicates disruption of E1 or E2, was confirmed by an independent PCR for the same region.

2.5. Statistical analysis

Exploratory analysis was carried out to describe sample characteristics. The level of methylation of each CpG site per sample was estimated as the proportion of methylated cytosines, ranging from 0 (without methylation) to 1 (totally methylated). To evaluate the influence of other factors on the methylation level of this region, we considered the average methylation of the five CpG sites per sample as outcome, because the binding of E2 proteins to E2BS#3 and E2BS#4 physically encompass the CpG sites at genomic positions 31, 37, 43, 52, and 58. Bivariate models were elaborated to evaluate the association between the level of average methylation with presence of intact or disrupted E2/E1 open reading frames (ORFs), and patient age at diagnosis, analyzed as a continuous variable. As our outcome variable (average methylation level) was a continuous proportion with non-normality assumption, a Quasi-likelihood regression with logit link was used. In multivariate models, we considered tumor type (SCC/ADN) and lineages as control variables because samples were not randomly selected. A p value < 0.05 was considered significant. Statistical analyzes were conducted with R v.3.2.4.

3. Results

Clinical and pathological characteristics of the 69 patients herein studied are summarized in Table 1. Age at diagnosis ranged from 19 to 79 years, with a mean of 44.61 (SD = 13.45) and a median of 43.00. In 21 tumors (31.34%) disruptions were detected contrary

Table 1
Characteristics of Study Population.

Characteristics	N
Overall population	69
Patient Age (year)	
Mean (SD): 44.61 (13.45)	
Median: 43	
Tumor types	
ADN	25
SCC	44
No information	-
HPV16 lineage	
A	26
B/C	19
D	24
No information	-
FIGO Stage	
I	22
II-IV	47
No information	-
Disruption of E1/E2	
Yes	21
No	46
No information	2

ADN: adenocarcinomas, SCC: squamous cell carcinomas.

All CpGs: average methylation of all five CpG.

N = number of samples.

to 46 (68.66%) in which these regions were intact. E1 disruption occurred in 13 tumors, while E2 was disrupted in two, and both genes were disrupted in six (Supplemental Table 2).

Following bisulfite treatment, the methylation level of each CpG site per sample was estimated by pyrosequencing for viral genomic positions 31, 37, 43, 52 and 58 (in respect to GenBank K02718.1). The methylation level (%) of each CpG site, and the average of methylation of these sites per sample are listed in Supplemental Table 2. The lowest methylation level was found at nt 31 (mean = 18%), overlapping with an Sp1 binding site. The other CpG sites showed a similar methylation level with a mean per site ranging from 21.50% to 23.00% (Fig. 1a). No statistical difference between the CpG sites were observed. The average methylation of the five CpG sites per sample ranged from <1 to 88.00% (Fig. 1b). Similar as previously reported (Badal et al., 2003), a higher average methylation was found for the CaSki cell line (96.00%) and a lower one for SiHa (3.20%).

Bivariate analysis showed that presence of disruption and patient age were positively associated with average methylation at the five CpG sites per tumor sample (p = 0.003 and p = 0.006, respectively; Table 2). With multivariate analysis, both variables (E1/E2 integrity and patient age) were also positively associated with average methylation level (p = 0.002 and p = 0.021, respectively; Table 2) following adjusting for tumor type (SCC or ADN) and HPV16 lineage. These findings showed that presence of intact E1/E2 open reading frames was associated with higher levels of methylation. Additionally, older patients showed higher levels of methylation than younger ones independently of HPV DNA disruption.

4. Discussion

Our results showed that the range of average methylation of the CpG sites located in the 3' LCR of HPV16, per invasive cervical cancer sample, was similar to previously reported (from 2.00% to 84.00%) in 17 squamous cell carcinomas analyzed by bisulfite treatment and pyrosequencing (Jacquin et al., 2013). Conflicting reports on the 3' LCR methylation pattern of HPV16 in invasive cervical cancer are likely to have resulted from using different methodologies and sample size. This is because methylation of this HPV16 region

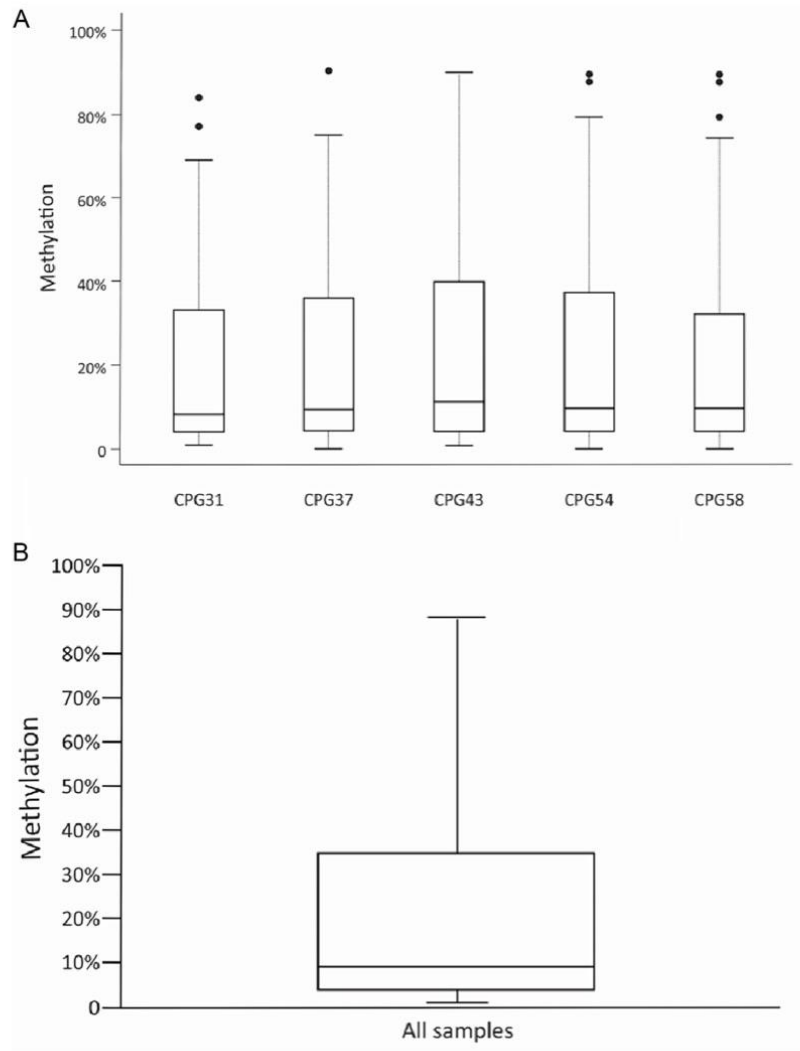


Fig. 1. Average methylation at 3' LCR of HPV16. In A, the average methylation is represented by individual CpG sites, and in B is shown the overall average methylation for all five CpG sites. Methylation levels are represented in boxes displaying upper and lower quartiles separated by the median line and whisker plots; black circles indicate outliers.

Table 2
Bivariate and multivariate analyzes of the average methylation of the five CpG sites at 3' LCR.

	Bivariate Analysis			Multivariate analysis ^a		
	Coefficient	Std. Error	p value	Coefficient	Std. Error	p value
Disruption of E1/E2						
Yes	1	–		1	–	
No	1.373	0.438	0.003	1.422	0.443	0.002
Age	0.034	0.012	0.006	0.028	0.012	0.021

^aAdjusted for tumor type (SCC vs ADN) and HPV16 lineage (A vs. B/C/D).

has been investigated with non-quantitative methods like Bisulfite Restriction Analysis (COBRA), methylation-specific PCR (MS-PCR) and PCR-based methylation sensitive restriction analysis (MSRA) (Ding et al., 2009; Leung et al., 2015; Mazumder Indra et al., 2011; Vinokurova and von Knebel Doeberitz, 2011), a reason why these findings cannot be compared with the quantitative data herein reported.

Our study showed association between methylation patterns and presence of disruption. Viral DNA integration with disruption of *E1* or *E2* and loss of the repressive activity of *E2* protein over the earlier promoter has been considered an essential step during malignant transformation. However, in 20% of cervical cancers, HPV16 DNA integration was not detected, potentially retaining *E2* function (Cricca et al., 2007; Cullen et al., 1991). In addition, experimental data showed that methylation of CpG sites at the 3' LCR *E2*BS blocked *E2* binding in association with loss of *E2* repressive activity (Tan et al., 1994; Thain et al., 1996). Our findings were coincident with this proposition, showing that a higher methylation at the 3' LCR in invasive cervical carcinomas was associated with a potentially functional *E2* gene. However, is important to highlight these findings do not allow to establish a direct association between the presence of *E2* protein and the higher methylation of CpG sites at 3' LCR. This is because the presence of an intact *E2* ORF does not imply in the presence of *E2* protein, as reported in samples from SCC and the CaSki cell line (Xue et al., 2012).

A positive correlation between methylation level at the 3' LCR of HPV16 and patient age was found in this study. In humans, aging has been associated with global hypomethylation and site-specific hypermethylation, mainly in CpG islands (Christensen et al., 2009). To date, our findings represent the first report showing the likely influence of aging on HPV16 DNA methylation in cancer samples, independently of gene disruptions. Previous reports on the association between HPV DNA methylation and patient age did not consider the presence of an episomal or integrated viral genome, and showed contradictory findings for CpGs placed in LCRs and in other HPV genomic regions. Leung et al. reported association between patient age and increase of methylation level for HPV18 *E2*BS4 (in 5' LCR), but not for HPV16 in samples of SCC (Leung et al., 2015). Mirabello et al. reported increase in methylation in CpGs along the HPV16 genome associated with patient age in CIN2-3 (Cervical Intraepithelial Neoplasia 2 and 3) but not in control samples (women with HPV16 clearance in the period of <2 years) (Mirabello et al., 2013).

