



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE BIOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA

ANTÔNIO HUMBERTO PEREIRA DA SILVA JÚNIOR

**ANÁLISE DA VARIABILIDADE GENÉTICA DO ONCOGENE E5 DOS HPV_s 16 E
31 CIRCULANTES NO ESTADO DE PERNAMBUCO, BRASIL**

Recife
2018

ANTÔNIO HUMBERTO PEREIRA DA SILVA JÚNIOR

ANÁLISE DA VARIABILIDADE GENÉTICA DO ONCOGENE E5 DOS HPV's 16 E 31
CIRCULANTES NO ESTADO DE PERNAMBUCO, BRASIL

Tese apresentada ao Programa de Pós-Graduação em Genética da Universidade Federal de Pernambuco, como requisito para a obtenção do título de Doutor em Genética.

Área de Concentração: Genética.

Orientador: Prof. Dr. Antonio Carlos de Freitas.

Co-orientador (a): Prof^a. Dra. Bárbara Simas Chagas.

Recife

2018

Catálogo na fonte:
Bibliotecário Bruno Márcio Gouveia - CRB-4/1788

Silva Júnior, Antônio Humberto Pereira da
Análise da variabilidade genética do oncogênese E5 dos HPVs 16 e 31 circulantes
no Estado de Pernambuco, Brasil / Antônio Humberto Pereira da Silva Júnior. – 2018.

140 f. : il.

Orientador: Prof. Dr. Antonio Carlos de Freitas.
Coorientadora: Prof^a. Dr^a. Bárbara Simas Chagas.
Tese (doutorado) – Universidade Federal de Pernambuco. Centro de
Biotecnologias. Programa de Pós-graduação em Genética, Recife, 2018.
Inclui referências e anexos.

1. Papilomavírus. 2. Útero – Câncer. I. Freitas, Antonio Carlos de
(Orientador). II. Chagas, Bárbara Simas (Coorientadora). III. Título.

618.14

CDD (22.ed.)

UFPE/CB – 2018 - 413

ANTÔNIO HUMBERTO PEREIRA DA SILVA JÚNIOR

ANÁLISE DA VARIABILIDADE GENÉTICA DO ONCOGENE E5 DOS HPVs 16 E 31
CIRCULANTES NO ESTADO DE PERNAMBUCO, BRASIL

Tese apresentada ao Programa de Pós-Graduação em Genética da Universidade Federal de Pernambuco, como requisito para a obtenção do título de Doutor em Genética.

Aprovada em: 30/08/2018.

BANCA EXAMINADORA

Dr. Antonio Carlos de Freitas (Orientador)
Universidade Federal de Pernambuco

Dr. Tercílio Calsa Junior (Examinador Interno)
Universidade Federal de Pernambuco

Dr. Valdir de Queiroz Balbino (Examinador Interno)
Universidade Federal de Pernambuco

Dr. Marcelo Henrique Santos Paiva (Examinador Externo)
Universidade Federal de Pernambuco

Dr. Sérgio de Sá Leitão Paiva Júnior (Examinador Externo)
Universidade Federal Rural de Pernambuco

Aos meus Pais

Sr. Antônio Humberto Pereira da Silva (In memoriam)

& Sra. Mariluce Batista Castro

AGRADECIMENTOS

Agradeço a Deus por toda a saúde concedida ao longo da minha vida, para a execução das mais dolorosas tarefas, e pelos livramentos inerentes do nosso cotidiano. Toda a fé que possuo e todas as orações estão, sem dúvidas, acima de qualquer ato científico, e isso me fortalece como Ser Humano diariamente.

Ao Orientador deste projeto, o Professor Dr. Antonio Carlos de Freitas e às co-orientadoras Dra. Bárbara Simas Chagas e Dra. Ana Pavla Almeida Diniz Gulgel, que se fizeram presentes durante toda essa jornada e, além do incentivo diário, se mostraram sobretudo, grandes amigas.

Ao Dr. Marcus Vinícius de Aragão Batista pela orientação nas análises *in silico*, contribuindo com a sua expertise na área da bioinformática e com o engrandecimento deste projeto.

A todos os companheiros do Laboratório de Estudos Moleculares e Terapia Experimental (LEMTE – UFPE), pelo convívio diário, discussões científicas e pelas amizades construídas.

Um agradecimento, mais do que especial, aos meus familiares. Aos meus pais, o Sr. Antônio Humberto Pereira da Silva (*in memoriam*) e a Sra. Mariluce Batista Castro, por todo amor concedido a mim. Ao meu irmão, o Sr. José Carlos e à sua família, aqui representados por sua esposa Analice Castro e pelos meus sobrinhos Luiza Castro e Lucas Castro.

Aos encantos e desencantos proporcionados pelo Sujeito Amor.

Aos companheiros que fazem parte da Universidade Federal de Campina Grande.

Meus sinceros agradecimentos

Christo Nihil Praeponere

RESUMO

As infecções causadas pelos HPVs representam a principal causa para o surgimento do câncer cervical. Os genótipos mais comuns e com ampla distribuição no nordeste brasileiro são os do HPV16 e 31. O estudo das variantes do oncogene E5 destes genótipos, tem mostrado a sua participação em inúmeros processos celulares, agindo na regulação dos fatores de crescimento e de proliferação celular, além de atuarem no escape imunológico. Neste estudo, investigamos a circulação de variantes do oncogene E5 dos HPVs 16 e 31 e avaliamos *in silico*, as possíveis repercussões biológicas destes polimorfismos. As sequências gênicas obtidas de amostras cervicais foram sequenciadas e os polimorfismos identificados com o auxílio do software MEGA6. A estrutura secundária foi obtida pelo software PSIPRED, a pressão de seleção estimadas pelo software PAML e os epítomos mapeados com o auxílio da plataforma IEDB. Dentre as amostras HPV positivas, 24/93 foram HPV16 (sete polimorfismos identificados) e 35/93 foram HPV31 (dez polimorfismos identificados). Para o HPV16, as mutações I44L e I65V foram relacionadas ao aumento e à redução da estabilidade da proteína, respectivamente, e para o HPV31, o polimorfismo P56A foi predito como uma mutação estabilizante. As mutações encontram-se sob diferentes formas de seleção e diversos epítomos foram mapeados com alelos MHC-I e MHC-II. Estes resultados contribuem para o entendimento da distribuição geográfica das variantes e no desenvolvimento de novas estratégias imunoterapêuticas.

Palavras-Chave: HPV16. HPV31. Oncogene E5. Alpha-papilomavírus.

ABSTRACT

Infections caused by HPVs are the leading cause of cervical cancer. The most common and widely distributed genotypes in the Brazilian northeast are those of HPV16 and 31. The study of variants of the E5 oncogene of these genotypes has shown their participation in numerous cellular processes, acting on the regulation of growth factors and cell proliferation, as well as acting on immune escape. In this study, we investigated the circulation of E5 oncogene variants of HPVs 16 and 31 and evaluated in silico the possible biological repercussions of these polymorphisms. Gene sequences obtained from cervical samples were sequenced and polymorphisms identified with the aid of MEGA6 software. The secondary structure was obtained by the PSIPRED software, the selection pressure estimated by the PAML software and the epitopes mapped with the aid of the IEDB platform. Among the HPV positive samples, 24/93 were HPV16 (seven polymorphisms identified) and 35/93 were HPV31 (ten polymorphisms identified). For HPV16, the I44L and I65V mutations were related to the increase and reduction of protein stability, respectively, and for HPV31, the P56A polymorphism was predicted as a stabilizing mutation. These mutations are under different forms of selection and several epitopes have been mapped with MHC-I and MHC-II alleles. These results contribute to the understanding of the geographical distribution of variants and the development of new immunotherapeutic strategies.

Keywords: HPV16. HPV31. E5 Oncogene. Alpha-papillomavirus.

LISTA DE FIGURAS

Revisão de Literatura

Figura 1 -	Representação esquemática do genoma completo do HPV16.....	26
Figura 2 -	Representação esquemática das espécies Alpha-9, suas linhagens e sublinhagens.....	30
Figura 3 -	Processo infeccioso desenvolvido pelo HPV no epitélio cervical.....	31
Figura 4 -	Representação esquemática da proteína E5 com 83 aminoácidos....	38
Figura 5 -	Modelo hexamérico da proteína E5 do HPV16. A figura mostra um canal iônico gerado pela proteína E5 do HPV16 <i>in vitro</i>	39
Figura 6 -	Representação esquemática do software MEGA6 utilizado para a análise e identificação das sequências das variantes do oncogene E5 dos HPVs 16 e 31.....	52
Figura 7 -	Layout do software PSIPRED utilizado para a análise e construção da estrutura secundária da oncoproteína E5 dos HPVs 16 e 31.....	54
Figura 8 -	Plataforma IEDB.	56
Figura 9 -	Árvore topológica das variantes do HPV31 e comparações de pares de genomas completos individuais. Nesta construção, são distinguíveis as linhagens A, B e C e suas respectivas variantes.....	60
Figura 10 -	Topologia da árvore HPV16 e comparações pareadas de genomas completos individuais. Nesta construção, são distinguíveis as linhagens A, B, C e D, além de suas respectivas variantes.....	61

Artigo I

Figura 1 -	Árvore filogenética baseado na Inferência Bayesiana obtida a partir da análise da sequência do oncogene 31E5.....	73
Figura 2 -	Predição da estrutura secundária da oncoproteína 31E5. Resultados gráficos obtidos do servidor <i>on line</i> PSIPRED.....	76
Figura 3 -	Modelo tridimensional da oncoproteína 31E5 estabelecida a partir do método <i>ab initio</i>	78

Artigo II

Figura 1 -	Árvore filogenética baseado no método de Neighbor-Joining obtida a partir da análise da sequência do gene E5 de HPV16. O suporte	
------------	--	--

	dos ramos foi determinado por 1000 réplicas de bootstrap, sendo mostrados apenas os valores maiores que 50%.....	100
Figura 2 -	Predição da estrutura secundária da oncoproteína 16E5. Os gráficos foram obtidos com a utilização do servidor PSIPRED.....	102
Figura 3 -	Modelo tridimensional da proteína E5 de HPV16 estabelecida a partir do método <i>ab initio</i> . A estrutura evidencia os três domínios de alfa-hélice.....	103
Figura 4 -	Alinhamento estrutural da proteína E5 de HPV16 de referência (marrom), HPV16E5_14PE (azul) e HPV16E5_49PE (rosa).....	104

LISTA DE TABELAS

Revisão de Literatura

Tabela 1 - Estimativas de novos casos de câncer cervical nos Estados da Região Nordeste do Brasil, para o biênio 2018-2019.....	22
Tabela 2 - Estimativas de novos casos de câncer cervical nas Capitais dos Estados da Região Nordeste do Brasil, para biênio 2018-2019.....	23
Tabela 3 - Papel geral dos genes de expressão precoce (Earlier) e tardios (Late) que compõem a estrutura genômica dos HPVs.....	26
Tabela 4 - Comparação entre os tipos de métodos para inferência de filogenias.....	57

Artigo I

Tabela 1 - Resultado histopatológico das amostras de escovados cervicais, relacionadas à presença de polimorfismos em E5 de HPV31.....	71
Tabela 2 - Variações nucleotídicas do oncogene E5 do HPV31.....	71
Tabela 3 - Estimativas de pressão seletiva no oncogene E5 de HPV31 com base na probabilidade, calculada pelo programa PAML.....	74
Tabela 4 - Resultados do teste LRT para estimativas de pressão seletiva, que verifica qual modelo se ajusta melhor aos dados.....	74
Tabela 5 - Impacto das mutações na estabilidade da proteína E5.....	78

Artigo II

Tabela 1 - Resultado histopatológico das amostras de escovados cervicais, relacionadas à presença de polimorfismos em E5 de HPV16.....	95
Tabela 2 - Variações nucleotídicas do oncogene E5 do HPV16. A tabela mostra as variações presentes nas sublinhagens do HPV16, depositadas no banco de dados do NCBI.....	96
Tabela 3 - Variações nucleotídicas do oncogene E5 do HPV16.....	96
Tabela 4 - Estimativas de pressão seletiva no gene HPV16 E5 com base na máxima verossimilhança.....	100
Tabela 5 - Impacto das mutações na estrutura e/ou função da proteína E5 de HPV16.....	105
Tabela 6 - Análise dos sítios polimórficos e das mudanças na sequência de utilização de códons preferenciais nas variantes do oncogene E5	

do HPV16.....	105
---------------	-----

LISTA DE ABREVIATURAS E SIGLAS

AIC	Cr�terios de Informa��o Akaike
AP-1	Prote�na ativadora 1
APC	C�lulas apresentadoras de ant�geno
ASCUS	C�lulas escamosas at�picas de significado indeterminado
BCR-ABL	Prote�na quim�rica com atividade tirosino-quinase
BEAST	An�lise evolutiva bayesiana de amostragem de �rvores
BIC	Cr�terios de informa��o bayesiano
BLAST	Ferramenta b�sica de pesquisa de alinhamento local
C-CBL	Ubiquitina ligase c-cbl
CD1D	Glioprote�na expressa na superf�cie das apcs
C/EBP	Prote�na de liga��o � sequ�ncia ccaat
C-FOS	Proto-oncogene de ativa��o imediata
C-JUN	Prote�na quinase c-jun
COX	Ciclo-oxigenase
CTL	Linf�cito t citot�xico
DC	C�lulas dendr�ticas
EGF	Fator de crescimento epid�rmico
EGFR	Receptor para o fator de crescimento epitelial
ERK MAPK	Prote�na-quinase ativada por mit�geno erk
ETA	Receptores de endotelina acoplado � prote�na g
ET-1	Ligante espec�fico de endotelina
EV	Epidermiodisplasia verruciforme
EVER1 E EVER2	Genes relacionados � epidermiodisplasia verruciforme
FASL	Ligante do fator de necrose tumoral
FASTA	Alinhamento r�pido
FGFR2	Receptor do fator de crescimento de fibroblastos
GENES E	<i>Early</i> (precoce)
GENES L	<i>Late</i> (tardio)
GM	Ganglios�deo
HIV	V�rus da imunodefici�ncia humana
HLA	Ant�geno leucocit�rio humano

HPV	Papilomavírus humano
HR HPV	Papilomavírus de alto risco oncogênico
HSIL	Lesão intraepitelial escamosa alta
IEDB	Banco de dados de epitopo imunológico
IC	Carcinoma invasivo
ICTV	Comitê internacional de taxonomia de vírus
IFN	Interferon
IL	Interleucina
INCA	Instituto nacional do câncer
KDA	Quilodáton
KGFR	Receptor do fator de crescimento dos queratinóticos
KNB3	Proteína carioferina
LCR	Região longa de controle
LR HPV	Papilomavírus de baixo risco oncogênico
LSIL	Lesão intraepitelial escamosa baixa
M2	Proteína presente no invólucro do vírus influenza
MAPK	Proteína-quinase ativada por mitógeno
MEGA	Análise genética evolutiva molecular
MCMC	Monte carlo cadeia de markov
MHC	Complexo principal de histocompatibilidade
ML	Máxima verossimilhança
MP	Máxima parcimônia
mRNA	Ácido ribonucleico mensageiro
NCR	Região não codificante
NIC	Neoplasia intraepitelial cervical
NK	<i>Natural killer</i> (assassinas naturais)
NJ	<i>Neighbor joining</i>
ORF	Quadro aberto de leitura
ORI	Origem de replicação
PAE	Sítio de poliadenilação
PAML	Análise filogenética por máxima verossimilhança
PDB	Banco de dados de proteínas
PRB	Proteína retinoblastoma

PSIPRED	Bancada de análise de sequência de proteína
PV	Papilomavírus
SDM	Mutação direcionada ao site
SIL	Lesão intraepitelial escamosa
SNPS	Polimorfismo de nucleotídeo único
T CELL	Células t
TH	Células t humorais
TREG	Células t regulatórias
TLRS	Receptores toll-like
TNF	Fator de necrose tumoral
TRAIL	Ligante do fator de necrose tumoral
VARIANTE AA	Variante asiático-americano
VARIANTE AS	Variante asiático
VARIANTE AF-1	Variante africano 1
VARIANTE AF-2	Variante africano-2
VARIANTE E	Variante europeia
VEGF	Fator de crescimento endotelial vascular

SUMÁRIO

1	INTRODUÇÃO.....	18
1.1	OBJETIVOS.....	19
1.1.1	Objetivo Geral.....	19
1.1.2	Objetos Específicos.....	19
2	REVISÃO DE LITERATURA.....	21
2.1	OS PAPILOMAVÍRUS (PVS).....	21
2.1.1	Histórico.....	21
2.1.2	Câncer cervical: dados epidemiológicos e infecção.....	22
2.1.3	Estrutura viral e organização genômica.....	25
2.2	HPVS ENVOLVIDOS NA CARCINOGENESE CERVICAL.....	27
2.3	CLASSIFICAÇÃO TAXONÔMICA.....	27
2.4	HETEROGENEIDADE DOS HPVS E A SUA AÇÃO NA CARCINOGENESE CERVICAL.....	29
2.5	VISÃO GERAL DA RESPOSTA IMUNOLÓGICA NO TECIDO CERVICAL E ESTRATÉGIAS DE EVASÃO DO HPV.....	33
2.6	PAPEL GERAL DAS ONCOPROTEÍNAS E1, E2, E4, E6 E E7 NA CARCINOGENESE CERVICAL.....	35
2.7	ONCOPROTEÍNA E5.....	37
2.7.1	Estrutura e aspectos biológicos da oncoproteína e5.....	37
2.7.2	Via celular de E5 durante a replicação viral e atividade oncogênica....	41
2.7.3	Oncoproteína E5 e a evolução dos hpvs.....	43
2.7.4	Oncoproteína E5 e resposta imunológica.....	44
2.7.5	Importância da oncoproteína E5 no diagnóstico e no tratamento.....	45
2.8	EPÍTOPOS: VISÃO GERAL.....	48
2.9	APLICABILIDADE DAS FERRAMENTAS APLICADAS NA ANÁLISE <i>IN SILICO</i> PARA A GERAÇÃO E INTERPRETAÇÃO DE DADOS BIOLÓGICOS.....	50
2.9.1	Fasta e Blast.....	50
2.9.2	Alinhamento múltiplo.....	51
2.9.3	Análise das sequências gênicas do hpv16 e hpv31.....	51
2.9.4	PSIPRED: <i>protein sequence analysis workbench</i>.....	52

2.9.5	<i>Immune Epitope Database and Analysis Resource (IEDB)</i>	54
2.9.6	MrBayes 3.2, jModelTest 2 e Tracer 1.5.....	55
2.9.7	Análise da pressão de seleção.....	59
2.9.8	Estrutura terciária da proteína.....	62
3	RESULTADOS	64
3.1	ARTIGO 1 - DIVERSIDADE GENÉTICA DO ONCOGENE E5 DO PAPILOMAVÍRUS HUMANO 31 IDENTIFICADOS EM AMOSTRAS CLÍNICAS CIRCULANTES NO NORDESTE DO BRASIL.....	64
3.2	ARTIGO 2 - <i>IN SILICO</i> EVALUATION OF HUMAN PAPILLOMAVIRUS TYPE 16 E5 ONCOGENE FROM CERVICAL SAMPLES OF NORTHEAST BRAZILIAN WOMAN.....	87
4	DISCUSSÃO GERAL	116
5	CONCLUSÃO	122
	REFERÊNCIAS	124
	ANEXO A – CÓPIA DA APROVAÇÃO PELO COMITÊ DE ÉTICA INSTITUCIONAL	137
	ANEXO B – CÓPIA DO TCLE	138

1 INTRODUÇÃO

As infecções provocadas pelos papilomavírus humano (HPV) correspondem ao fator de risco mais importante na etiopatogênese do câncer cervical, e de suas lesões precursoras, e é considerada uma das doenças de transmissão sexual viral mais comum em mulheres sexualmente ativas. Devido à grande morbidade e mortalidade causadas por esses vírus, o Ministério da Saúde considera que os estudos sobre a biologia da infecção provocada pelo HPV, deve ser tratada como uma linha prioritária de estudos em saúde pública, por afetarem a qualidade e a expectativa de vida da população.

De um modo geral, os HPVs participam do processo de infecção da pele e das mucosas, atuando no desenvolvimento e progressão das lesões cervicais, sendo assim considerado um fator importante no curso da carcinogênese cervical. Dentre os tipos de HPVs mais comuns, encontrados associados aos casos de câncer cervical, destacam-se os 16, 18, 31 e 58. Destes, os HPVs 16 e 31 já foram descritos com uma ampla circulação no Nordeste do Brasil, estando também, associados a outros tipos de cânceres, como os anogenital, orofaríngeo, prostático, colorectal e o peniano.

A Organização Mundial de Saúde aponta para o câncer cervical como uma morbidade que afeta em mais de 80% da população dos países em desenvolvimento. A sua prevalência entre diferentes países, difere na quantidade e na ordem de magnitude do tipo viral presente na região. Países desenvolvidos apresentam uma menor taxa de infecções causadas pelos HPVs, enquanto os países em desenvolvimento ou subdesenvolvidos, apresentam uma elevada incidência de casos notificados pelas agências de saúde.

A infecção persistente associada a outros cofatores, dentre eles, o comportamento sexual de risco, múltiplos parceiros, multiparidade, co-infecções, uso de tabaco e de anticoncepcionais, são co-fatores que predispõe ao desenvolvimento da infecção, embora, em muitos casos, a infecção possa regredir devido a ação do sistema imune.

Os estudos mais recentes sobre a diversidade genética dos HPVs, apontam para a existência de inúmeras variantes gênicas, com diferentes potenciais de infectividade. Esta diversidade de tipos e subtipos virais, constituem um importante

veículo de interpretação para os achados clínicos e, como consequência, o estudo e a elaboração de novas estratégias terapêuticas.

Dentre os oncogenes que compõe a estrutura genética dos HPVs, o estudo do oncogene E5 se torna particularmente interessante, visto a sua atuação em diversos processos biológicos. O seu produto expresso, a oncoproteína E5, corresponde a uma proteína de baixo peso molecular e que, normalmente, encontra-se associada à membrana do retículo endoplasmático, do aparato de Golgi e do envelope nuclear, agindo, por exemplo, na desregulação do ciclo celular, inibição da apoptose, potencialização dos oncogenes E6 e E7, ação sobre as moléculas envolvidas no controle da diferenciação celular, sobrevivência e crescimento, além de atuarem sobre as moléculas que compõem o sistema imunológico (MHC classe I e MHC classe II). Estes achados tornam o gene E5, um importante alvo para a análise molecular de sua estrutura, e de sua proteína, das variantes gênicas dos HPV 16 e 31, considerados de alto risco oncogênico (HR HPVs).

O presente trabalho visou o estudo da heterogeneidade do oncogene E5 dos HPVs 16 e 31 e a circulação destas variantes no Estado de Pernambuco. Esta análise é justificada pela ampla distribuição destes genótipos nesta região geográfica e, a análise particular deste oncogene, se deu em decorrência da escassez de estudos sobre a sua diversidade genética e distribuição geográfica no País. Além disso, o presente trabalho se propôs a realizar uma análise *in silico* da estrutura da oncoproteína E5, de ambos os genótipos, investigando, assim, a estrutura secundária e terciária da oncoproteína, a geração de epítomos imunogênicos, além de conduzirmos uma investigação sobre a análise filogenética das variantes identificadas neste estudo.

1.1 OBJETIVOS

1.1.1 Objetivo Geral

Analisar a variabilidade genética do oncogene E5 e realizar uma análise *in silico* da estrutura da proteína E5, dos HPVs 16 e 31, circulantes no Estado de Pernambuco.

1.2.2 Objetivos Específicos

- Identificar os polimorfismos presentes no gene E5, obtidas a partir de amostras clínicas, de pacientes com ou sem lesões pré-cancerosas ou com carcinoma invasivo, coletadas no Hospital das Clínicas da UFPE (HC-UFPE);
- Analisar *in silico* a estrutura secundária e terciária da oncoproteína E5 dos HPVs 16 e 31, comparando-as e identificando possíveis alterações estruturais em relação à sequência referência de ambos os tipos virais para, assim, hipotetizar sobre os seus potenciais efeitos na biologia da infecção;
- Analisar se as mudanças nucleotídicas do oncogene E5 dos HPVs 16 e 31 estão inseridas em possíveis regiões de epítomos imunogênicos com moléculas pertencentes ao MHC classe I e MHC classe II;
- Realizar uma análise filogenética dos HPVs 16 e 31, baseado na sequência de nucleotídeos do oncogene E5;
- Realizar uma análise de pressão de seleção ao qual estas variantes estão submetidas;
- Analisar a utilização de códons preferenciais nos sítios polimórficos identificados no oncogene E5 dos HPVs 16 e 31.