The positive correlation between patient age and 3' LCR methylation indicates that aging affects tumor epigenetics. In normal tissues was observed a global decline of CpG methylation with human age (Zampieri et al., 2015) and an inverse correlation between the amount of transcripts of DNA Methyltransferases (DNMT1 and DNMT3b) with age (Ciccarone et al., 2016). In cervical cancer and CIN1-3 a higher expression of DNMT1 protein was observed when compared with normal squamous epithelium (Sawada et al., 2007), and in cervical cancer cell lines was shown that *E6* and *E7* increased the expression of DNA Methyltransferases and the methylation of tumor suppressor genes (Au Yeung et al., 2010; Li et al., 2015). The correlation between methylation and age are probably associated with the neoplastic process and the over expression of the DNMTs, our data in cervical cancer and those from Mirabello et al. (2013) (for CIN2-3 and women with HPV16 clearance) showed that this correlation is observed only in high grade lesions and cancer but not in women with HPV16 clearance.

Although this study might be restricted to a relatively small sample of different HPV16 lineages and histological tumor types, our findings highlighted the importance of considering patient age as a factor that might affect methylation status in HPV DNA. This is because methylation has been used as a potential marker for presence of high grade lesions or invasive cancer. Further studies

should be carried out for testing the association of patient age with methylation in other HPV types and HPV genomic regions.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgments

We want to thank Dr. Héctor N. Seunez for a critical revision of the manuscript. This work was supported by National Institute for Cancer Control (INCT do Cancer: <http://www.inct-cancer-control.com.br/>), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ), Ministry of Health – Brazil, Pan-American Health Organization (PAHO).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.virusres.2017.01.022>.

References

- Au Yeung, C.L., Tsang, W.P., Tsang, T.Y., Co, N.N., Yau, P.L., Kwok, T.T., 2010. HPV-16 *E6* upregulation of DNMT1 through repression of tumor suppressor p53. *Oncol. Rep.* 24, 1599–1604.
- Badal, V., Chuang, L.S., Tan, E.H., Badal, S., Villa, L.L., Wheeler, C.M., Li, B.F., Bernard, H.U., 2003. CpG methylation of human papillomavirus type 16 DNA in cervical cancer cell lines and in clinical specimens: genomic hypomethylation correlates with carcinogenic progression. *J. Virol.* 77, 6227–6234.
- Badal, S., Badal, V., Calleja-Macias, I.E., Kalantari, M., Chuang, L.S., Li, B.F., Bernard, H.U., 2004. The human papillomavirus-18 genome is efficiently targeted by cellular DNA methylation. *Virology* 324, 483–492.
- Bird, A., 2002. DNA methylation patterns and epigenetic memory. *Genes Dev.* 16, 6–21.
- Burk, R.D., Harari, A., Chen, Z., 2013. Human papillomavirus genome variants. *Virology* 445, 232–243.
- Chaiwongkot, A., Vinokurova, S., Pientong, C., Ekalaksananan, T., Kongyiygoes, B., Kleebkaow, P., Chumworathayi, B., Patarapadungkit, N., Reuschenbach, M., von Knebe Doeberitz, M., 2013. Differential methylation of *E2* binding sites in episomal and integrated HPV 16 genomes in preinvasive and invasive cervical lesions. *Int. J. Cancer* 132, 2087–2094.
- Christensen, B.C., Houseman, E.A., Marsit, C.J., Zheng, S., Wrensch, M.R., Wiemels, J.L., Nelson, H.H., Karagas, M.R., Padbury, J.F., Bueno, R., Sugarbaker, D.J., Yeh, R.F., Wiencke, J.K., Kelsey, K.T., 2009. Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. *PLoS Genet.* 5, e1000602.
- Ciccarone, F., Malavolta, M., Calabrese, R., Guastafierro, T., Bacalini, M.G., Reale, A., Franceschi, C., Capri, M., Hervonen, A., Hurme, M., Grubeck-Loebenstein, B., Koller, B., Bernhardt, J., Schön, C., Slagboom, P.E., Toussaint, O., Sikora, E., Gonos, E.S., Breusing, N., Grune, T., Jansen, E., Dollé, M., Moreno-Villanueva, M., Sindlinger, T., Bürkle, A., Zampieri, M., Caiafa, P., 2016. Age-dependent expression of DNMT1 and DNMT3 B in PBMCs from a large European population enrolled in the MARK-AGE study. *Aging Cell* 15, 755–765.
- Collins, S.I., Constantinou-Williams, C., Wen, K., Young, L.S., Roberts, S., Murray, P.G., Woodman, C.B., 2009. Disruption of the *E2* gene is a common and early event in the natural history of cervical human papillomavirus infection: a longitudinal cohort study. *Cancer Res.* 69, 3828–3832.
- Cricca, M., Morselli-Labate, A.M., Venturoli, S., Ambretti, S., Gentilomi, G.A., Gallinella, G., Costa, S., Musiani, M., Zerbin, M., 2007. Viral DNA load, physical status and *E2/E6* ratio as markers to grade HPV16 positive women for high-grade cervical lesions. *Gynecol. Oncol.* 106, 549–557.
- Cullen, A.P., Reid, R., Campion, M., Lorincz, A.T., 1991. Analysis of the physical state of different human papillomavirus DNAs in intraepithelial and invasive cervical neoplasia. *J. Virol.* 65, 606–612.
- Desaintes, C., Demeret, C., 1996. Control of papillomavirus DNA replication and transcription. *Semin. Cancer Biol.* 7, 339–347.
- Ding, D.C., Chiang, M.H., Lai, H.C., Hsiung, C.A., Hsieh, C.Y., Chu, T.Y., 2009. Methylation of the long control region of HPV16 is related to the severity of cervical neoplasia. *Eur. J. Obstet. Gynecol. Reprod. Biol.* 147, 215–220.
- Favre, M., Breitbart, F., Croissant, O., Orth, G., 1977. Chromatin-like structures obtained after alkaline disruption of bovine and human papillomaviruses. *J. Virol.* 21, 1205–1209.
- Fernandez, A.F., Rosales, C., Lopez-Nieva, P., Grana, O., Ballestar, E., Ropero, S., Espada, J., Melo, S.A., Lujambio, A., Fraga, M.F., Pino, I., Javierre, B., Carmona, F.J., Acquadro, F., Steenbergen, R.D., Snijders, P.J., Meijer, C.J., Pineau, P., Dejean, A.,