2 REVISÃO DE LITERATURA

2.1 OS PAPILOMAVÍRUS (PVS)

2.1.1 Histórico

Em meados da década de 1950 e início dos anos 1960, quando as observações sobre os poliomavírus e os papilomavírus foram realizadas com o auxílio da microscopia eletrônica, pode-se constatar que ambos os vírus eram constituídos por um genoma de DNA circular dupla fita e por um capsídeo com formato icosaédrico não-envelopado. Como consequência destes achados, ambos os vírus foram classificados dentro de uma mesma família denominada *Papovaviridae* (BERNARD, 2005).

Por volta dos anos 1970, os estudos sobre a biologia da infecção provocada pelos papilomavírus, atraíram um grande interesse, por se tratarem de agentes que provocam o aparecimento de verrugas. Estas lesões benignas foram observadas em especial nos humanos e outros mamíferos, e por muito tempo foi considerado um problema de ordem estética, e não um problema de saúde pública (ZUR HAUSEN, 2002; BERNARD et al., 2010).

Embora, as dificuldades de encontrar um sistema de cultura celular, adequado para propagar estes vírus, afetassem o progresso dos estudos sobre a biologia da infecção, e limitasse o estabelecimento de uma taxonomia baseada nas propriedades biológicas dos vírus, os primeiros Papilomavírus (PVs) começaram a ser isolados (KREIDER et al., 1985; HUMMEL et al., 1995). Nos estágios iniciais das pesquisas, a coleta de verrugas e a análise das lesões dos pacientes com Epidermiodisplasia verruciforme (EV), foram utilizadas para isolar grandes quantidades de partículas virais, objetivando-se isolar o DNA genômico dos vírus. As análises prosseguiram com a interpretação do padrão de bandas feitas por digestão enzimática e por hibridização por Southern Blot (VILLIERS et al., 2004).

Nos anos 1980, com o desenvolvimento de novas técnicas de biologia molecular, dezenas de HPVs foram detectados em lesões de mucosa de natureza benigna e maligna, como as encontradas no câncer cervical e nas suas lesões precursoras (ZUR HAUSEN, 2002). Nesta mesma década, estudos da sequência e análise funcionais, demonstraram que as similaridades encontradas entre os poliomavírus e os papilomavírus, que os faziam pertencer à mesma família, eram

muito superficiais para se estabelecer uma possível relação. Um dos achados, estava na análise do tamanho do genoma; enquanto os poliomavírus apresentam um genoma de tamanho de 5 Kb, os papilomavírus, possuem um genoma com tamanho aproximado de 8 Kb (REBRIKOV et al., 2002).

Nos anos 2000, a família *Papillomaviridae* tornou-se oficialmente reconhecida pelo *International Committee on Taxonomy of Viruses* (ICTV) (VILLIERS et al., 2004). Os dados atuais mostram a existência de mais de 200 tipos de HPVs, com o seu genoma isolado e totalmente sequenciados (VAN DOORSLAER et al., 2013).

2.1.2 Câncer Cervical: Dados Epidemiológicos e Visão Geral da Diversidade dos HPVs

O câncer cervical representa a quarta causa mais comum de cânceres em mulheres ao redor do mundo, tendo sido estimado cerca de 528 mil novos casos e 266 mil mortes, no ano de 2012 (FERLAY et al., 2015). No Brasil foram estimados 16.370 novos casos de câncer cervical para o biênio 2018-2019, com um risco estimado de 15,43 casos a cada 100 mil mulheres, ocupando a terceira posição dos tipos de cânceres mais frequentes no país e, no Nordeste do Brasil, ocupa a segunda posição, com um risco estimado de 20,47 para cada 100 mil mulheres (tabela 1) (INCA, 2018).

Tabela 1 – Estimativas de novos casos de câncer cervical nos Estados da Região Nordeste do Brasil, para biênio 2018-2019.

Estado	Câncer Cervical	Taxa Bruta por 100 mil hab.	Ranking (Nº de casos)
Bahia	1230	15,72	1º
Maranhão	1090	30,55	2º
Pernambuco	1030	20,84	3º
Ceará	990	21,24	4º
Piauí	430	26,19	5º
Paraíba	370	17,57	6º
Alagoas	320	18,49	7º
Rio Grande do Norte	320	17,93	7º
Sergipe	250	20,78	8º

Fonte: INCA, 2018.

No Estado de Pernambuco estima-se 1030 novos casos sejam diagnosticados no ano de 2018, ficando atrás apenas dos Estados da Bahia e do Maranhão. Analisando as estimativas por Capital, o município de Recife ocupa a quarta posição entre as capitais do nordeste brasileiro (tabela 2) (INCA, 2018).

Tabela 2 – Estimativas de novos casos de câncer cervical nas Capitais dos Estados da Região Nordeste do Brasil, para biênio 2018-2019.

Estado	Câncer Cervical	Ranking
Salvador	300	1º
São Luís	240	2º
Fortaleza	290	3º
Recife	180	4º
Teresina	140	5º
Maceió	130	6º
Natal	100	7º
João Pessoa	80	8º
Aracaju	60	9º

Fonte: INCA, 2018.

A presença do Papilomavírus (HPV) é considerada necessária, mas não o suficiente para provocar o desenvolvimento do câncer cervical (BOSCH et al., 2002). Os tipos virais que estão associados com o desenvolvimento e progressão das lesões são denominados HPVs de alto risco (*High Risk HPV*), ou simplesmente HR-HPV (MUÑOZ et al., 2006).

Em muitos casos, a evolução do câncer cervical se apresenta lentamente. Muitas lesões podem regredir, porém, as lesões precursoras do câncer cervical, recebem a denominação de Neoplasias Intraepiteliais Cervicais (NICs), que podem evoluir gradativamente, passando por três estágios, conhecido como lesão grau 1 (baixo risco oncogênico – Low Risk HPV (LR HPV)), e as lesões de alto risco oncogênico, chamadas de lesões grau 2 e grau 3 (High Risk), simbologicamente representadas por NICI, NICII e NICIII, até o estabelecimento da neoplasia maligna (STEENBERGEN et al., 2005).

No Brasil diversos programas de prevenção, controle e rastreamento do câncer cervical, o que permite a conscientização e a detecção precoce de anormalidades (ANGHEBEM-OLIVEIRA; MERLIN, 2010). Para o rastreamento do câncer cervical, temos o exame citopatológico, através do teste Papanicolau, recomendado pelo Ministério da Saúde e adotado em mulheres sexualmente ativas, prioritariamente entre 25 a 64 anos, sendo esta, a principal estratégia dos programas de rastreamento da doença. O exame permite encontrar as lesões pré-invasivas, o que demonstra a necessidade e a importância da precocidade e exatidão do diagnóstico (STEENBERGEN et al., 2005). Por outro lado, a análise histopatológica, vem sendo considerado o padrão-ouro para o diagnóstico, embora ocorram variações na interpretação das atipias celulares, e na liberação de resultando diagnósticos inconsistentes (ANGHEBEM-OLIVEIRA; MERLIN, 2010).

Atualmente, mais de 200 genomas de HPVs foram totalmente sequenciados. Deste total, quinze genótipos de HPVs estão associados com o alto risco oncogênico (MUÑOZ et al., 2006). Na Região Nordeste do Brasil, os genótipos mais prevalentes em lesões cervicais foram os do HPV16, HPV18, HPV31, HPV33 e HPV58 (CHAGAS et al., 2011; CHAGAS et al., 2013; GURGEL et al., 2015).

A transmissão se dá através de relações sexuais (vaginal, oral e/ou anal) onde os vírus aproveitam a presença de microtraumas na mucosa para alcançar as células basais (SAPP; BIENKOWSKA-HABA, 2009). Nos indivíduos imunocompetentes, o vírus pode ser eliminado do organismo antes de causar a lesão; caso contrário, teremos a formação de lesões intraepiteliais de baixo grau (*Low squamous intraepithelial lesion* - LSIL), que poderá evoluir para lesões de alto grau (*High squamous intraepithelial lesion* - HSIL). Podemos, em seguida, ter uma infecção persistente com integração do genoma viral no genoma do hospedeiro, que é uma etapa importante para o surgimento do câncer cervical (WOODMAN et al., 2007).

As ações virais visam a modulação da proliferação, apoptose e da resposta imune do hospedeiro. A resposta imune celular corresponde ao principal tipo de proteção contra a infecção pelo HPV. No caso das infecções persistentes, a resposta imune celular é inibida, o que contribui para o não reconhecimento do vírus pelo organismo hospedeiro e consequente manifestação do seu perfil infeccioso (YU et al., 2012).

Nas lesões de pele benignas, associadas ao HPV, o genoma viral encontra-se separado do DNA celular, ao contrário das lesões malignas, onde o DNA viral integra-se aos cromossomos do hospedeiro (KAUFMAN et al., 2000). Neste último caso, o processo de integração ocorre mediante uma quebra do genoma viral, em geral, nas regiões dos genes E1 e E2, o que resulta na perda de função de ambos os genes, acompanhada de uma desregulação dos genes E6 e E7, levando a uma transformação da célula hospedeira. Um fato importante, é que a região ao qual o genoma viral se integra não parece ser essencial para o desenvolvimento carcinogênico, uma vez que se observa uma grande variação de locais de inserção (BURD, 2003).

Dessa maneira, a região de integração do genoma viral compreende a região LCR (*Long Control Region*) e os oncogenes E6 e E7, estes últimos associados com a degradação de p53 e pRb, respectivamente, o que gera instabilidade genômica e

desvio das rotas que culminam com a apoptose (ZUR HAUSEN, 1996; ZUR HAUSEN, 2002).

2.1.3 Estrutura Viral e Organização Genômica

O HPV é um pequeno vírus de DNA não-envelopado, com diâmetro de 55 nm, DNA circular de fita dupla, com cerca de 8000 pares de base (pb) associados às histonas e tem como característica exibir tropismo por células epiteliais (ZUR HAUSEN, 1996).

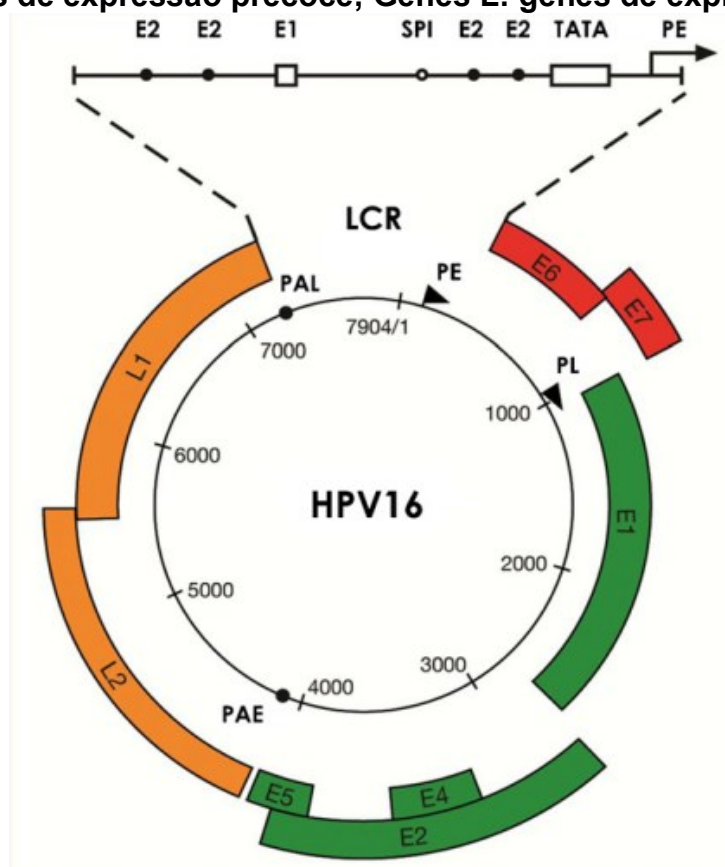
O seu genoma apresenta oito quadros abertos de leitura (*Open Reading Frames* – ORFs) e pode ser dividido nas seguintes regiões: (1) uma região de controle denominada de Região Longa de Controle (LCR); (2) regiões de expressão precoce dos genes, denominados genes *Earlier* (genes E); (3) regiões de expressão tardia dos genes, denominados genes *Late* (genes L); e (4) uma região de controle, denominada *Non-Coding Region* (NCR) (ZUR HAUSEN, 1996) (figura 1).

A região LCR é um segmento de aproximadamente 850 pb, situada entre os gens L1 e E6, corresponde a uma fração de 10% do genoma total, contendo, na sua estrutura, a origem de replicação viral (denominada de ORI) e alguns sítios de transcrição gênica. Além desta região, temos os genes E, representados pelas regiões E1, E2, E4, E5, E6 e E7, que estão relacionados à regulação do ciclo viral e constituem mais de 50% do tamanho total do genoma. Finalmente, os genes L, responsáveis por sintetizar a estrutura do capsídeo viral (ZUR HAUSEN, 1996) (tabela 3).

A região NCR, localizada entre os genes E5 e L1, também atua na regulação da transcrição e replicação viral. Contém um sítio de poliadenilação (pAE), que é responsável pela regulação dos genes precoce do genoma viral (ZUR HAUSEN, 1996).

Os vírus classificados como de baixo risco de transformação maligna, por exemplo os genótipos HPV6, -11, -40, -42, -43, -44, -54, -61, -70, -72 e o -81, não possuem a capacidade de integração do seu material genético com a célula hospedeira, divergindo, portanto, dos vírus classificados como de alto risco oncogênico, como aqueles representados pelos genótipos HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 58, 59, 68, 73 e 82 (SUMMERSGILL et al., 2000).

Figura 1 – Representação esquemática do genoma completo do HPV16. A estrutura é representada por um DNA circular fita dupla, com cerca de 8kb. Genes E: genes de expressão precoce; Genes L: genes de expressão tardia.



Extraído de Doorbar et al. (2012).

Tabela 3 – Papel geral dos genes de expressão precoce (Earlier) e tardios (Late) que compõem a estrutura genômica dos HPVs.

Gene	Significado Biológico
L1	Síntese da proteína maior do capsídeo, empacotamento do genoma e liberação de partículas virais.
L2	Síntese da proteína menor do capsídeo, empacotamento do genoma e liberação de partículas virais.
E1	Atividade de DNA helicase e ligação de ATP DNA dependente. Atua no controle da replicação do DNA viral na célula infectada. Forma um complexo com a proteínas E2, o que permite a sua atividade catalítica.
E2	Atua como regulador da transcrição, ativando ou reprimindo a ativação dos genes precoces do genoma viral, além de atuar, juntamente com a proteína E1, no controle do ciclo de replicação viral.
E4	Expressão primária no epitélio em diferenciação, associado ao citoesqueleto de queratina das células epiteliais em cultura. Atua na liberação das partículas virais.
E5	Estimula a proliferação celular, através da ativação dos receptores para o fator de crescimento epitelial (EGFR), e nas etapas de iniciação da carcinogênese, além de interagir com moléculas do sistema imune.
E6	Papel na transformação celular juntamente com a proteína E7 e na degradação da proteína p53 e de outras proteínas que atuam na supressão tumoral. Também é responsável pela inibição da apoptose.
E7	Semelhante à proteína E6, atua na transformação dos queratinócitos e provoca a inativação de pRb, induzindo, assim, a proliferação celular.

2.2 HPVS ENVOLVIDOS NA CARCINOGENESE CERVICAL

A infecção pelo HPV pode se dar por atividade sexual, contato com tecido infectado e transmissão vertical. Além disso, muitos estudos apontam alguns cofatores que estão diretamente associados ao surgimento do câncer cervical, tais como múltiplos parceiros, multiparidade, tabagismo, imunossupressão e uso de anticoncepcionais (QUEIROZ et al., 2007).

Atualmente são descritos mais de 200 genótipos de HPV já identificados, onde cerca de 40 destes estão relacionadas as infecções do trato genital feminino e 15 deles são potencialmente oncogênicos, sendo os tipos 16 e 18 os de maior prevalência e oncogenicidade. Dentre os genótipos considerados de alto risco oncogênico (HR-HPV), destacam-se os HPV -16, -18, -31, -33, -35, -39, -45, -51, -52, -58, -59, -68, -73 e -82, e os de baixo risco (LR-HPV), os tipos HPV -6, -11, -40, -42, -43, -44, -54, -61, -70, -72 e -81 (CLIFFORD et al., 2003).

Os tipos HPV16, -18, -31 e -52, são encontrados com maior frequência em mulheres com ou sem lesão. Nas lesões de baixo grau histológico (*Low Squamous Intraepithelial Lesion* - LSIL), os genótipos mais encontrados são os 16, 51, 31 e 52, e nas lesões de alto grau (*High Squamous Intraepithelial Lesion* - HSIL) no carcinoma invasivo, os genótipos mais frequentes são os 16, 18, 31, 58 e 33. Estes dados estão diretamente relacionados à geografia da localidade estudada. Nos países em desenvolvimento encontram-se uma ampla distribuição os HPV16, 18, 58 e 33 e, nos países desenvolvidos, os genótipos 16, 18, 31 e 33, prevalecem (SANJOSE et al., 2010).

2.3 CLASSIFICAÇÃO TAXONÔMICA

A classificação taxonômica dos tipos de papilomavírus baseia-se na análise da ORF do gene L1. Sendo assim, um novo isolado de HPV é reconhecido quando a sequência de DNA da ORF do gene L1 difere mais do que 10% das espécies proximamente relacionadas. Diferenças entre 2% e 10% de homologia definem um subtipo e diferenças menores do que 2%, definem uma variante (VILLIERS et al., 2004).

Outro critério taxonômico adotado trata da utilização do termo Gênero, onde os HPVs deverão compartilhar no mínimo 60% de identidade da sequência de nucleotídeos da ORF do gene L1. Para a ordenação na categorização de Espécie,

os HPVs deverão compartilhar no mínimo 71% a 89% de identidade da sequência de nucleotídeos da ORF completa de L1 (VILLIERS et al., 2004).

O rápido aumento no número de PVs isolados demonstrou a necessidade de um sistema de classificação dentro da família *Papillomaviridae*, cujo objetivo é de estabelecer relações entre os tipos de HPVs e investigar, através da classificação taxonômica, as propriedades biológicas e patológicas dos vírus (BURK et al., 2013).

Mais recentemente, Burk et al. (2013) propôs uma classificação e uma nomenclatura para a diversidade intratípica baseado na análise da divergência completa das sequências completa do genoma dos HPVs, classificando em linhagens àquelas que apresentam divergência entre 1-10% em seu genoma e sublinhagens, àquelas com divergências entre 0,5-1% em seu genoma completo, e cada variante, anotada com um termo alfa-numérico.

Em relação ao estudo das variantes virais, muitos tipos de HPVs tem sido encontrado e isolado em amostras clínicas, e suas sequências gênicas comparadas. Existe uma enorme expectativa na diferenciação destes isolados e no estudo da diversidade genética encontrada nos HPVs, uma vez que esta heterogeneidade pode repercutir nas implicações clínicas, como na diversidade de sinais patológicos, apresentados pelos diversos tipos de HPVs (CHAGAS et al., 2013; GURGEL et al., 2015; FREITAS et al., 2012; FREITAS et al., 2017).

Na família *Papillomaviridae*, atualmente, estão descritos trinta e sete gêneros, embora a diversidade de PVs ainda esteja subestimada (VAN DOORSLAER et al., 2013). Os gêneros descritos até o presente são os *Alphapapilomavírus*, *Betapapilomavírus*, *Gamapapilomavírus*, *Deltapapilomavírus*, *Epsilonpapilomavírus*, *Zetapapilomavírus*, *Etapapilomavírus*, *Thetapapilomavírus*, *Lotapapilomavírus*, *Kappapapilomavírus*, *Lambdapapilomavírus*, *Mupapilomavírus*, *Nupapilomavírus*, *Xipapilomavírus*, *Omikronpapilomavírus*, *Pipapilomavírus*, *Rhopapilomavírus*, *Sigmapapilomavírus*, *Taupapilomavírus*, *Upsilonpapilomavírus*, *Phipapilomavírus*, *Chipapilomavírus*, *Psipapilomavírus*, *Omegapapilomavírus*, *Dyodeltapapilomavírus*, *Dyoepsilonpapilomavírus*, *Dyozetapapilomavírus*, *Dyoetapapilomavírus*, *Dyothetapapilomavírus*, *Dyoiotapapilomavírus*, *Dyolambdapapilomavírus*, *Dyomupapilomavírus*, *Dyonupapilomavírus*, *Dyomikronpapilomavírus*, *Dyophipapilomavírus*, *Dyorhopapilomavírus*, *Dyosigmapapilomavírus* e *Dyochipapilomavírus* (BURK et al., 2013).

Em relação ao Gênero *Alphapapilomavírus*, dos mais de 200 tipos de HPV que já foram bem caracterizados, aproximadamente 60 destes são predominantemente encontrados no epitélio da mucosa, o que demonstra que as infecções causadas pelos Alpha-PVs estão diretamente associadas no desenvolvimento de ambas as doenças, benigna e maligna, tais como condiloma acuminado, papilomatose respiratória e diversos cânceres como os cervical, anal, cabeça e pescoço (VILLIERS et al., 2004; BERNARD et al., 2010).

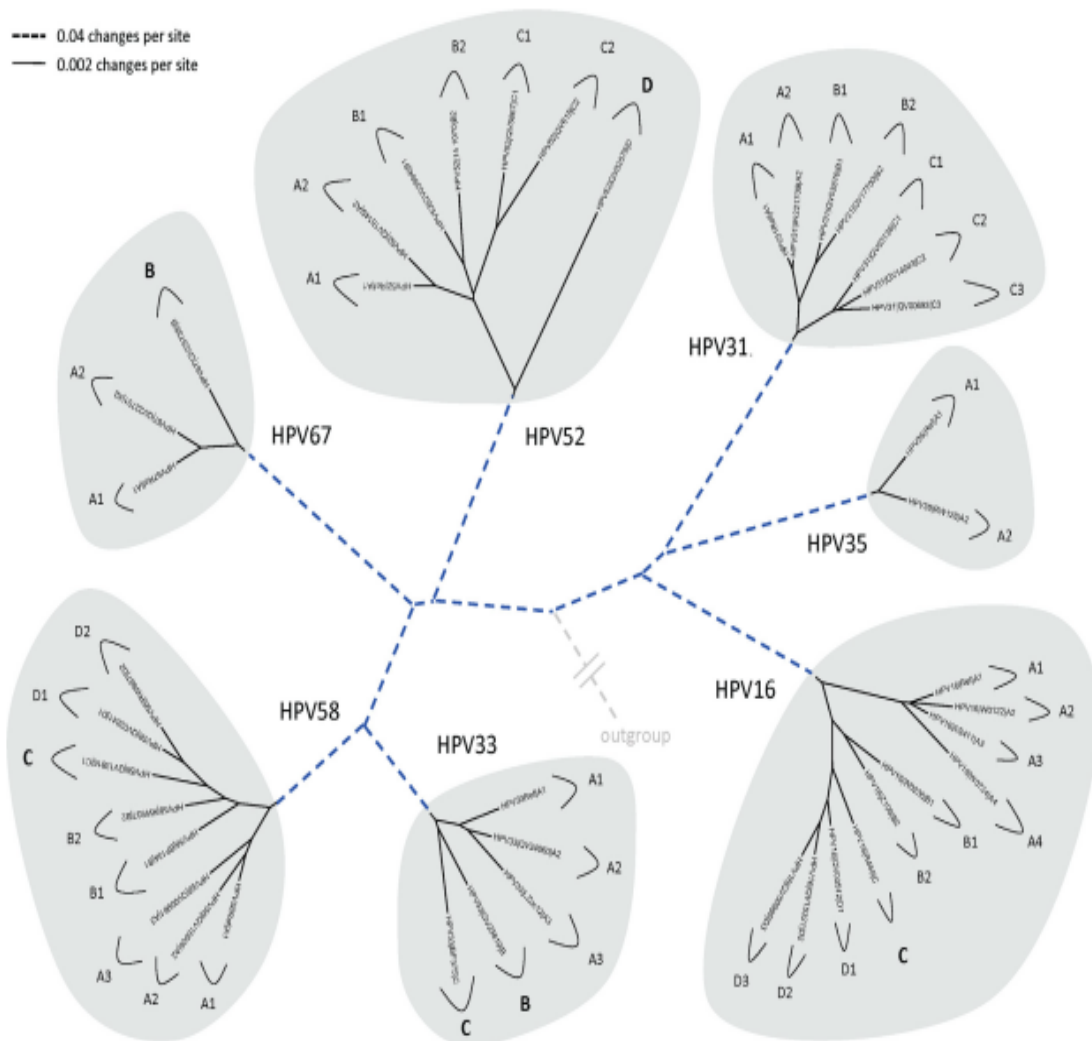
Dentre os HR-HPVs envolvidos na carcinogênese cervical, aqueles pertencentes ao alpha-HPV (em especial os pertencentes ao grupo alpha-5, alpha-6, alpha-7 e alpha-9), representam cerca de 90% de participação no envolvimento do câncer cervical (LI et al., 2011; BURK et al., 2013; FREITAS et al., 2017).