- Lloveras, B., Capella, G., Quer, J., Buti, M., Esteban, J.I., Allende, H., Rodriguez-Frias, F., Castellsague, X., Minarovits, J., Ponce, J., Capello, D., Gaidano, G., Cigudosa, J.C., Gomez-Lopez, G., Pisano, D.G., Valencia, A., Piris, M.A., Bosch, F.X., Cahir-McFarland, E., Kieff, E., Esteller, M., 2009. The dynamic DNA methylomes of double-stranded DNA viruses associated with human cancer. *Genome Res.* 19, 438–451.
- Gama-Sosa, M.A., Midgett, R.M., Slagel, V.A., Githens, S., Kuo, K.C., Gehrke, C.W., Ehrlich, M., 1983. Tissue-specific differences in DNA methylation in various mammals. *Biochim. Biophys. Acta* 740, 212–219.
- Hong, D., Ye, F., Lu, W., Hu, Y., Wan, X., Chen, Y., Xie, X., 2008. Methylation status of the long control region of HPV 16 in clinical cervical specimens. *Mol. Med. Rep.* 1, 555–560.
- IARC Working Group on the Evaluation of Carcinogenic Risks to Human, 2009. A Review of Human Carcinogens, Biological Agents, IARC Monographs on the Evaluation of Carcinogenic Risks to Humans v. 100B, Lyon.
- Jacquin, E., Baraquin, A., Ramanah, R., Carcopino, X., Morel, A., Valmary-Degano, S., Bravo, I.G., de Sanjosé, S., Riethmuller, D., Mougín, C., Prétet, J.L., 2013. Methylation of human papillomavirus Type 16 CpG sites at E2-binding site 1 (E2BS1), E2BS2, and the Sp1-binding site in cervical cancer samples as determined by high-resolution melting analysis-PCR. *J. Clin. Microbiol.* 51, 3207–3215.
- Kalantari, M., Calleja-Macias, I.E., Tewari, D., Hagmar, B., Lie, K., Barrera-Saldana, H.A., Wiley, D.J., Bernard, H.U., 2004. Conserved methylation patterns of human papillomavirus type 16 DNA in asymptomatic infection and cervical neoplasia. *J. Virol.* 78, 12762–12772.
- Kocher, T.D., Thomas, W.K., Meyer, A., Edwards, S.V., Paabo, S., Villablanca, F.X., Wilson, A.C., 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proc. Natl. Acad. Sci. U. S. A.* 86, 6196–6200.
- Leung, T.W., Liu, S.S., Leung, R.C., Chu, M.M., Cheung, A.N., Ngan, H.Y., 2015. HPV 16 E2 binding sites 1 and 2 become more methylated than E2 binding site 4 during cervical carcinogenesis. *J. Med. Virol.* 87, 1022–1033.
- Li, L., Xu, C., Long, J., Shen, D., Zhou, W., Zhou, Q., Yang, J., Jiang, M., 2015. E6 and E7 gene silencing results in decreased methylation of tumor suppressor genes and induces phenotype transformation of human cervical carcinoma cell lines. *Oncotarget* 6, 23930–23943.
- Lim, U., Song, M.A., 2012. Dietary and lifestyle factors of DNA methylation. *Methods Mol. Biol.* 863, 359–376.
- Liu, Y., Lu, Z., Xu, R., Ke, Y., 2016. Comprehensive mapping of the human papillomavirus (HPV) DNA integration sites in cervical carcinomas by HPV capture technology. *Oncotarget* 7, 5852–5864.
- Lung, M.S., Zhang, N., Murray, V., 2012. Site-directed mutagenesis of human papillomavirus 18 promoter elements and tissue-specific expression in cervical carcinoma cells. *Virus Genes* 44, 395–402.
- Mazumder Indra, D., Singh, R.K., Mitra, S., Dutta, S., Chakraborty, C., Basu, P.S., Mondal, R.K., Roychoudhury, S., Panda, C.K., 2011. Genetic and epigenetic changes of HPV16 in cervical cancer differentially regulate E6/E7 expression and associate with disease progression. *Gynecol. Oncol.* 123, 597–604.
- McBride, A.A., 2013. The papillomavirus E2 proteins. *Virology* 445, 57–79.
- Mirabello, L., Schiffman, M., Ghosh, A., Rodriguez, A.C., Vasiljevic, N., Wentzensen, N., Herrero, R., Hildesheim, A., Wacholder, S., Scibior-Bentkowska, D., Burk, R.D., Lorincz, A.T., 2013. Elevated methylation of HPV16 DNA is associated with the development of high grade cervical intraepithelial neoplasia. *Int. J. Cancer* 132, 1412–1422.
- O'Connor, M., Chan, S.Y., Bernard, H.U., 1995. Transcription factor binding sites in the long control region of genital HPVs. In: Myers, G., Bernard, H.U., Delius, H., Baker, K., Icenogle, J., Halpern, A., Wheeler, C. (Eds.), *Human Papillomaviruses, Compendium, Part III*. Los Alamos National Laboratory, Los Alamos, pp. 21–40.
- Rajeevan, M.S., Swan, D.C., Duncan, K., Lee, D.R., Limor, J.R., Unger, E.R., 2006. Quantitation of site-specific HPV 16 DNA methylation by pyrosequencing. *J. Virol. Methods* 138, 170–176.
- Sawada, M., Kanai, Y., Arai, E., Ushijima, S., Ojima, H., Hirohashi, S., 2007. Increased expression of DNA methyltransferase 1 (DNMT1) protein in uterine cervix squamous cell carcinoma and its precursor lesion. *Cancer Lett.* 251, 211–219.
- Steger, G., Corbach, S., 1997a. Dose-dependent regulation of the early promoter of human papillomavirus type 18 by the viral E2 protein. *J. Virol.* 71, 50–58.
- Steger, G., Corbach, S., 1997b. Dose-dependent regulation of the early promoter of human papillomavirus type 18 by the viral E2 protein. *J. Virol.* 71, 50–58.
- 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 TFIIID. *J. Virol.* 68, 6411–6420.
- Thain, A., Jenkins, O., Clarke, A.R., Gaston, K., 1996. CpG methylation directly inhibits binding of the human papillomavirus type 16 E2 protein to specific DNA sequences. *J. Virol.* 70, 7233–7235.
- Thierry, F., Howley, P.M., 1991. Functional analysis of E2-mediated repression of the HPV18 P105 promoter. *New Biol.* 3, 90–100.
- Vernon, S.D., Unger, E.R., Miller, D.L., Lee, D.R., Reeves, W.C., 1997. Association of human papillomavirus type 16 integration in the E2 gene with poor disease-free survival from cervical cancer. *Int. J. Cancer* 74, 50–56.
- Vidal, J.P., Felix, S.P., Chaves, C.B., Patury, P., Franco, V.F., de Moraes, E.A., de Carvalho, N.A., Carvalho, A.C., Almeida Neto, O.F., Vieira, L.M., Correa, F.M., Martins, L.M., Negrão, L.F., de Almeida, A., Moreira, L.M., 2016. Genetic diversity of HPV16 and HPV18 in Brazilian patients with invasive cervical cancer. *J. Med. Virol.* 88, 1279–1287.
- Vinokurova, S., von Knebel Doeberitz, M., 2011. Differential methylation of the HPV 16 upstream regulatory region during epithelial differentiation and neoplastic transformation. *PLoS One* 6, e24451.
- Xu, B., Chotewutmontri, S., Wolf, S., Klos, U., Schmitz, M., Durst, M., Schwarz, M., 2013. Multiplex identification of human papillomavirus 16 DNA integration sites in cervical carcinomas. *PLoS One* 8, e66693.
- Xue, Y., Lim, D., Zhi, L., He, P., Abastado, J.P., Thierry, F., 2012. Loss of HPV16 E2 protein expression without disruption of the E2 ORF correlates with carcinogenic progression. *Open Virol. J.* 6, 163–172.
- Zampieri, M., Ciccarone, F., Calabrese, R., Franceschi, C., Bürkle, A., Caiata, P., 2015. Reconfiguration of DNA methylation in aging. *Mech. Ageing Dev.* 151, 60–70.

9.4. Anexo 4. Artigos publicados em colaboração.

Int J Clin Exp Pathol 2016;9(2):1520-1532
www.ijcep.com / ISSN:1936-2625 / IJCEP0019868

Original Article

***In situ* FoxP3+ and IL-10 over-expression is associated with high grade anal lesions in HIV infected patients**

Nathalia S Oliveira¹, Cecilia V Andrade², Beatriz Grinsztejn³, Ruth K Friedman³, Cynthia B Cunha³, Valdeia G Veloso³, Jose R Coutinho³, Janice M Chicarino-Oliveira³, Dennis De C Ferreira⁴, Luiza CR Pereira¹, Sergio M Amaro-Filho⁵, Miguel A Moreira⁶, Gerard J Nuovo⁶, Alcina F Nicol¹

¹Laboratory of Interdisciplinary Medical Research (LIPMED), IOC-Fiocruz, Rio de Janeiro, Brazil; ²Department of Pathology, Fernandes Figueira Institute-Fiocruz, Rio de Janeiro, Brazil; ³LapCin DST/AIDS-National Institute of Infectology-Fiocruz, Rio de Janeiro, Brazil; ⁴Department of Endodontics, Pos-graduate Program in Dentistry, Estácio de Sá University, Rio de Janeiro, RJ, Brazil; ⁵Division of Genetic, National Institute of Cancer, Rio de Janeiro, Brazil; ⁶Comprehensive Cancer Center, Ohio State University, Columbus, Ohio, USA

Received November 16, 2015; Accepted January 12, 2016; Epub February 1, 2016; Published February 15, 2016

Abstract: Human papillomavirus (HPV) is the main etiologic agent of lower genital tract cancers. The natural history of HPV infection and the immune response to HIV/ HPV co-infection, particularly in the anal mucosa, is poorly understood. The aim was evaluate the *in situ* immune response in anorectal biopsies from HIV-infected patients. A total of 114 biopsies were analyzed by Tissue Micro-Array: 15 were from HIV-negative individuals with normal squamous epithelium and 99 from HIV-positive individuals (21 normal squamous epithelium, 39 with anal intra-epithelial lesion grade 1 and 39 with anal intra-epithelial lesion grade 2/3). PCR and sequencing were used to identify HPV DNA. Staining for CD4, CD8, Foxp3+, Tbet and IL-10 were analyzed via immunohistochemistry. HIV-positive patients with AIN 2/3 showed a lower number of CD4+ cells (< 50 cells/mm³) compared to HIV negative subjects (P = 0.01). HIV infected individuals showed a higher expression of FoxP3+ and IL-10 that correlated with the severity of the lesion (P = 0.002). A positive coefficient correlation was found between FoxP3+ and IL-10 (r = 0.34; P = 0.027). HPV DNA was detected in 93.4% (101/107) of the samples and the most common types were HPV 16 (26.9%), followed by HPV 6 (15.7%), HPV 59 (13%) and HPV 18 (10.2%). Our results showed a strong association between the increased T-reg cells and IL-10 expression in HIV-positive patients with AIN 2/3. HPV 16 was the most prevalent type. Our study suggests that the immune regulatory *in situ* profile may favor HPV persistence in HIV-positive individuals. Further *in situ* studies should be done in order to elucidate the development of anal cancer in HPV/ HIV co-infected individuals.

Keywords: HPV/ HIV-1, anal intra epithelial neoplasia, immune response, genotype

Introduction

HIV-1 infected individuals have a higher incidence, prevalence and persistence of HPV infection. This is attributed to an increased susceptibility to infection and a decreased ability to eliminate HPV via the immune response, due to a deficiency of the CD4+ T cells [1].

Regulatory T cells (T-regs) can be responsible for mediating the tolerance of the immune response by regulating the Th1 and Th2 responses. One mechanism of T-regs-mediated immune tolerance is production of IL-10 and TGF- β 1. IL-10 can impair antigen presentation by dendritic cells (DC), and thus prevent T cell-

mediated immune response. IL-10 inhibits the inflammatory response and contributes to the development of malignant lesions in HIV-infected individuals by down regulation of expression of Th1 cytokines and triggering a Th2 cytokine pattern [2].

Prospective studies are required to better clarify the cellular immune response in HIV-1/ HPV co-infections, and other factors, that may contribute to progression to HPV-associated anal cancer. We hypothesize that HIV-induced expression of an inflammatory immune response favors HPV persistence and consequent development of anal intraepithelial lesions. The aim of the present study was to evaluate the im-

Original Article

Correlation of CD8 infiltration and expression of its checkpoint proteins PD-L1 and PD-L2 with the stage of cervical carcinoma

Maria CO Carvalho^{1*}, Alcina F Nicol^{2*}, Nathalia S Oliveira³, Cynthia B Cunha⁴, Sergio M Amaro-Filho⁵, Fabio B Russomano⁶, Elizabeth A Portari⁷, Gerard J Nuovo⁸

¹ Doctoral Program in Clinical Research on Infectious Diseases, INI-Fiocruz, Rio de Janeiro, Brazil; ² Laboratory of Interdisciplinary Medical Research (LIPMED), IOC-Fiocruz, Rio de Janeiro, Brazil; ³ LapCin DST/AIDS-National Institute of Infectology-INI-FIOCRUZ, Rio de Janeiro, Brazil; ⁴ Department of Pathology, Federal Fluminense University, Niteroi, Rio de Janeiro, Brazil; ⁵ Division of Genetic, National Institute of Cancer, Rio de Janeiro, Brazil; ⁶ Cervical Pathology and Colposcopy Unit - Instituto Nacional de Saúde da Mulher, da Criança e do Adolescente, Fernandes Figueira - FIOCRUZ, Rio de Janeiro, Brazil; ⁷ Department of Pathology, Fernandes Figueira Institute - FIOCRUZ, Rio de Janeiro, Brazil; ⁸ Comprehensive Cancer Center, Ohio State University, Columbus, Ohio, USA. * Equal contributors.

Received May 24, 2016; Accepted July 25, 2016; Epub October 1, 2016; Published October 15, 2016

Abstract: The importance of CD8 infiltration for cancer prognosis has been underscored lately by the increased use of checkpoint inhibitors in patients with invasive tumors. The objective of this study was to determine if CD8 infiltration and the expression of two key checkpoint proteins, PD-L1 and PD-L2, varied between early stage (FIGO IA-IIA) and advanced stage (FIGO IIB-IVA) cervical cancer. A tissue micro-array with 61 cervical specimens was analyzed through immunohistochemistry for PD-L1, PD-L2 and CD8 using CD1a (antigen presenting cells) as an internal control. Infiltration of both CD8 and CD1a was evident in control, benign tissues that showed little to no PD-L1 and PD-L2 reactivity. In samples of invasive cervical cancers, there was a three-fold increase in the number of CD8 cells with an increase in the expression of both PD-L1 and PD-L2 ($P < 0.001$ for each vs. control). A slight decrease in the number of CD1a cells was observed in malignant tissues compared to controls. No significant difference was evident for the infiltration of CD8⁺ cells or the expression of either PD-L1 and PD-L2 between the samples of early stage (FIGO IA-IIA) and advanced stage (FIGO IIB-IVA) cancers. Cytotoxic T cell infiltration and expression of two of its key checkpoint proteins (PD-L1 and PD-L2) remained constant as cervical cancers advance from early stage to late stage tumors. This suggests that the immune response may be equivalent in early and late stage cervical cancers.

Keywords: Invasive cancer, PD-L1, PD-L2, CD8, CD1a, HPV

Introduction

HPV infection is a prerequisite for the development of cervical cancer [1]. Given that cervical cancer is the classic example of a malignant tumor associated with a productive viral infection, it is not surprising that the host immune response, especially as it relates to the cytotoxic CD8-mediated response, has received much attention in the pathogenesis and evolution of both cervical intraepithelial neoplasia (CIN) and cervical cancer [2, 3]. The correlation of the CD8⁺ T cell infiltration and the regression of intraepithelial lesions have been documented by immunohistochemical studies [2, 3]. CD8⁺ cells are the dominant inflammatory cell type in cervical cancer and their activation by HPV associated vaccines can reduce tumor

growth [4, 5]. Much attention has been focused on the activation of CD8⁺ cells that are commonly found amongst invasive cancer cells. Two recent studies have highlighted that about one third of late stage cancers contain CD8⁺ cell infiltrates, yet these cytotoxic cells appear to be predominantly quiescent due to the concomitant expression of checkpoint proteins such as PD-1 and PD-L1 [6, 7]. Importantly, "unblockage" of CD8⁺ cell quiescence by inhibitors of these checkpoints have been observed to promote regression of metastatic lung, renal cell and colorectal cancers as well as melanomas [6]. However, little is known about the role of PD-L1 and PD-L2 in invasive cervical cancer.

Programmed cell death protein 1 (PD-1) is a transmembrane glycoprotein expressed on B

9.5. Anexo 5. Informações totais das amostras HPV16 positivas analisadas por pirosequenciamento

N	Paciente ID	Idade	Linhagem	Histologia	FIGO	Integridade E1/E2	Sítio CpG (% de Metilação)				
							31	37	43	52	58
1	5085293	37	D	ADN	IB1	Intacto	5	18	11	12	13
2	5094135	48	A	SCC	IVB	Intacto	25	22	23	25	10
3	5088433	44	A	SCC	IIIB	Com ruptura	5	6	6	7	8
4	5093188	35	A	SCC	IIIB	Com ruptura	35	35	39	33	25
5	5097835	43	A	ADN	IA1	Intacto	50	72	71	89	89
6	5086968	54	A	ADN	IIIB	Com ruptura	48	39	45	45	45
7	5116689	35	A	SCC	IIB	Intacto	3	5	4	6	5
8	5116386	46	A	ADN	IIB	Intacto	67	71	75	79	74
9	5117224	40	D	SCC	IIB	Intacto	2	3	3	4	3
10	5117165	35	A	SCC	IIIB	Com ruptura	3	5	4	5	5
11	5117179	57	A	SCC	IIIB	Intacto	4	65	68	73	66
12	5090224	41	D	ADN	IIB	Intacto	12	34	28	24	30
13	5090756	61	A	SCC	IB1	Com ruptura	2	2	0	2	2
14	5106670	68	D	SCC	IB2	Intacto	13	35	21	34	30
15	5111198	50	D	SCC	IIIB	Com ruptura	2	5	2	4	3
16	5104424	35	A	SCC	IIIB	Com ruptura	2	3	2	4	3
17	5105814	69	A	SCC	IIIB	Intacto	28	75	45	73	69
18	5094372	26	A	ADN	IB2	Com ruptura	2	3	2	4	2
19	5108135	47	C	ADN	IIIB	Com ruptura	3	3	3	4	4
20	5110544	37	D	ADN	IIIB	Intacto	25	42	32	42	38
21	5105770	48	D	ADN	IB1	Intacto	14	26	18	27	24
22	5085702	52	A	ADN	IB2	Intacto	35	36	33	35	32
23	5086436	36	A	ADN	IB2	Com ruptura	2	4	3	3	0
24	5087726	60	A	ADN	IB2	Intacto	50	51	51	48	43
25	5105071	49	D	ADN	IIB	Intacto	11	24	17	24	21
26	5105810	39	D	ADN	IIA	Com ruptura	6	7	8	6	5

27	5101433	70	D	ADN	IIA	Intacto	83	90	89	88	88
28	5091344	19	D	SCC	IIIB	Intacto	5	8	7	7	5
29	5087480	47	D	SCC	IIA1	Com ruptura	7	9	9	8	8
30	5090037	47	A	SCC	IIIB	Intacto	11	18	12	18	18
31	5100426	28	D	ADN	IB1	Intacto	2	2	2	3	4
32	5090713	35	A	SCC	IIIB	Intacto	49	39	43	44	39
33	5092267	39	D	SCC	IIB	Com ruptura	0	0	7	0	0
34	5088801	42	A	SCC	IIIB	Intacto	4	6	11	11	9
35	5079019	58	B	SCC	IIB	Com ruptura	2	3	2	2	2
36	5086785	22	B	SCC	IIB	Com ruptura	15	16	15	16	14
37	5091947	50	B	SCC	IIIB	Intacto	53	55	59	56	54
38	5104997	62	B	SCC	IIIB	Intacto	63	64	70	68	66
39	5118914	33	B	SCC	IB2	Intacto	10	11	12	12	13
40	5090104	43	C	SCC	IB1	Intacto	4	5	5	5	5
41	5099582	52	C	SCC	IIB	Intacto	14	19	16	21	14
42	5111752	36	C	SCC	IVA	Intacto	7	8	6	8	7
43	5117150	43	C	SCC	IB1	Com ruptura	1	3	3	3	4
44	5120241	62	C	SCC	IIIB	Intacto	76	74	79	75	79
45	5089877	41	C	SCC	IB1	Intacto	3	4	3	4	3
46	5120450	62	B	SCC	IB1	#	45	56	55	38	47
47	5088557	27	B	SCC	IIB	#	1	3	2	3	2
48	5106484	48	B	SCC	IB1	Intacto	27	26	28	28	26
49	5112065	32	C	SCC	IB1	Intacto	34	23	20	24	18
50	5094649	28	C	SCC	IIIB	Intacto	5	3	3	3	2
51	5101703	35	C	SCC	IIA	Intacto	3	4	3	4	4
52	5108696	31	B	SCC	IIIB	Com ruptura	1	2	2	3	2
53	5099777	55	A	ADN	IIIB	Intacto	2	2	2	2	3
54	5086234	66	A	ADN	IIB	Intacto	3	4	3	4	4
55	5123418	28	A	ADN	IB1	Intacto	47	50	53	51	52
56	5120437	50	D	ADN	IB2	Intacto	2	2	2	3	3
57	5121765	50	D	ADN	IVB	Intacto	9	9	9	10	10
58	5093119	39	D	SCC	IB2	Intacto	42	40	41	44	34

59	5121074	31	D	SCC	IIB	Com ruptura	2	2	2	2	2
60	5108163	37	D	SCC	IIB	Intacto	2	3	2	3	3
61	5121734	21	D	SCC	IIB	Intacto	68	69	72	70	71
62	5097577	79	A	ADN	IIB	Com ruptura	13	28	21	10	10
63	5106401	47	A	ADN	IB1	Com ruptura	11	12	13	11	11
64	5090471	37	A	SCC	IIIB	Com ruptura	7	7	16	8	9
65	5113612	59	A	SCC	IIIB	Intacto	5	5	8	5	7
66	5121046	28	D	ADN	IB2	Intacto	4	5	6	5	5
67	5124396	39	D	ADN	IIB	Intacto	32	34	42	37	24
68	5108552	50	D	SCC	IIB	Intacto	2	3	5	4	3
69	5113434	78	D	SCC	IIB	Intacto	45	50	54	52	47