Os *Alphapapilomavírus* são classificados nas seguintes espécies: **Alpha-9** (HPV16, HPV31, HPV33, HPV35, HPV52, HPV58, HPV67) (figura 2); **Alpha-7** (HPV18, HPV 39, HPV45, HPV59, HPV68, HPV70, HPV85, HPV97); **Alpha-5** (HPV26, HPV51, HPV69, HPV82); **Alpha-6** (HPV30, HPV53, HPV56, HPV66); **Alpha-11** (HPV34, HPV73); **Alpha-10** (HPV6, HPV11); **Alpha-13** (HPV54); **Alpha-3** (HPV61) (BURK et al., 2013).

2.4 HETEROGENEIDADE DOS HPVS E A SUA AÇÃO NA CARCINOGENESE CERVICAL

O desenvolvimento do câncer cervical segue os estágios de exposição sexual a um HPV oncogênico, persistência da infecção e progressão da lesão pré-cancerosa para o câncer invasivo (BOUVARD et al., 2009; WENTZENSEN et al., 2009; SCHIFFMAN; WENTZENSEN, 2010; SCHIFFMAN; WENTZENSEN, 2013). Em cada estágio, pode-se ter a influência dos fatores ambientais, da biologia do hospedeiro e das variações genômicas dos vírus (SCHIFFMAN; WENTZENSEN, 2010; SCHIFFMAN; WENTZENSEN, 2013). A figura 3 retrata o processo infeccioso desenvolvido pelo Papilomavírus Humano no epitélio cervical.

Figura 2 – Representação esquemática das espécies Alpha-9, suas linhagens e sublinhagens.

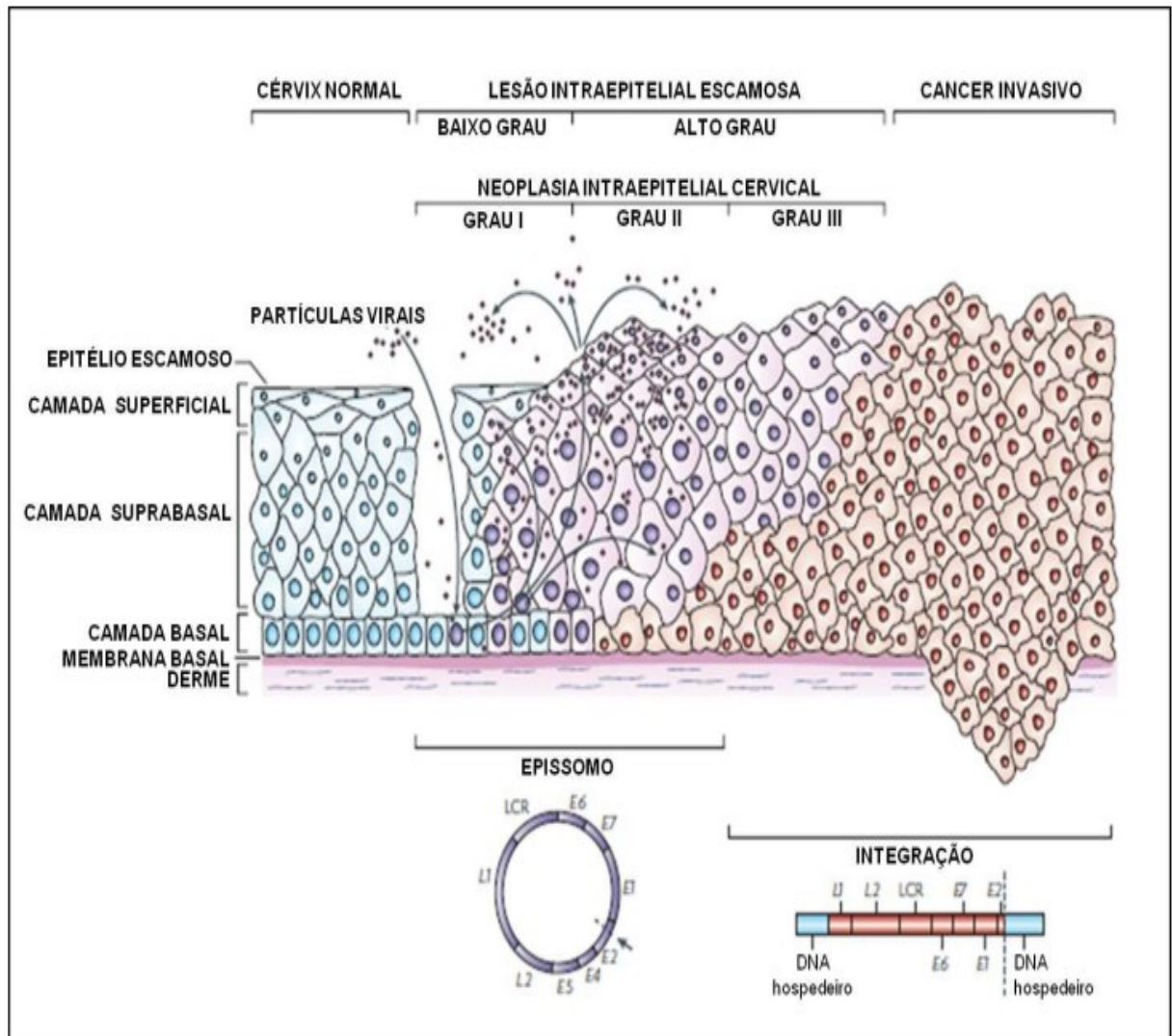


Fonte: Chen et al. (2011).

A contribuição das variantes virais no desenvolvimento do câncer cervical, está nas diferenças de infectividade, dado a partir das exposições a um tipo ou outro de HPV, refletindo em respostas diferenciadas do sistema imune, na proliferação e diferenciação viral (SCHIFFMAN; WENTZENSEN, 2013; FREITAS et al., 2012).

As análises do potencial patogênico das variantes de HPV podem sofrer influência de diversos fatores, que tornam as interpretações dos dados extremamente complicados. Tal fato se deve às complexas interações dos vírus com a genética dos seus hospedeiros, bem como a prevalência de variantes de HPV circulantes em diferentes populações e regiões geográficas (SOUZA et al., 2009).

Figura 3 – Processo infeccioso desenvolvido pelo HPV no epitélio cervical. A figura aponta para a progressão das lesões e seus estados de NICs.



Adaptado de Woodman et al. (2007).

O estudo da heterogeneidade dos HPVs envolvidos na carcinogênese cervical, tem demonstrado que a principal via de variação genética vem surgindo ao longo do acúmulo de alterações nas sequências de nucleotídeos, sendo esse processo conhecido como fixação de linhagem. Sendo assim, o acúmulo destas variações nucleotídicas contribuem para a diversificação e especiação dos HPVs. Por exemplo, estudos com os HPVs 16 e 18, demonstram que ambos sofreram um mecanismo de coevolução, onde percebe-se a existências de três ramificações filogenéticas, designadas como Africana, Caucasiana e Asiática. As linhagens de HPV16 ainda passaram por um reagrupamento em cinco categorias distintas, chamadas de Europeu (E), Asiático (As), Asiático-Americano (AA), Africano-1 (Af-1)

e Africano-2 (Af-2); enquanto que as linhagens de HPV18, foram categorizadas em três grupos distintos, sendo o Europeu (A), Asiático-Americano (AA) e o Africano (Af) (BURK et al. 2013).

Estas variações genômicas podem ser compreendidas, como importantes marcadores encontrados no genoma dos HPVs e podem ser devidamente utilizados em estudos de etiologia e epidemiologia, objetivando-se investigar a sua transmissão intra e interpopulacional. Por exemplo, a utilização da variação genômica encontrada no HPV16 pode ser utilizado para diferenciar as infecções persistentes de uma nova infecção, provocada por outra variante do mesmo genótipo HPV16 (MAYRAND et al., 2000).

No genoma celular, diferentes graus de patogenicidade podem ocorrer devidos às infecções provocadas pelas variantes dos HPVs. Alguns efeitos podem ser observados como alterações na taxa de replicação e transcrição, mudanças nos potenciais carcinogênicos dos oncogenes virais, além de alterações na eficiência da infecção (BERNARD et al., 2006).

No sistema de classificação mais atual, os HPVs oncogênicos de maior prevalência no Brasil, os tipos 16, 18, 31, 33 e 58, passaram por uma reformulação em relação a sua classificação em variantes e sublinhagens. O HPV16 possui quatro variantes designadas por A, B, C e D e dez sublinhagens denominadas A1, A2, A3, A4, B1, B2, C, D1, D2 e a D3; O HPV 18 possui três variantes, designadas A, B e C e nove sublinhagens denominadas A1, A2, A3, A4, A5, B1, B2, B3 e a C; O HPV 31 possui três variantes, designadas A, B e C e sete sublinhagens denominadas A1, A2, B1, B2, C1, C2 e C3; O HPV 33 possui três variantes, designadas A, B e C e cinco sublinhagens denominadas A1, A2, A3, B e C; O HPV 58 possui quatro variantes, designadas A, B, C e D e oito sublinhagens denominadas A1, A2, A3, B1, B2, C e D1, D2 (BURK et al. 2013).

Esta heterogeneidade mostra que a substituição de nucleotídeos pode culminar com o surgimento de novas variantes de HPVs e afetar o seu potencial oncogênico, tornando algumas variantes mais oncogênicas em relação a outras. Por isso, um melhor entendimento do processo infeccioso mediado por estas variantes pode contribuir na avaliação do seu grau de patogenicidade (FREITAS et al., 2012).

Por exemplo, no estudo de prevalência das variantes Af e AA dos HPV16 e 18, percebeu-se uma elevada disseminação destas variantes na América Latina,

África e Ásia e, coincidentemente, ocorreu o aumento no número de casos relacionados aos cânceres anogenitais (VILLA et al., 2000; BERUMEN et al., 2001).

No que se refere à patogenicidade causada pelas variantes, uma importante análise funcional trata da presença de uma substituição no aminoácido L83V na oncoproteína E6 do HPV16, e sugere a possibilidade desta variante diferir em patogenicidade do seu protótipo de HPV16 devido à sua forte associação com a persistência da infecção. Esta mutação foi observada na população da Grã-Bretanha e na Suécia, e foi associada a uma elevada taxa de progressão da infecção (LONDESBOROUGH et al., 1996; NINDL et al., 1996; BONTKES et al., 1998). Embora estes estudos apontem para o potencial infeccioso desta variante L83V, os mesmos resultados não foram reproduzidos nos estudos realizados na população da Holanda, Alemanha, Itália e América do Norte (ZEHBE et al., 2001; RAJEEVAN et al., 2005).

Outra mudança na sequência de aminoácidos foi relatada na oncoproteína E6 do HPV16. A mutação E31G foi sugerida como uma alteração capaz de aumentar o risco de progressão da lesão, devido à sua interferência em subtipos protéicos do sistema HLA (ETHERINGTON et al., 1999).

Em relação ao gene E7 do HPV16, existem poucas mutações relatadas. O estudo mais significativo apontou para a variação N29S envolvida, principalmente, com o carcinoma invasivo (SONG et al., 1997).

Assim sendo, o papel funcional das variantes dos HPVs, pode estar diretamente associada com a diversidade patogênica, o que pode interferir na alteração do funcionamento dos elementos regulatórios dentro do genoma, influenciando a persistência da infecção e o risco aumentado para o surgimento do carcinoma invasivo (SCHIFFMAN et al., 2005), além de influenciar a ativação dos oncogenes virais, a exemplo de E6 e E7, que naturalmente atuam sobre importantes componentes do ciclo celular e elementos regulatórios e de reparo do DNA, como as proteínas p53 e pRb (HUBERT et al., 2005).

2.5 VISÃO GERAL DA RESPOSTA IMUNOLÓGICA NO TECIDO CERVICAL E ESTRATÉGIAS DE EVASÃO DO HPV

No tecido cervical, as células do sistema imune inato como os macrófagos, células dendríticas, células de Langerhans, neutrófilos, linfócitos T e células *Natural Killer* (NK) conseguem reconhecer estruturas não próprias do hospedeiro, por meio

de receptores celulares, como os *Toll-like receptores* (TLRs). Esse reconhecimento leva à expressão de citocinas inflamatórias e de diversas quimiocinas, tais como as interleucinas (IL) IL1 β , IL6, IL8 e IL12, e também o fator de necrose tumoral (TNF α) e os interferons (IFNs) α e β (ZHANG et al., 2007).