#sem informação

9.6. Anexo 6. Informações totais das amostras HPV18 positivas analisadas por pirosequenciamento

N	Paciente ID	Idade	Linhagem	Histologia	FIGO	Integridade E1/E2	Sítio CpG (% de Metilação)				
							44	50	54	60	66
1	5107209	44	A	SCC	IIB	Com ruptura	3	4	6	0	0
2	5120016	55	#	SCC	IIIB	Com ruptura	3	5	6	4	0
3	5107040	62	A	SCC	IIIB	Com ruptura	3	4	2	6	3
4	5086003	43	B	SCC	IIIB	Com ruptura	7	9	8	16	9
5	5088120	46	B	Adeno	IIB	Com ruptura	5	6	7	9	9
6	5090668	56	A	SCC	IIB	Com ruptura	13	14	17	30	16
7	5108421	38	A	Adeno	IIB	Com ruptura	4	4	3	9	5
8	5091942	62	A	SCC	IIIB	Com ruptura	2	3	3	6	3
9	5094850	44	A	Adeno	IIB	Com ruptura	20	28	31	39	31
10	5095443	68	A	SCC	IB1	Com ruptura	12	8	10	21	14
11	5105554	42	A	SCC	IB2	intacto	9	11	13	21	12
12	5117413	47	A	SCC	IIIB	intacto	13	17	16	29	11
13	5117418	33	A	SCC	IIB	Com ruptura	4	4	10	12	6
14	5106684	52	A	SCC	IIB	Com ruptura	15	28	3	31	3
15	5095205	39	B	SCC	IIIB	Com ruptura	29	42	47	62	53
16	5108347	58	B	SCC	IIIB	Com ruptura	2	3	6	4	2
17	5094296	37	B	SCC	IIIB	Com ruptura	5	8	7	11	8
18	5116472	21	A	SCC	IIIB	Com ruptura	7	7	7	18	9
19	5116924	31	A	Adeno	IB2	Com ruptura	4	2	2	6	3
20	5119739	48	#	SCC	IIB	Com ruptura	4	5	0	4	1
21	5114518	48	A	SCC	IIIB	intacto	5	6	6	10	5
22	5117247	34	B	SCC	IIIB	Com ruptura	16	17	16	31	16
23	5120238	35	A	Adeno	IIIB	Com ruptura	1	1	0	4	1
24	5118718	42	#	Adeno	IIB	Com ruptura	2	2	3	3	1
25	5120334	33	B	SCC	IB2	Com ruptura	1	1	3	4	0
26	5112600	64	A	Adeno	IIIB	Com ruptura	3	4	3	6	4
27	5105904	39	A	SCC	IIB	Com ruptura	2	2	2	3	2
28	5102530	68	A	Adeno	IVB	intacto	30	45	50	71	56
29	5101826	50	A	SCC	IIA	intacto	9	10	17	14	5
30	5100265	75	A	SCC	IIIB	Com ruptura	2	3	4	5	2
31	5098111	45	A	#	IIIB	Com ruptura	17	17	26	43	31
32	5099630	45	A	SCC	IIIB	Com ruptura	2	3	2	3	2
33	5098055	55	A	Adeno	IIIB	Com ruptura	2	2	3	4	2

34	5120009	67	A	SCC	IIB	Com ruptura	2	2	1	4	3
35	5119823	64	A	#	IIB	Com ruptura	21	34	34	43	31
36	5118273	31	B	Adeno	IIB	Com ruptura	35	52	61	72	58
37	5106732	34	A	SCC	IIB	Com ruptura	2	3	3	4	2

#sem informação

9.7. Anexo 7. Informações totais das amostras HPV16 positivas analisadas por NGS para os dados de metilação do ensaio HPV16_L2.

N	Paciente ID	Idade	Linhagem	Histologia	Grau	Figo	Integridade E1/E2	Carga Viral	Sítio CpG (% de Metilação em L2 de HPV16)				
									4240	4249	4261	4270	4277
1	5085528	58	A	SCC	G2	IIIB	Intacto	2632	68	28	49	39	53
2	5085618	32	B	SCC	G2	IB2	Com ruptura	#	79	40	68	52	73
3	5087519	43	D	SCC	#	IIB	Intacto	89502	81	35	66	55	71
4	5087843	60	A	SCC	G2	IIB	#	#	36	47	69	56	73
5	5089148	43	D	SCC	G3	IIIB	Com ruptura	4751	93	80	90	86	90
6	5089753	58	A	SCC	#	IVA	Com ruptura	34287	59	28	42	37	51
7	5090037	47	A	SCC	G2	IIIB	Intacto	220213	40	17	26	23	27
8	5091250	31	D	SCC	G3	IIB	Intacto	7073	80	40	66	53	65
9	5095013	49	D	SCC	G2	IB1	Com ruptura	#	78	52	74	61	72
10	5095089	73	D	SCC	G2	IIB	Com ruptura	#	88	46	55	26	72
11	5096998	53	#	SCC	G2	IIIB	Com ruptura	#	34	11	19	16	21
12	5098552	45	D	SCC	G2	IIB	Com ruptura	#	63	24	41	34	51
13	5099629	49	A	SCC	#	IIIB	Com ruptura	#	92	47	82	68	78
14	5099665	71	A	SCC	#	IIIB	Intacto	2864548	60	24	42	32	43
15	5100510	41	A	SCC	G2	IIIB	Com ruptura	#	90	19	86	82	87
16	5101056	23	A	SCC	G3	IB2	Com ruptura	#	91	67	85	73	82
17	5102950	63	A	SCC	G2	IB2	Com ruptura	787	94	53	91	87	91
18	5103120	81	D	SCC	G3	IVB	Com ruptura	12517	88	60	81	75	82
19	5103143	29	A	SCC	G2	IB2	Intacto	26298	57	28	45	37	47
20	5104308	40	D	SCC	G2	IIIB	Intacto	128516	79	39	72	56	75
21	5104487	26	A	SCC	G2	IVA	Com ruptura	#	22	10	16	14	18
22	5104997	62	B	SCC	#	IIIB	Intacto	1300	84	57	80	66	80
23	5105685	65	A	SCC	G2	IIIB	Intacto	6322	80	49	68	57	71
24	5108502	28	D	SCC	G3	IIB	Intacto	368	88	58	86	70	84
25	5109849	38	A	SCC	G2	IIB	Intacto	138	68	38	59	47	61

26	5110094	20	A	SCC	G3	IB2	Intacto	426616	16	7	12	10	13
27	5111370	49	A	SCC	G2	IB2	Com ruptura	407	64	14	60	57	60
28	5111412	52	A	SCC	G2	IIIB	Com ruptura	260	77	42	65	55	63
29	5111994	54	D	SCC	#	IVA	Intacto	#	36	15	23	19	27
30	5113607	45	D	SCC	G3	IIB	Intacto	137	64	23	47	35	49
31	5114609	73	A	SCC	G2	IIB	Intacto	150061	51	22	35	27	37
32	5116689	35	A	SCC	#	IIB	Intacto	21159	93	79	92	86	93
33	5117591	43	A	SCC	G2	IB1	Intacto	#	90	78	86	81	87
34	5119407	53	D	SCC	G2	IIB	Intacto	5100	33	18	26	23	27
35	5119829	51	D	#	G3	IIIB	Intacto	#	90	61	90	71	89
36	5120284	68	#	SCC	G2	IB2	Intacto	#	93	81	91	87	92
37	5120450	62	B	SCC	G2	IB1	#	#	11	4	7	6	9
38	5120475	57	A	SCC	G3	IIB	Intacto	#	23	11	17	14	18
39	5123497	32	A	SCC	G2	IIIB	Com ruptura	#	85	72	84	80	84
40	5085293	37	D	Adeno	G2	IB1	Intacto	#	60	26	48	39	62
41	5085702	52	A	Adeno	G2	IB2	Intacto	3547	94	65	90	85	93
42	5085887	54	A	Adeno	G2	IIB	Intacto	1214	86	39	77	58	79
43	5087726	60	A	Adeno	G2	IB2	Intacto	1752	90	72	87	81	89
44	5090224	41	D	Adeno	G2	IIB	Intacto	55720	51	15	32	25	45
45	5091224	59	A	Adeno	#	IA1	Com ruptura	#	70	30	53	46	62
46	5095648	49	A	Adeno	G2	IB2	Intacto	1124	75	21	58	49	71
47	5095750	42	A	Adeno	G1	IB2	Intacto	25	94	84	93	90	93
48	5097577	79	A	Adeno	G3	IIB	#	#	36	64	91	79	83
49	5097835	43	A	Adeno	G2	IA1	Intacto	#	87	86	93	84	87
50	5099777	55	A	Adeno	G3	IIIB	Intacto	1203	68	31	58	47	64
51	5100426	28	D	Adeno	G2	IB1	Intacto	#	41	20	29	24	30
52	5101433	70	D	Adeno	G2	IIA	Intacto	#	86	65	66	68	87
53	5105071	49	D	Adeno	G1	IIB	Intacto	13409	29	15	22	20	28
54	5105770	48	D	Adeno	G2	IB1	Intacto	507661983	53	22	36	30	46
55	5110544	37	D	Adeno	G2	IIIB	Intacto	1262	79	30	56	47	78
56	5111181	44	A	Adeno	G2	IVB	Com ruptura	#	75	26	51	45	79
57	5112080	62	A	#	G3	IIB	Intacto	#	71	27	56	40	59

58	5113742	45	A	Adeno	G2	IIB	Intacto	550430	52	20	35	27	44
59	5114664	52	A	Adeno	G2	IIB	Intacto	#	77	37	62	51	74
60	5116386	46	A	Adeno	G2	IIB	Intacto	70130	92	76	92	86	92
61	5120437	50	D	Adeno	G2	IB2	Intacto	880	47	14	24	21	35
62	5121046	28	D	Adeno	G1	IB2	Intacto	93	67	28	55	43	67
63	5121765	50	D	#	G3	IVB	Com ruptura	#	42	18	31	25	42
64	5123418	28	A	Adeno	G1	IB1	Intacto	279	94	82	92	89	92
65	5123749	46	D	Adeno	G2	IIA1	Intacto	387	49	17	32	25	41
66	5124396	39	D	Adeno	G3	IIB	Intacto	1679	83	39	70	51	74
67	5094284	57	#	SCC	G2	IIIB	Com ruptura	#	#	#	#	#	#
68	5095309	55	#	SCC	G2	IIB	Com ruptura	#	#	#	#	#	#
69	5111836	32	A	SCC	G2	IB2	Com ruptura	#	#	#	#	#	#
70	5121031	45	D	SCC	G2	IIB	Com ruptura	#	#	#	#	#	#
71	5123721	39	D	SCC	G2	IIA	Intacto	910	#	#	#	#	#

#sem informação; Carga viral em 10⁶ células

9.8. Anexo 8. Informações totais das amostras HPV16 positivas analisadas por NGS para os dados de metilação do ensaio HPV16_L1.

N	Paciente ID	Idade	Linhagem	Histologia	Grau	Figo	Integridade E1/E2	Carga Viral	Sítio CpG (% de Metilação em L1 de HPV16)			
									5602	5608	5611	5617
1	5085528	58	A	SCC	G2	IIIB	Intacto	2632	71	27	61	29
2	5085618	32	B	SCC	G2	IB2	Com ruptura	#	81	34	76	41
3	5087519	43	D	SCC	#	IIB	Intacto	89502	90	44	86	46
4	5087843	60	A	SCC	G2	IIB	#	#	94	79	94	75
5	5089148	43	D	SCC	G3	IIIB	Com ruptura	4751	96	83	95	77
6	5089753	58	A	SCC	#	IVA	Com ruptura	34287	74	34	63	33
7	5090037	47	A	SCC	G2	IIIB	Intacto	220213	42	16	32	16
8	5091250	31	D	SCC	G3	IIB	Intacto	7073	89	45	85	50
9	5095013	49	D	SCC	G2	IB1	Com ruptura	#	80	67	79	67
10	5095089	73	D	SCC	G2	IIB	Com ruptura	#	91	55	90	38
11	5096998	53	#	SCC	G2	IIIB	Com ruptura	#	36	11	28	12
12	5098552	45	D	SCC	G2	IIB	Com ruptura	#	77	26	69	27
13	5099629	49	A	SCC	#	IIIB	Com ruptura	#	93	63	91	57
14	5099665	71	A	SCC	#	IIIB	Intacto	2864548	59	20	48	21
15	5100510	41	A	SCC	G2	IIIB	Com ruptura	#	85	58	84	61
16	5101056	23	A	SCC	G3	IB2	Com ruptura	#	97	82	94	76
17	5102950	63	A	SCC	G2	IB2	Com ruptura	787	96	86	96	90
18	5103120	81	D	SCC	G3	IVB	Com ruptura	12517	88	43	83	49
19	5103143	29	A	SCC	G2	IB2	Intacto	26298	59	28	52	30
20	5104308	40	D	SCC	G2	IIIB	Intacto	128516	95	86	95	85
21	5104487	26	A	SCC	G2	IVA	Com ruptura	#	94	40	82	41
22	5104997	62	B	SCC	#	IIIB	Intacto	1300	92	71	91	73
23	5105685	65	A	SCC	G2	IIIB	Intacto	6322	84	58	82	59
24	5108502	28	D	SCC	G3	IIB	Intacto	368	90	67	91	72
25	5109849	38	A	SCC	G2	IIB	Intacto	138	71	45	70	50

26	5110094	20	A	SCC	G3	IB2	Intacto	426616	95	82	94	81
27	5111370	49	A	SCC	G2	IB2	Com ruptura	407	73	40	92	40
28	5111412	52	A	SCC	G2	IIIB	Com ruptura	260	76	56	74	53
29	5111994	54	D	SCC	#	IVA	Intacto	#	54	16	47	16
30	5113607	45	D	SCC	G3	IIB	Intacto	137	65	24	55	27
31	5114609	73	A	SCC	G2	IIB	Intacto	150061	58	27	49	27
32	5116689	35	A	SCC	#	IIB	Intacto	21159	94	82	92	78
33	5117591	43	A	SCC	G2	IB1	Intacto	#	94	85	92	84
34	5119407	53	D	SCC	G2	IIB	Intacto	5100	50	21	42	21
35	5119829	51	D	#	G3	IIIB	Intacto	#	84	56	88	69
36	5120284	68	#	SCC	G2	IB2	Intacto	#	94	83	93	81
37	5120450	62	B	SCC	G2	IB1	#	#	21	10	17	10
38	5120475	57	A	SCC	G3	IIB	Intacto	#	94	86	93	84
39	5123497	32	A	SCC	G2	IIIB	Com ruptura	#	92	81	90	81
40	5085293	37	D	Adeno	G2	IB1	Intacto	#	59	25	50	23
41	5085702	52	A	Adeno	G2	IB2	Intacto	3547	94	80	92	79
42	5085887	54	A	Adeno	G2	IIB	Intacto	1214	86	37	82	45
43	5087726	60	A	Adeno	G2	IB2	Intacto	1752	90	76	89	77
44	5090224	41	D	Adeno	G2	IIB	Intacto	55720	53	18	41	19
45	5091224	59	A	Adeno	#	IA1	Com ruptura	#	76	31	64	32
46	5095648	49	A	Adeno	G2	IB2	Intacto	1124	81	31	72	34
47	5095750	42	A	Adeno	G1	IB2	Intacto	25	96	89	94	89
48	5097577	79	A	Adeno	G3	IIB	#	#	89	59	88	67
49	5097835	43	A	Adeno	G2	IA1	Intacto	#	89	73	91	78
50	5099777	55	A	Adeno	G3	IIIB	Intacto	1203	71	30	61	31
51	5100426	28	D	Adeno	G2	IB1	Intacto	#	43	18	35	18
52	5101433	70	D	Adeno	G2	IIA	Intacto	#	80	44	90	66
53	5105071	49	D	Adeno	G1	IIB	Intacto	13409	34	16	26	16
54	5105770	48	D	Adeno	G2	IB1	Intacto	507661983	46	19	35	18
55	5110544	37	D	Adeno	G2	IIIB	Intacto	1262	80	36	74	39
56	5111181	44	A	Adeno	G2	IVB	Com ruptura	#	75	13	59	14
57	5112080	62	A	#	G3	IIB	Intacto	#	78	32	70	35

58	5113742	45	A	Adeno	G2	IIB	Intacto	550430	65	27	55	27
59	5114664	52	A	Adeno	G2	IIB	Intacto	#	75	32	66	34
60	5116386	46	A	Adeno	G2	IIB	Intacto	70130	#	#	#	#
61	5120437	50	D	Adeno	G2	IB2	Intacto	880	51	16	37	17
62	5121046	28	D	Adeno	G1	IB2	Intacto	93	68	25	55	26
63	5121765	50	D	#	G3	IVB	Com ruptura	#	45	21	38	21
64	5123418	28	A	Adeno	G1	IB1	Intacto	279	93	84	92	85
65	5123749	46	D	Adeno	G2	IIA1	Intacto	387	59	21	47	20
66	5124396	39	D	Adeno	G3	IIB	Intacto	1679	86	52	81	53
67	5094284	57	#	SCC	G2	IIIB	Com ruptura	#	63	38	47	14
68	5095309	55	#	SCC	G2	IIB	Com ruptura	#	64	33	58	31
69	5111836	32	A	SCC	G2	IB2	Com ruptura	#	16	6	14	7
70	5121031	45	D	SCC	G2	IIB	Com ruptura	#	93	80	92	76
71	5123721	39	D	SCC	G2	IIA	Intacto	910	94	77	92	79

#sem informação; Carga viral em 10⁶ células

9.9. Anexo 9. Informações totais das amostras HPV16 positivas analisadas por NGS para os dados de metilação do ensaio HPV16_LCR.

N	Paciente ID	Idade	Linhagem	Histologia	Grau	Figo	Integridade E1/E2	Carga Viral	Sítio CpG (% de Metilação em LCR de HPV16)					
									7862	31	37	43	52	58
1	5085528	58	A	SCC	G2	IIIB	Intacto	2632	11	30	30	31	32	30
2	5085618	32	B	SCC	G2	IB2	Com ruptura	#	9	41	36	37	36	34
3	5087519	43	D	SCC	#	IIB	Intacto	89502	5	41	39	39	39	31
4	5087843	60	A	SCC	G2	IIB	#	#	11	24	25	26	26	22
5	5089148	43	D	SCC	G3	IIIB	Com ruptura	4751	6	9	11	12	11	11
6	5089753	58	A	SCC	#	IVA	Com ruptura	34287	18	25	30	32	31	28
7	5090037	47	A	SCC	G2	IIIB	Intacto	220213	19	21	24	26	27	25
8	5091250	31	D	SCC	G3	IIB	Intacto	7073	26	22	25	28	29	27
9	5095013	49	D	SCC	G2	IB1	Com ruptura	#	9	47	51	50	50	34
10	5095089	73	D	SCC	G2	IIB	Com ruptura	#	60	67	69	71	73	69
11	5096998	53	#	SCC	G2	IIIB	Com ruptura	#	9	14	22	23	16	15
12	5098552	45	D	SCC	G2	IIB	Com ruptura	#	14	24	29	29	27	26
13	5099629	49	A	SCC	#	IIIB	Com ruptura	#	5	9	12	11	12	11
14	5099665	71	A	SCC	#	IIIB	Intacto	2864548	12	16	21	22	17	20
15	5100510	41	A	SCC	G2	IIIB	Com ruptura	#	6	46	47	47	46	39
16	5101056	23	A	SCC	G3	IB2	Com ruptura	#	20	24	25	25	28	23
17	5102950	63	A	SCC	G2	IB2	Com ruptura	787	4	28	29	28	30	24
18	5103120	81	D	SCC	G3	IVB	Com ruptura	12517	4	6	9	10	11	8
19	5103143	29	A	SCC	G2	IB2	Intacto	26298	16	29	32	32	33	32
20	5104308	40	D	SCC	G2	IIIB	Intacto	128516	4	10	11	11	11	10
21	5104487	26	A	SCC	G2	IVA	Com ruptura	#	3	11	13	13	12	11
22	5104997	62	B	SCC	#	IIIB	Intacto	1300	7	56	58	59	59	57
23	5105685	65	A	SCC	G2	IIIB	Intacto	6322	16	39	43	44	42	40
24	5108502	28	D	SCC	G3	IIB	Intacto	368	7	62	61	65	64	64
25	5109849	38	A	SCC	G2	IIB	Intacto	138	15	49	50	51	51	49