As células do sistema imune são, então, recrutadas, além da secreção de fatores anti-microbianos e os elementos da resposta imune inata e adaptativa. As células dendríticas (DC), por exemplo, atuam como células apresentadoras de antígenos (APCs) que, após o reconhecimento do antígeno por uma citocina, ativa a expressão dos componentes protéicos relacionados ao complexo principal de histocompatibilidade (MHC; HLA em humanos), para a apresentação do antígeno e o seu processamento pelas células T (DEGLI-ESPOSTI; SMYTH, 2005). As DCs precisam migrar para as áreas onde existem as células T, no tecido linfóide secundário, para serem submetidos a etapa final de maturação através da expressão de moléculas co-estimuladoras, tais como CD40 e CD80. Este mecanismo induz a proliferação de células T, a produção de IL-2 e a regulação positiva dos genes anti-apoptóticos (FIFE; BLUESTONE, 2008).

A resposta das células T CD4⁺ (e as citocinas que produzem), passam a ser classificadas como Th1 mediada por células (incluindo IL1, IL2, IL15, TNF α , e γ -IFN); humoral Th2 (incluindo IL4, IL5, IL6, IL10 e IL13) e Treg/Th3 (que compreende IL10, TGF- β e γ -IFN) e Th17 (incluindo IL17, IL23 e IL32). Já a resposta citotóxica mediada pelas células T CD8⁺ desencadeia o desenvolvimento de linfócitos T citotóxicos (CTL), que segregam enzimas proteolíticas e porfirinas (BROERE et al., 2011).

Os mecanismos de evasão do sistema imunológico, mediado pelo HPV, resulta em danos nos diferentes estágios da imunidade inata e adaptativa, que vão desde a entrega do antígeno à expressão das citocinas (CONESA-ZAMORA, 2013).

O HPV é raramente detectado em locais sistêmicos ou em outros tipos de células que não o epitélio escamoso. A exposição dos seus antígenos é minimizada, o que torna difícil a detecção pelo sistema imunológico do hospedeiro. Outro fator que contribui para essa evasão, é que o HPV depende do programa de diferenciação dos queratinócitos, o que explica a baixa expressão gênica dos genes E, e somente na camada basal do epitélio, onde a célula tem a capacidade de se proliferar, garantindo a replicação viral. Em contraste, os genes L, que são mais imunogênicos, se expressam numa camada mais afastada, onde as APCs não são

abundantes e onde os queratinócitos extravasam para dentro do lúmen do colo do útero, favorecendo a propagação do vírus (TINDLE, 2002).

Do ponto de vista evolutivo, considera-se que os HPVs evoluíram para retardar a expressão dos genes relacionados à síntese das proteínas do capsídeo L1 e L2. Outro mecanismo de evasão utilizado pelo HPV, em contraste a outros vírus, é que o HPV retarda ou não lisa as células, evitando o reconhecimento pelas APCs e tem alguns de seus genes expressos no núcleo dos queratinócitos infectados, a exemplo do gene E7 do HPV16, tornando-o inacessível às APCs, além de utilizar motivos comuns a proteínas humanas de evitar o reconhecimento do antígeno (NATALE et al., 2000).

As oncoproteínas do HPV têm a capacidade de reduzir a secreção de IFN pelos queratinócitos, a exemplo do oncogene E6, que diminui a transativação do IFN- β , enquanto E7 bloqueia a expressão dos genes indutíveis de IFN- α (HASAN et al., 2007).

Os HPVs de alto risco e os de baixo risco, diferem na sua capacidade de regular negativamente a expressão das proteínas que fazem parte do MHC classe I. Recentemente, foi demonstrado que o gene E7 tem a capacidade de modificar a ação das histonas desacetilases (HELLER et al., 2011).

Diante da variedade de mecanismos de evasão imunológica, deve-se salientar que o sistema imunitário exerce uma seletiva pressão contra os diversos tipos de tumores que, por sua vez, são obrigados a desenvolver mecanismos eficazes de evasão. Em suma, esses mecanismos correspondem a uma redução da expressão do antígeno, bloqueio da apresentação dos antígenos aos componentes do complexo MHC classe I, redução da apoptose e das moléculas imunossupressoras e, finalmente, um bloqueio do recrutamento das células imunitárias (CONESA-ZAMORA, 2013).

2.6 PAPEL GERAL DAS ONCOPROTEÍNAS E1, E2, E4, E6 E E7 NA CARCINOGENESE CERVICAL

Dentre as oncoproteínas virais que se relacionam com o potencial carcinogênico aumentado dos HR-HPVs, estão as oncoproteínas E5, E6 e E7. É notória a participação destas proteínas de expressão precoce, no curso da carcinogênese, a partir de resultados obtidos *in vivo* e *in vitro*, que apontam para a

inativação de supressores tumorais, potencialização de quinases dependente de ciclinas, participação na transdução de sinais, (des)regulação de fatores envolvidos na apoptose, ativação da telomerase e de oncogenes, redução da expressão de moléculas de adesão celular, imortalização, regulação da diferenciação celular, suscetibilidade à metástase e escape imunológico (FREITAS et al., 2017).

O gene E1 é uma proteína que, nos estágios iniciais da carcinogênese, apresenta uma baixa expressão, sendo necessário, a formação de um complexo com a proteína E2, que permite, assim, a sua ligação na origem de replicação, denominada “ORI”, na LCR. Este heterodímero formado, agora, desenvolve sua atividade catalítica, atuando na síntese do DNA (BERGVALL et al., 2013).

A oncoproteína E2 atua na ativação ou na repressão dos genes de expressão precoce, competindo com fatores transcricionais pela ligação na região da origem de replicação da LCR. Desta forma, a expressão dos genes precoces, podem ser reprimidos ou ativados. A oncoproteína E2 pode ter a sua atividade celular modificada devido às mutações ou integração do genoma viral no genoma da célula hospedeira, ocasionando a diminuição da sua atividade repressora e consequente expressão dos oncogenes E6 e E7 (McBRIDE, 2013).

O papel do gene E4 no curso da infecção ainda não está totalmente elucidado. Sabe-se que a sua proteína é encontrada nas camadas mais diferenciadas do epitélio mucoso, e que se apresenta fusionada com outros cinco aminoácidos sintetizado pelo gene E1, formando uma proteína E1^{E4}, responsável por alterar a conformação dos filamentos do citoesqueleto dos queratinócitos e, assim, contribuir com a liberação das partículas virais (DOORBAR, 2013).

A proteínas E6, por sua vez, pode encontrar-se distribuída no citosol ou no núcleo dos queratinócitos infectados. Uma de suas propriedades mais bem conhecidas é de se ligar a um componente da família E3 da ubiquitina-quinase e induzir a formação de um complexo trimérico com a proteína p53, promovendo a sua degradação proteolítica e impedindo sua atividade regulatória sobre o ciclo celular.

Os mecanismos que levam um HPV de alto risco a induzir uma progressão neoplásica, a partir de uma lesão benigna, está relacionado, principalmente, à capacidade desses genótipos se integrarem ao genoma da célula hospedeira. A consequência desta sucessão de eventos abrange uma desregulação da proliferação celular, além de perdas na capacidade de reparo do DNA. Juntos,

contribuem para o acúmulo de mutações e o bloqueio da apoptose (GHITTONI et al., 2010).

2.7 ONCOPROTEÍNA E5

2.7.1 Estrutura e Aspectos Biológicos da Oncoproteína E5

A oncoproteína E5 do HPV possui um tamanho que varia entre 40 a 83 aminoácidos, e apresenta um baixo peso molecular de 10 kDa. É uma proteína de natureza hidrofóbica e com a capacidade de se ligar a diversas membranas celulares, tais como o Aparato de Golgi e o Retículo Endoplasmático, além de se inserir na membrana dos endossomos, alterando a acidez do meio (VENUTI et al., 2011).

O seu modo de ação na carcinogênese cervical ainda está sob investigação, principalmente devido a algumas características peculiares desta oncoproteína, como o seu pequeno tamanho e hidrofobicidade, que a torna uma proteína de difícil detecção, aliado aos seus baixos níveis de expressão gênica nos tecidos-alvo infectados (KAVATI, 2012; ANAYA-SAAVEDRA et al., 2013).

Nem todos os HPVs codificam a oncoproteína E5. Através das análises filogenéticas e das sequências de nucleotídeos do gene E5, várias diferentes classes da proteína E5 foram identificadas (BRAVO; ALONSO, 2004; DIMAIO; PETTI, 2013). Nas infecções por HPV16, uma proteína conservada de aproximadamente 80 aminoácidos (designada E5 α), é sintetizada, e apresenta atividade transformante em cultura de células (MICHELS; ZUR HAUSEN, 2009).

Nos HR-HPVs, 16E5 tem sido alvo de diversas análises, mas ainda, a base molecular de sua propriedade biológica não está completamente compreendida. Isoladamente, o gene E5 possui baixa ação transformante, porém quando em cooperação com outras oncoproteínas virais, a exemplo de E6 e E7, aumenta o risco carcinogênico potencializando os efeitos destes outros oncogenes, além de exercer um importante papel biológico no ciclo de replicação viral (BRAVO; ALONSO, 2004).

Nesta relação, temos o aumento da capacidade da E7 de imortalizar os queratinócitos e, conseqüentemente, um maior potencial carcinogênico, apresentando maior vantagem na infecção e maior probabilidade de transformação oncogênica (VENUTI et al., 2011; DIMAIO; PETTI, 2013).

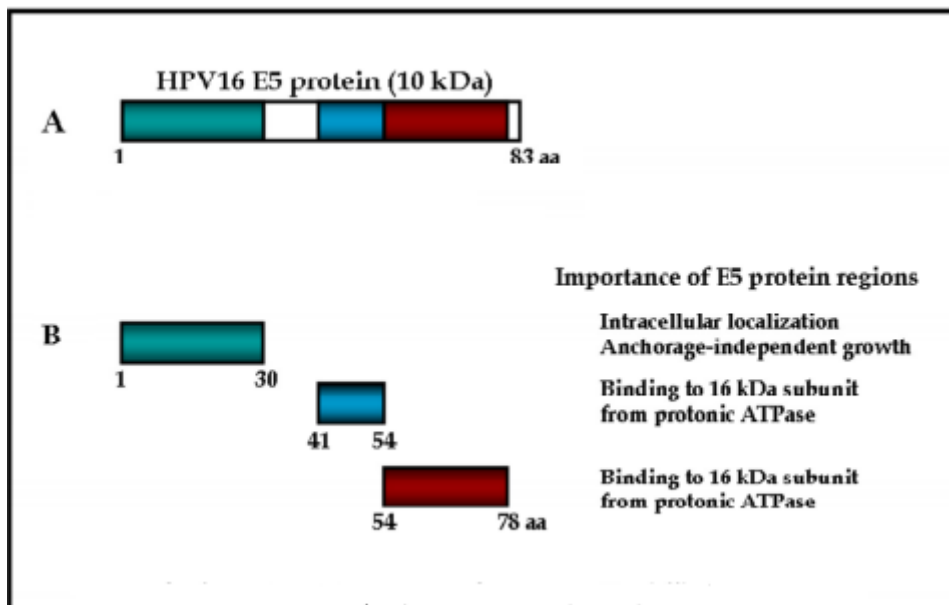
O seu pequeno tamanho e natureza hidrofóbica, faz com que seja necessário que a proteína E5 se desestruture para conseguir se ligar e interagir com suas

proteínas alvo, modificando a sua interface. Seus grupos amina e carboxila formam interações de hidrogênio, adotando uma estrutura energeticamente favorável, onde as cadeias laterais de aminoácidos ficam do lado de fora do eixo helicoidal, podendo dessa forma, interagir com os segmentos transmembranares de outras proteínas celulares (DIMAIO; PETTI, 2013).

A primeira região hidrofóbica de E5 é responsável pelo crescimento independente de ancoragem, e está associada à capacidade dessas células invadirem queratinócitos e a matriz extracelular (figura 4). Através da modelagem molecular, percebe-se que a oncoproteína E5 possui 3 hélices transmembranares e pequenas regiões hidrofílicas nas regiões C-terminal e N-terminal (figura 5) (NATH et al., 2006; WETHERILL et al., 2012; ANAYA-SAAVEDRA et al., 2013).