26	5110094	20	A	SCC	G3	IB2	Intacto	426616	5	12	42	42	45	40
27	5111370	49	A	SCC	G2	IB2	Com ruptura	407	3	14	13	12	12	12
28	5111412	52	A	SCC	G2	IIIB	Com ruptura	260	3	16	15	15	16	15
29	5111994	54	D	SCC	#	IVA	Intacto	#	15	16	20	21	19	19
30	5113607	45	D	SCC	G3	IIB	Intacto	137	7	17	17	18	17	17
31	5114609	73	A	SCC	G2	IIB	Intacto	150061	18	32	35	37	35	33
32	5116689	35	A	SCC	#	IIB	Intacto	21159	6	11	12	12	12	11
33	5117591	43	A	SCC	G2	IB1	Intacto	#	23	73	73	75	74	73
34	5119407	53	D	SCC	G2	IIB	Intacto	5100	22	26	31	32	31	29
35	5119829	51	D	#	G3	IIIB	Intacto	#	10	75	66	80	74	78
36	5120284	68	#	SCC	G2	IB2	Intacto	#	#	#	#	#	#	#
37	5120450	62	B	SCC	G2	IB1	#	#	#	#	#	#	#	#
38	5120475	57	A	SCC	G3	IIB	Intacto	#	11	14	16	16	16	17
39	5123497	32	A	SCC	G2	IIIB	Com ruptura	#	20	68	68	70	69	68
40	5085293	37	D	Adeno	G2	IB1	Intacto	#	23	24	29	31	28	28
41	5085702	52	A	Adeno	G2	IB2	Intacto	3547	7	33	32	32	33	31
42	5085887	54	A	Adeno	G2	IIB	Intacto	1214	20	33	34	35	34	32
43	5087726	60	A	Adeno	G2	IB2	Intacto	1752	10	42	43	42	41	37
44	5090224	41	D	Adeno	G2	IIB	Intacto	55720	33	29	38	39	36	33
45	5091224	59	A	Adeno	#	IA1	Com ruptura	#	31	33	42	44	40	38
46	5095648	49	A	Adeno	G2	IB2	Intacto	1124	40	11	39	37	28	40
47	5095750	42	A	Adeno	G1	IB2	Intacto	25	4	33	32	32	34	29
48	5097577	79	A	Adeno	G3	IIB	#	#	#	#	#	#	#	#
49	5097835	43	A	Adeno	G2	IA1	Intacto	#	6	42	46	50	48	44
50	5099777	55	A	Adeno	G3	IIIB	Intacto	1203	40	35	43	45	42	39
51	5100426	28	D	Adeno	G2	IB1	Intacto	#	31	30	35	39	34	34
52	5101433	70	D	Adeno	G2	IIA	Intacto	#	19	67	68	70	69	66
53	5105071	49	D	Adeno	G1	IIB	Intacto	13409	39	33	42	45	41	38
54	5105770	48	D	Adeno	G2	IB1	Intacto	507661983	39	38	46	49	45	43
55	5110544	37	D	Adeno	G2	IIIB	Intacto	1262	13	26	27	29	28	26
56	5111181	44	A	Adeno	G2	IVB	Com ruptura	#	4	18	12	19	18	12
57	5112080	62	A	#	G3	IIB	Intacto	#	15	25	27	28	28	27

58	5113742	45	A	Adeno	G2	IIB	Intacto	550430	40	37	46	48	43	44
59	5114664	52	A	Adeno	G2	IIB	Intacto	#	23	24	29	29	30	28
60	5116386	46	A	Adeno	G2	IIB	Intacto	70130	6	50	52	53	53	49
61	5120437	50	D	Adeno	G2	IB2	Intacto	880	13	20	19	20	20	21
62	5121046	28	D	Adeno	G1	IB2	Intacto	93	20	24	28	29	28	27
63	5121765	50	D	#	G3	IVB	Com ruptura	#	16	24	25	27	27	26
64	5123418	28	A	Adeno	G1	IB1	Intacto	279	19	45	47	47	48	46
65	5123749	46	D	Adeno	G2	IIA1	Intacto	387	20	26	28	30	30	28
66	5124396	39	D	Adeno	G3	IIB	Intacto	1679	7	35	36	41	39	36
67	5094284	57	#	SCC	G2	IIIB	Com ruptura	#	#	#	#	#	#	#
68	5095309	55	#	SCC	G2	IIB	Com ruptura	#	8	76	77	70	71	76
69	5111836	32	A	SCC	G2	IB2	Com ruptura	#	6	12	12	11	10	12
70	5121031	45	D	SCC	G2	IIB	Com ruptura	#	3	11	10	10	10	10
71	5123721	39	D	SCC	G2	IIA	Intacto	910	30	46	50	50	49	46

#sem informação; Carga viral em 10⁶ células

9.10. Anexo 10. Informações totais das amostras HPV16 positivas analisadas por NGS para os dados de metilação do ensaio HPV18_L2.

N	Paciente ID	Idade	Linhagem	Histologia	Grau	Figo	Integridade E1/E2	Sítio CpG (% de Metilação em L2 de HPV18)			
								5331	5333	5367	5406
1	5088120	46	B	Adeno	G1	IIB	Com ruptura	95	95	95	93
2	5089012	45	A	Adeno	#	IIB	Com ruptura	62	56	65	59
3	5089419	57	A	Adeno	G1	IIB	Com ruptura	95	89	95	94
4	5094123	32	#	Adeno	G1	IB1	#	92	78	95	88
5	5095990	53	A	Adeno	G3	IIIB	Com ruptura	84	81	90	80
6	5096890	38	A	Adeno	G1	IB1	Intacto	86	80	93	84
7	5097594	38	A	Adeno	G2	IB1	Intacto	51	46	70	52
8	5098055	55	A	Adeno	G3	IIIB	Com ruptura	95	95	95	94
9	5101451	43	A	Adeno	G1	IB1	Com ruptura	62	60	70	58
10	5102558	73	A	Adeno	G2	IIB	Com ruptura	96	95	96	94
11	5112600	64	A	Adeno	G1	IIIB	Com ruptura	96	95	96	94
12	5114003	37	A	Adeno	G1	IB1	#	96	95	96	94
13	5114837	50	A	Adeno	G2	IIIB	#	71	67	81	70
14	5116924	31	A	Adeno	G2	IB2	Com ruptura	26	24	38	29
15	5118273	31	B	Adeno	G3	IIB	Com ruptura	95	94	93	91
16	5118718	42	#	Adeno	G3	IIB	Com ruptura	95	95	95	93
17	5120238	35	A	Adeno	#	IIIB	Com ruptura	94	94	94	93
18	5122170	46	A	Adeno	G3	IIB	#	57	56	68	57
19	5085161	55	A	SCC	G3	IB1	Com ruptura	95	94	95	94
20	5087961	46	A	SCC	G3	IIIB	Com ruptura	95	95	95	93
21	5088177	38	B	SCC	G3	IIB	Com ruptura	81	78	89	77
22	5093245	38	A	SCC	G2	IIIB	#	95	94	95	93
23	5094147	32	A	SCC	G2	IIB	#	96	95	95	93
24	5094296	37	B	SCC	G2	IIIB	Com ruptura	72	67	64	58

25	5095205	39	B	SCC	G2	IIIB	Com ruptura	59	55	66	42
26	5096435	48	A	SCC	G3	IIIB	Com ruptura	94	92	92	90
27	5097002	58	A	SCC	G2	IVA	Intacto	15	14	18	15
28	5098561	36	#	SCC	G2	IB2	Intacto	94	94	95	93
29	5098970	41	A	SCC	G3	IB1	#	74	70	82	73
30	5099630	45	A	SCC	#	IIIB	Com ruptura	93	91	95	92
31	5099856	40	A	SCC	G2	IIIB	Com ruptura	95	95	95	94
32	5100265	75	A	SCC	#	IIIB	Com ruptura	42	41	44	42
33	5103540	58	A	SCC	G2	IIIB	Intacto	94	92	95	90
34	5105554	42	A	SCC	G3	IB2	Intacto	93	92	93	93
35	5106684	52	A	SCC	G2	IIB	Com ruptura	25	38	24	23
36	5106732	34	A	SCC	G3	IIB	Com ruptura	95	94	94	92
37	5107040	62	A	SCC	G2	IIIB	Com ruptura	42	41	48	44
38	5107209	44	A	SCC	G2	IIB	Com ruptura	96	95	96	94
39	5114518	48	A	SCC	G2	IIIB	Intacto	74	69	65	67
40	5115724	35	A	SCC	G2	IIB	Com ruptura	95	94	94	92
41	5116472	21	A	SCC	G2	IIIB	Com ruptura	93	92	94	91
42	5116780	44	A	SCC	G2	IIIB	#	95	94	95	93
43	5117413	47	A	SCC	G2	IIIB	Intacto	93	90	91	73
44	5119739	48	#	SCC	G2	IIB	Com ruptura	97	97	96	96
45	5120009	67	A	SCC	G2	IIB	#	96	93	97	94
46	5122632	43	A	SCC	G3	IB2	Com ruptura	98	97	97	96
47	5122818	52	B	SCC	G2	IIIB	Com ruptura	98	97	98	97
48	5104549	44	A	SCC	G2	IIIB	#	#	#	#	#
49	5112764	48	#	SCC	G2	IIIB	Com ruptura	#	#	#	#

#sem informação; Carga viral em 10⁶ células

9.11. Anexo 11. Informações totais das amostras HPV16 positivas analisadas por NGS para os dados de metilação do ensaio HPV18_L1.