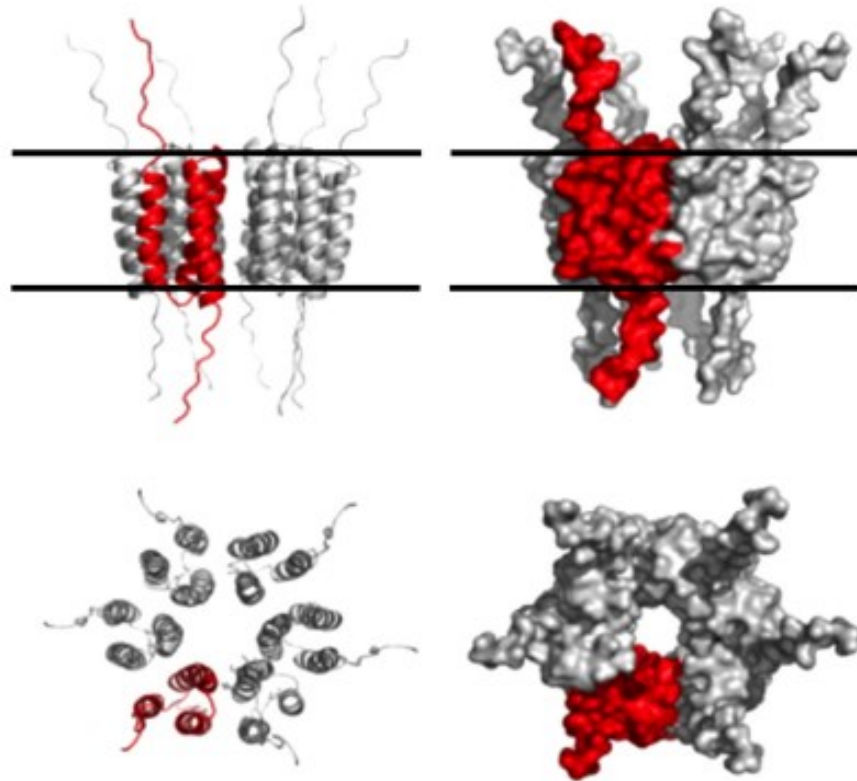
A proteína E5 não possui atividade enzimática intrínseca, pois atua na modulação da atividade de outras proteínas celulares, afetando principalmente a atividade e metabolismo dos receptores dos fatores de crescimento celular. Dentre os fatores de crescimento com metabolismo alterado pela ação da E5, se destaca o Fator de Crescimento Epidérmico (EGF) (DIMAIO; PETTI, 2013).

Figura 4 – Representação esquemática da proteína E5 com 83 aminoácidos.



Adaptado de Saavedra et al. (2012).

Figura 5 – Modelo hexamérico da proteína E5 do HPV16. A figura mostra um canal iônico gerado pela proteína E5 do HPV16 *in vitro*.



Adaptado de Wetherill et al. (2012).

Durante a multiplicação celular, uma das vias de regulação da mitose se dá através da atividade da bomba de prótons com ação ATPase. Esta bomba é responsável pela acidificação dos endossomos, através do acúmulo de íons H^+ , o que gera a degradação dos receptores de superfície dos EGFRs, e impossibilita a continuidade do ciclo mitótico. Neste contexto, a proteína 16E5 associa-se à subunidade 16kDa do vacúolo de bomba de prótons, provocando o seu bloqueio, o que gera uma inibição da acidificação do endossomos e, consequentemente, a degradação dos EGFRs, aumentando a sinalização para a mitose (VENUTI et al., 2011; KAVATI et al., 2012; ANAYA-SAAVEDRA et al., 2013).

Esse mecanismo de sinalização proporciona o aumento do tempo de meia-vida dos EGFRs, bem como o de outros receptores de fatores de crescimento. Assim, atribui-se à proteína E5 a capacidade de estimular a proliferação celular através dos EGFRs e a expansão dos queratinócitos infectados (VENUTI et al., 2011)

Em contraste, outros estudos sugerem que a interação da E5 do HPV16 com a subunidade 16kDa ocorre em níveis baixos, e que a oncoproteína E5 desempenha seu papel no tráfico endocítico, não por acidificação endossomal, mas

principalmente por inibição da fusão das vesículas endossomais, independentemente da subunidade 16kDa (SUPRYNOWICZ et al., 2010).

Outra via importante de ação da proteína 16E5 está em evitar a degradação dos EGFRs, através da interação com uma ubiquitina-ligase, a c-Cbl. A c-Cbl é responsável por controlar negativamente os receptores de membrana, através de sua ligação com a tirosina-quinase dos receptores, sinalizando-os para ubiquitinação e degradação proteossômica. 16E5 inibe a interação do EGFR com a c-Cbl, diminuindo dessa forma a ubiquitinação dos receptores, e mantendo a sinalização para a constante transdução de sinal (DIMAIO; PETTI, 2013).

A via de sinalização do EGFR também pode ser ativada por E5 de forma independente de EGF, quando as células são colocadas sob condição de estresse. Nesta situação, 16E5 ativa proteínas quinases como MAPK, p38 e ERK 1/2. As MAPKs aumentam a transcrição dos oncogenes c-fos e c-jun, que estimulam a transcrição de E6, E7 e de outros genes envolvidos no crescimento celular, através da formação do complexo AP-1, um fator de transcrição essencial para o ciclo do HPV e crescimento celular. A ligação de E5 com outras proteínas membranares, como EVER1 e EVER2, também aumenta a quantidade de AP-1, e consequentemente de c-fos e c-jun, podendo gerar mutação destes genes e levando à tumores de pele (PATEL et al., 2008; SAAVEDRA et al., 2012).

16E5 possui a capacidade de modificar a composição lipídica e a dinâmica das membranas celulares, através do aumento da expressão dos componentes lipídicos caveolina-1 e gangliosídeo GM1. A associação caveolina-1/GM1 afeta o tráfico vesicular, sinalização celular e aumenta a resposta proliferativa de EGF (SUPRYNOWICZ et al., 2010).

Dentre outras funções celulares atribuídas à oncoproteína E5 estão: aumento da expressão dos genes COX2 (ciclo-oxigenase 2), aumento dos receptores EP4 de prostaglandinas, o que leva a um aumento na expressão do Fator de Crescimento Endotelial Vascular (VEGF), necessários para o crescimento do tumor, angiogênese e metástase (ANAYA-SAAVEDRA et al., 2013).

Há ainda, os receptores de endotelina acoplados à proteína G (ETA), que assim como as EGFRs possuem atividade mitogênica. A E5 aumenta a atividade de ETA através da indução de seu ligante específico endotelina-1 (ET-1), aumentando a sinalização mitogênica dos queratinócitos. Outro mecanismo que a E5 utiliza para

desregular o crescimento celular é através da inibição das proteínas p21 e p27, proteínas envolvidas na inibição do ciclo celular (DIMAIO; PETTI, 2013).

A proteínas E5 dos HPVs alto risco também está envolvida em alterações morfológicas, como aumento do tamanho nuclear, aumento do conteúdo do DNA, poliploidia e binucleação (fusão célula-célula), e instabilidade cromossômica (DIMAIO; PETTI, 2013; VENUTI et al., 2011).

Alguns estudos sugerem que cerca de 10% do total de 16E5 em queratinócitos imortalizados (HaCat) se encontram na membrana plasmática, apresentando sua parte n-terminal intracelular, e a c-terminal extracelular, o que parece estar envolvido no processo de fusão entre as células e formação de HaCats binucleados. Contraditoriamente, outros estudos citam a presença de 16E5 apenas no retículo endoplasmático, e a região n-terminal sendo extracelular, e a c-terminal como intracelular. A hipótese da região c-terminal ser intracelular é que cerca de 10 aminoácidos dessa região, interagem com a KN β 3, uma proteína citoplasmática. O complexo KN β 3/E5 está envolvido no processo de tráfico de vesículas (VENUTI et al., 2011).

Sugere-se que a função primária de E5 é inibir a apoptose em queratinócitos nos estágios iniciais da infecção, através da destruição do ligante do fator de necrose tumoral (*FasL*) e/ou do ligante do fator de necrose tumoral relacionado à indução de apoptose (TRAIL). De fato, queratinócitos que expressam 16E5 são imunes à ação anti-apoptótica destes dois fatores (VENUTI et al., 2011).

2.7.2 Via Celular de E5 durante a replicação viral e atividade oncogênica

Os HPVs têm como principal característica a infecção do epitélio estratificado e o seu ciclo de vida, está intimamente associado com os mecanismos de diferenciação dos queratinócitos. Até o estabelecimento de uma infecção persistente, os HPVs enfrentam e precisam ultrapassar inúmeras barreiras encontradas no epitélio. Dentre elas, estão: a) uma descamação contínua do epitélio estratificado, por isso a necessidade que os vírus apresentam de invadir as camadas suprabasais para a propagação das células infectadas; b) os papilomavírus se utilizam da maquinaria celular para a sua própria replicação e precisam manter os mecanismos de divisão celular em pleno funcionamento, juntamente com uma inibição, embora não seja completa, dos mecanismos de diferenciação, evitando a

ocorrência da apoptose; e c) para garantir a sobrevivência e a infecção persistente, os HPVs precisam driblar o sistema imunológico (VENUTI et al., 2011).

A conservação do gene E5 pelos HR HPVs sugere que ele é requerido para os estágios iniciais de infecção e para a progressão do ciclo de vida dos vírus durante a infecção, o que é notado através da análise dos níveis de expressão do oncogene E5 aumentados durante a diferenciação das células suprabasais do epitélio estratificado que se encontra infectado pelo HPV (Longworth e Laimins, 2004). Este aumento nos níveis de expressão também é observado em culturas de queratinócitos humanos em diferenciação, contendo o genoma do HPV16 e HPV31. Com estes níveis elevados, teremos a expressão dos genes L, que sintetizarão o capsídeo viral (FEHRMANN et al., 2003; GENTHER et al., 2003).

Alguns estudos genéticos têm indicado que a interação entre E5/Bap31 é essencial para a formação de colônias de queratinócitos em diferenciação e que a proteína E5 é capaz de conservar esta capacidade proliferativa, e isto mantém, presumivelmente, estas células num estado propício à replicação do DNA viral. Em queratinócitos humanos imortalizados, 16E5 inibe a síntese de DNA em resposta à ativação do receptor do fator de crescimento dos queratinócitos e o receptor do fator de crescimento de fibroblastos (KGFR/FGFR2) (BELLEUDI et al., 2011).

Durante a infecção, as proteínas E5 cooperam com E6 para induzir a formação de vacúolos perinucleares, bastante característico das infecções por HPV no epitélio estratificado, o que culmina com o aumento da fragilidade celular, facilitando a entrada dos vírus (KRAWCZYK et al., 2011).

É notável, então, o papel essencial desenvolvido pela proteína E5 e a sua colaboração com outras proteínas virais no desenvolvimento de estratégias ao longo do ciclo de infecção e da replicação viral.

A principal atividade transformante dos HPVs reside nos genes E6 e E7, responsáveis pela imortalização dos queratinócitos e potente atividade oncogênica. Em cultura de células E5 do HPV apresenta uma fraca atividade transformante e requer a presença de algum cofator que potencialize a sua atividade. Em níveis fisiológicos de expressão gênica, 16E5 localiza-se primariamente na membrana do Retículo Endoplasmático, enquanto que, em níveis elevados de expressão, a oncoproteína passa a estar presente na membrana do Complexo de Golgi, Membrana Plasmática e Envelope Nuclear (SUPRYNOWICZ et al., 2010).

Em relação à atividade transformante *in vitro*, dos HR-HPVs, E5 pode representar um importante agente no curso da carcinogênese. Análise de variantes de 16E5 revelou um poder aumentado sobre a atividade mitogênica, com alto potencial carcinogênico (NATH et al., 2006; BRAVO e ALONSO et al., 2004; (SCHIFFMAN et al., 2005).

Em HPV18, o oncogene E5 é frequentemente deletado ao longo da integração do DNA viral no genoma da célula hospedeira e a sua detecção torna-se dificultada. Contudo, o genoma viral persiste na forma episomal e 16E5 pode ser detectado em lesões pré-malignas e malignas em amostras HPV16 positivas, sugerindo a participação deste gene na progressão da malignidade (CULLEN et al., 1991; SAHAB et al., 2012). Dessa forma, E5 contribui para nos estágios de promoção e progressão da carcinogênese (MAUFORT et al., 2007).

2.7.3 Oncoproteína E5 e a evolução dos HPVs

Os HPVs são, atualmente, classificados em cinco gêneros, denominados Alpha, Beta, Gamma, Mu e Nu. Nem todos os gêneros codificam a proteína E5, sendo esta função muito bem reconhecida nos HPVs do gênero Alphapapillomavirus (BERNARD et al., 2010).

Análises filogenéticas sugerem, até o presente, três papilomavírus ancestrais responsáveis pela heterogeneidade do grupo dos HPVs que provocam alterações genitais. Dentre os grupos atuais que emergiram destes ancestrais, estão os alphapapillomavírus, todos derivados de um ancestral em comum, com potencial oncogênico, e que codifica a proteína E5. As variantes de 16E5 possuem uma alta atividade mitogênica e é um dos mais frequentes detectados nas populações e associados às lesões cervicais (NATH et al., 2006).