N	Paciente ID	Idade	Linhagem	Histologia	Grau	Figo	Integridade E1/E2	Sítio CpG (% de Metilação em L1 de HPV18)								
								7011	7038	7041	7062	7068	7090	7110	7116	7122
1	5088120	46	B	Adeno	G1	IIB	Com ruptura	67	84	88	93	69	93	95	93	89
2	5089012	45	A	Adeno	#	IIB	Com ruptura	34	50	51	54	37	52	59	50	42
3	5089419	57	A	Adeno	G1	IIB	Com ruptura	83	95	95	96	86	96	97	96	93
4	5094123	32	#	Adeno	G1	IB1	#	86	95	94	95	86	94	97	96	93
5	5095990	53	A	Adeno	G3	IIIB	Com ruptura	68	77	78	79	65	79	83	79	75
6	5096890	38	A	Adeno	G1	IB1	Intacto	83	89	88	89	77	93	94	94	89
7	5097594	38	A	Adeno	G2	IB1	Intacto	25	48	48	54	40	64	72	60	46
8	5098055	55	A	Adeno	G3	IIIB	Com ruptura	77	90	92	94	69	94	97	95	89
9	5101451	43	A	Adeno	G1	IB1	Com ruptura	40	57	58	59	44	57	63	57	51
10	5102558	73	A	Adeno	G2	IIB	Com ruptura	#	#	#	#	#	#	#	#	#
11	5112600	64	A	Adeno	G1	IIIB	Com ruptura	80	93	94	95	78	94	96	94	90
12	5114003	37	A	Adeno	G1	IB1	#	87	95	95	95	82	95	97	96	93
13	5114837	50	A	Adeno	G2	IIIB	#	40	56	60	56	37	54	61	52	40
14	5116924	31	A	Adeno	G2	IB2	Com ruptura	35	45	46	48	34	54	60	52	46
15	5118273	31	B	Adeno	G3	IIB	Com ruptura	88	95	95	96	88	95	97	96	94
16	5118718	42	#	Adeno	G3	IIB	Com ruptura	60	89	92	91	68	91	94	91	85
17	5120238	35	A	Adeno	#	IIIB	Com ruptura	79	94	95	93	79	95	97	96	88
18	5122170	46	A	Adeno	G3	IIB	#	45	65	65	64	52	64	66	63	57
19	5085161	55	A	SCC	G3	IB1	Com ruptura	51	86	88	90	76	88	95	92	83
20	5087961	46	A	SCC	G3	IIIB	Com ruptura	66	91	91	92	70	75	90	82	66
21	5088177	38	B	SCC	G3	IIB	Com ruptura	45	80	83	85	55	83	86	77	66
22	5093245	38	A	SCC	G2	IIIB	#	84	94	94	94	75	94	96	93	90
23	5094147	32	A	SCC	G2	IIB	#	67	91	93	94	67	92	96	92	84

24	5094296	37	B	SCC	G2	IIIB	Com ruptura	18	33	32	42	23	43	52	41	31
25	5095205	39	B	SCC	G2	IIIB	Com ruptura	60	76	77	77	66	75	80	77	71
26	5096435	48	A	SCC	G3	IIIB	Com ruptura	70	86	84	88	72	88	92	90	85
27	5097002	58	A	SCC	G2	IVA	Intacto	20	37	35	38	20	34	37	30	23
28	5098561	36	#	SCC	G2	IB2	Intacto	88	95	95	95	87	95	97	96	94
29	5098970	41	A	SCC	G3	IB1	#	30	62	62	78	52	79	82	75	60
30	5099630	45	A	SCC	#	IIIB	Com ruptura	68	87	87	88	62	85	89	83	70
31	5099856	40	A	SCC	G2	IIIB	Com ruptura	65	90	92	91	60	86	90	82	68
32	5100265	75	A	SCC	#	IIIB	Com ruptura	39	49	50	52	43	52	56	52	48
33	5103540	58	A	SCC	G2	IIIB	Intacto	78	88	89	90	73	89	91	87	81
34	5105554	42	A	SCC	G3	IB2	Intacto	63	94	94	95	87	95	96	95	92
35	5106684	52	A	SCC	G2	IIB	Com ruptura	#	#	#	#	#	#	#	#	#
36	5106732	34	A	SCC	G3	IIB	Com ruptura	83	95	96	96	88	96	97	96	93
37	5107040	62	A	SCC	G2	IIIB	Com ruptura	38	47	49	51	38	53	56	51	48
38	5107209	44	A	SCC	G2	IIB	Com ruptura	#	#	#	#	#	#	#	#	#
39	5114518	48	A	SCC	G2	IIIB	Intacto	19	29	29	45	24	57	62	52	39
40	5115724	35	A	SCC	G2	IIB	Com ruptura	55	69	71	71	58	71	76	70	65
41	5116472	21	A	SCC	G2	IIIB	Com ruptura	80	93	93	93	75	93	95	94	89
42	5116780	44	A	SCC	G2	IIIB	#	#	#	#	#	#	#	#	#	#
43	5117413	47	A	SCC	G2	IIIB	Intacto	76	86	89	91	75	88	92	92	89
44	5119739	48	#	SCC	G2	IIB	Com ruptura	58	77	77	76	50	71	78	70	57
45	5120009	67	A	SCC	G2	IIB	#	83	95	96	97	88	95	97	96	94
46	5122632	43	A	SCC	G3	IB2	Com ruptura	89	94	95	96	85	95	96	96	92
47	5122818	52	B	SCC	G2	IIIB	Com ruptura	91	95	96	97	83	95	97	96	92
48	5104549	44	A	SCC	G2	IIIB	#	50	74	75	80	50	79	81	76	66
49	5112764	48	#	SCC	G2	IIIB	Com ruptura	54	85	87	88	57	83	89	80	65

#sem informação; Carga viral em 10⁶ células

9.12. Anexo 12. Informações totais das amostras HPV16 positivas analisadas por NGS para os dados de metilação do ensaio HPV18_LCR.

N	Paciente ID	Idade	Linhagem	Histologia	Grau	Figo	Integridade E1/E2	Sítio CpG (% de Metilação em LCR de HPV18)					
								7824	44	50	54	60	66
1	5088120	46	B	Adeno	G1	IIB	Com ruptura	4	7	10	5	6	6
2	5089012	45	A	Adeno	#	IIB	Com ruptura	10	10	12	9	13	11
3	5089419	57	A	Adeno	G1	IIB	Com ruptura	8	12	10	10	12	17
4	5094123	32	#	Adeno	G1	IB1	#	#	#	#	#	#	#
5	5095990	53	A	Adeno	G3	IIIB	Com ruptura	9	20	14	13	19	17
6	5096890	38	A	Adeno	G1	IB1	Intacto	15	31	20	19	26	26
7	5097594	38	A	Adeno	G2	IB1	Intacto	29	33	33	31	34	33
8	5098055	55	A	Adeno	G3	IIIB	Com ruptura	5	4	5	4	5	5
9	5101451	43	A	Adeno	G1	IB1	Com ruptura	14	8	8	7	8	10
10	5102558	73	A	Adeno	G2	IIB	Com ruptura	5	40	39	6	9	10
11	5112600	64	A	Adeno	G1	IIIB	Com ruptura	5	7	5	4	4	4
12	5114003	37	A	Adeno	G1	IB1	#	8	8	6	6	7	10
13	5114837	50	A	Adeno	G2	IIIB	#	4	4	4	5	9	6
14	5116924	31	A	Adeno	G2	IB2	Com ruptura	7	5	7	4	5	4
15	5118273	31	B	Adeno	G3	IIB	Com ruptura	34	53	41	42	51	49
16	5118718	42	#	Adeno	G3	IIB	Com ruptura	4	4	4	2	5	4
17	5120238	35	A	Adeno	#	IIIB	Com ruptura	4	4	32	4	4	4
18	5122170	46	A	Adeno	G3	IIB	#	3	3	3	2	4	3
19	5085161	55	A	SCC	G3	IB1	Com ruptura	4	4	3	2	2	4
20	5087961	46	A	SCC	G3	IIIB	Com ruptura	9	5	5	4	5	6
21	5088177	38	B	SCC	G3	IIB	Com ruptura	4	4	5	2	4	4
22	5093245	38	A	SCC	G2	IIIB	#	8	12	9	4	6	5
23	5094147	32	A	SCC	G2	IIB	#	#	#	#	#	#	#
24	5094296	37	B	SCC	G2	IIIB	Com ruptura	5	8	7	6	9	7

25	5095205	39	B	SCC	G2	IIIB	Com ruptura	14	42	34	33	43	44
26	5096435	48	A	SCC	G3	IIIB	Com ruptura	18	21	20	20	22	21
27	5097002	58	A	SCC	G2	IVA	Intacto	6	13	10	9	16	12
28	5098561	36	#	SCC	G2	IB2	Intacto	33	64	53	52	64	62
29	5098970	41	A	SCC	G3	IB1	#	29	37	35	30	36	37
30	5099630	45	A	SCC	#	IIIB	Com ruptura	4	5	5	4	7	5
31	5099856	40	A	SCC	G2	IIIB	Com ruptura	6	8	9	7	9	9
32	5100265	75	A	SCC	#	IIIB	Com ruptura	3	11	10	5	7	6
33	5103540	58	A	SCC	G2	IIIB	Intacto	8	25	16	11	21	18
34	5105554	42	A	SCC	G3	IB2	Intacto	3	4	10	3	5	4
35	5106684	52	A	SCC	G2	IIB	Com ruptura	24	42	35	35	42	40
36	5106732	34	A	SCC	G3	IIB	Com ruptura	#	#	#	#	#	#
37	5107040	62	A	SCC	G2	IIIB	Com ruptura	#	#	#	#	#	#
38	5107209	44	A	SCC	G2	IIB	Com ruptura	#	#	#	#	#	#
39	5114518	48	A	SCC	G2	IIIB	Intacto	5	6	6	4	6	6
40	5115724	35	A	SCC	G2	IIB	Com ruptura	2	4	15	4	6	6
41	5116472	21	A	SCC	G2	IIIB	Com ruptura	6	17	10	9	22	13
42	5116780	44	A	SCC	G2	IIIB	#	#	#	#	#	#	#
43	5117413	47	A	SCC	G2	IIIB	Intacto	6	12	10	13	18	14
44	5119739	48	#	SCC	G2	IIB	Com ruptura	3	3	4	3	16	4
45	5120009	67	A	SCC	G2	IIB	#	6	8	7	7	9	10
46	5122632	43	A	SCC	G3	IB2	Com ruptura	18	26	19	22	28	33
47	5122818	52	B	SCC	G2	IIIB	Com ruptura	10	18	14	14	18	17
48	5104549	44	A	SCC	G2	IIIB	#	#	#	#	#	#	#
49	5112764	48	#	SCC	G2	IIIB	Com ruptura	2	5	4	3	5	4

#sem informação; Carga viral em 10⁶ células.