A ORF do gene E5 do HPV tem sido classificada em quatro diferentes grupos, denominados alpha, beta, gamma e delta, cada um relacionado com um determinado tipo de manifestação clínica, em particular, com o potencial oncogênico (BRAVO e ALONSO, 2004). Assim, E5-alpha é codificado pelos alpha-HPVs de alto risco, enquanto que o E5-gamma e E5-delta, são codificados por HPVs de baixo risco (GARCÍA-VALLVÉ et al., 2005).

A proteína E5 também pode ser responsável por dar alguma vantagem na expressão dos vírus ao longo da infecção. Análises filogenéticas têm mostrado que as variantes destas proteínas parecem aumentar a transformação oncogênica

seguido do aumento das infecções persistentes. No entanto, a ORF de E5 está ausente no genoma de muitos HPVS, como o *beta*, o *gamma* e o *mu*, indicando que esta proteína não é essencial para o ciclo de vida destes vírus. Sua ORF é expressa durante as fases iniciais do ciclo de vida viral, na forma de um transcrito policistrônico (LONGWORTH; LAIMINS, 2004).

2.7.4 Oncoproteína E5 e resposta imunológica

A frequência de infecções pelo HPV e a progressão para o câncer cervical é dependente de fatores imunológicos: mulheres imunocompetentes tendem a eliminar o vírus, ao passo que mulheres imunossuprimidas tendem a persistir com o vírus, podendo desenvolver lesões que podem progredir, levando ao processo neoplásico. Fica claro dessa forma, a importância do envolvimento da resposta imune na eliminação da infecção viral (CAMARA et al., 2003; MONFRÉ, 2011).

A principal resposta imune utilizada contra o vírus é a Th1 (celular), que secreta IFN- γ e ativa os macrófagos. O MHC de classe I expresso na superfície das células, quando associado a antígenos intracelulares, ativa os linfócitos T CD8+ e sua consequente ação citotóxica. Paralelamente, o aumento da resposta de linfócitos T CD4+ está diretamente relacionado com a eliminação do HPV e condilomas (MONFRÉ, 2011).

O HPV age por diferentes mecanismos para escapar da resposta imune, entre eles, o fato do vírus não causar a lise das células, e dessa forma, não haver o recrutamento dos macrófagos (VENUTI et al., 2011). Tumores cervicais e lesões de alto grau estão associadas com a diminuição da resposta Th1, e consequentemente, de IFN- γ , IL-2 e TNF- α . Dessa forma, percebe-se que as concentrações de IL-2 e IFN- γ presentes nos estágios iniciais do tumor cervical, são suprimidos por IL-4 e IL-10 em estágios mais avançados (MONFRÉ, 2011).

A proteína E5 auxilia na evasão imune promovendo a diminuição da sinalização de infecção intracelular, através de interação com a cadeia pesada da molécula de MHC classe I, internalizando seu antígeno HLA I no aparelho de Golgi. Outra forma importante de reter o MHC de classe I com seu respectivo HLA no aparelho de Golgi é através da alcalinização dos compartimentos endomembrana (VENUTI et al., 2011).

Uma vez acumulado no complexo de Golgi, o MHC classe I não é expresso na superfície celular, e, portanto, as células infectadas acabam por não ser

apresentadas aos linfócitos T CD8+, e não ocorre a resposta adaptativa. (DIMAIO e PETTI, 2013; VENUTI et al., 2011). A oncoproteína E5 não inativa todos os tipos de HLA I: apenas inativa HLA-A e HLA-B, permitindo a expressão normal de HLA-C e HLA-E (HLA-E não desempenha função contra a infecção por HPV), a fim de evitar a ação das células NK, que costumam atacar células que não expressam MHC e HLA. A inibição do HLA de classe I pode ser reversível através da administração de Interferon (MONFRÉ, 2011).

Dessa forma, a fim de evitar a reativação do HLA I, as proteínas E6 e E7 são capazes de inibir o Interferon. Enquanto isso, 16E5 interfere na maturação de MHC classe II, também impedindo sua expressão na superfície celular. Porém, diferentemente do HLA I, a internalização do HLA II é irreversível (DIMAIO e PETTI, 2013; VENUTI et al., 2011). E5 inibe ainda citocinas dependentes de CD1d, através do tráfico de CD1d e posterior degradação por vias proteolíticas, afetando negativamente a imunidade inata e adaptativa (VENUTI et al., 2011).

2.7.5 Importância da oncoproteína E5 como estratégia imunoterapêutica

Como visto, E5 atua em diversas etapas do processo carcinogênico, afetando e desregulando diversas moléculas, e atuando em conjunto com outras proteínas, como E6 e E7. Além disso, E5 também se associa a proteína E6, aumentando a fragilidade dessas células e podendo formar vacúolos coilócíticos (coilócitos). Tanto os vacúolos perinucleares quanto os coilócitos são importantes características de infecção por HPV, levando ao diagnóstico da infecção (DIMAIO; PETTI, 2013).

De acordo com diversos estudos, há a possibilidade de se utilizar a E5 como alvo na terapia de câncer cervical, assim como as vias alteradas por essa proteína (VENUTI et al., 2011). Em modelos animais que receberam vacinas constituídas por epítomos de E5, houve um aumento de células T CD8+, conferindo proteção contra tumores e redução do crescimento tumoral. Por se tratar de uma proteína expressa em estágios iniciais da infecção e das lesões, possui potencial terapia contra a infecção pelo HPV e lesões pré-cancerosas, prevenindo contra a progressão para o câncer invasivo (VENUTI et al., 2011).

Diversos estudos têm avaliado a interação entre os HR-HPVs e o sistema imunológico do hospedeiro, objetivando ampliar diversas as estratégias terapêuticas e modular a resposta imune no combate à infecção (VICI et al., 2014).

Atualmente, não temos medicamentos antivirais contra as infecções provocadas pelo HPV, sendo necessário o tratamento via quimioterapia, radioterapia, cirurgia ou histerectomia. Embora existam essas possibilidades, podemos ter uma alta recorrência e um aumento no número de mortes (WHO, 2016).

Novas infecções por HPV podem ser prevenidas com a aplicação de vacinas profiláticas recentemente introduzidas, mas estas vacinas não têm um efeito terapêutico nos milhões de infectados com o HPV. O tratamento torna-se limitado a estes procedimentos intervencionistas e, portanto, o desenvolvimento de um agente antiviral eficaz e específico é extremamente desejável (WETHERILL et al., 2012).

O câncer associado ao HPV é frequentemente associado à integração do genoma do HPV nos cromossomos do hospedeiro. Em muitos casos, isso resulta em perda do E5 ORF. Entretanto, dado que a análise de mRNA indica alta expressão de E5 em lesões pré-cancerosas (CHANG et al., 2001).

Um recente estudo mostra que a oncoproteína E5 funciona como um canal iônico e esta atividade biológica pode ser explorada com uma finalidade terapêutica (WETHERIL et al., 2012). Canais iônicos codificados por vírus (vioporinas), já mostraram ser alvos de drogas viáveis, a exemplo do canal M2 do vírus influenza A, que o torna suscetível à inibição pelas drogas antivirais amantadina e rimantadina. No contexto das infecções provocadas pelo HPV16, o alvo E5 também foi relatado como suscetível à inibição por estes compostos, bem como um número de outras moléculas inibidoras de vioporina reconhecidas (SCHNELL; CHOU, 2008).

Em relação ao desenvolvimento de estratégias vacinais, duas vacinas profiláticas (bivalente e quadrivalente) foram produzidas para prevenir a infecção pelo HPV (MARIANI e PAGLIUSI, 2008), e recentemente foi aprovada uma vacina nonavalente, que protege de nove genótipos de HPV (6, 11, 16, 18, 31, 33, 45, 52, e 58) (JOURA et al., 2015). Estas vacinas geralmente visam ativar células imunes, seja de resposta inata ou adaptativa (ROSALES, 2014), a fim de simular a resposta imune natural e reverter os mecanismos imunomoduladores virais.

As oncoproteínas do HPV têm sido amplamente empregadas na imunoterapia do câncer cervical e os oncogenes E6 e E7 foram amplamente testados. No entanto, os estudos demonstram que as estratégias adotadas visando estas duas oncoproteínas ainda estão longe de evitar as lesões pré-cancerígenas. Dessa forma, os estudos sobre a oncoproteína E5 vem se tornando uma alternativa (LIAO et al.,

2015), pois visa estimular a resposta imunológica do hospedeiro no combate a infecção viral e progressão das lesões.

Dentre as metodologias testadas para o E5, as vacinas vetoriais virais são uma estratégia terapêutica promissora para a apresentação de antígenos patogênicos para sensibilizar o sistema imune do hospedeiro, induzindo forte resposta citotóxica. Estas vacinas são altamente imunogênicas, pois são capazes de penetrar nas células hospedeiras onde sintetizam os antígenos de interesse causando o rompimento das células infectadas (ROSALES, 2014).

Um dos primeiros estudos a avaliar a vacina terapêutica HPV16 E5 utilizou a bioengenharia de adenovírus recombinante para expressar E5 em modelo murino, cujo resultado observado foi a redução do crescimento do tumor *in vitro* e, além disso, a resposta imunológica foi T CD8+ dependente e T CD4+ independente, sugerindo sua ação antigênica com capacidade de reduzir o tumor (LIU et al., 2000).

Numa vacina de DNA que codifica as oncoproteínas de HPV16, E5 foi associado a E6 e E7, e foi geneticamente fundido à glicoproteína D do vírus do herpes. Esta vacina mostrou uma resposta T CD8+ específica para cada oncoproteína individualmente num modelo murinho, além de exibir um maior efeito anti-câncer quando combinadas, por exemplo, com IL-12 (DINIZ et al., 2010). Essas citocinas são co-estimuladoras da resposta imune e recrutam células centrais do sistema inato, como as células dendríticas e as *Natural Killers* (NK), o que resulta na maturação da resposta dos linfócitos T citotóxicos, por exemplo (PERALTA-ZARAGOZA et al., 2012). IL-12 atuando como um co-estimulador foi utilizado com sucesso em ensaios pré-clínicos e clínicos da terapia do câncer do colo do útero (PERALTA-ZARAGOZA et al., 2012; VICI et al., 2014).

Outros estudos mostram que a resposta imune mediada por células T, induzida pela oncoproteína E5 de HPV16, pode ser influenciada pelo HLA-A. O epítipo HPV16 E5 reduziu o potencial oncogênico e gerou células de memória T, restritas ao HLA-A*0201 (LIU et al., 2007).

A utilização de potentes peptídeos pode ser útil para induzir um hospedeiro à resposta imune contra o vírus, como mostrado *in silico* para a oncoproteína E5 (ROSALES, 2014). O estudo *in silico* avaliou os epítopos mais potentes da oncoproteína HPV16 E5, que eram capazes de estimular as atividades das células T e B e também eram totalmente imunogênicas através do MHC classe I e II (KUMAR et al., 2015).

Outras estratégias de vacina para o E5 permanecem não testadas. Um exemplo é o uso da vacina autóloga de células dendríticas (DC). Neste estudo, os monócitos autólogos foram diferenciados em DC *in vitro* que foram, em seguida, administradas de volta ao paciente após serem carregadas com antígenos de HPV 16/18, em sistema cultivado com a oncoproteína E7. Os resultados apontaram para um aumento na atividade das células T CD8+ e secreção de IFN- γ . DC é conhecida como a célula mais potente para a indução de CTL, e o uso de oncoproteínas HPV16/18 como antígenos para o desenvolvimento de vacinas DC mostraram bons resultados, mas não foram realizados ensaios para a oncoproteína E5 (LIU et al., 2016).

O uso de E5 por meio dessas estratégias deve fornecer mais conhecimento sobre a atividade E5 como um antígeno e ajudar na busca de uma terapia apropriada contra o câncer cervical. De uma forma geral, essas vacinas de E5 podem representar uma boa estratégia, embora mais estudos sejam necessários para validar este alvo molecular como uma boa ferramenta imunoterapêutica (PALADINI et al., 2016).

2.8 EPÍTOPOS: VISÃO GERAL

Os Epítomos ou determinantes antigênicos são sequências formadas por resíduos de aminoácidos importantes encontrados no antígeno que participam das interações antígeno-anticorpo (PELLEQUER et al., 1993). Estas sequências são reconhecidas pelos anticorpos com especificidade variável e quando esta interação é formada, tanto o anticorpo quanto o antígeno sofrem mudanças estruturais que podem neutralizar o antígeno, facilitando o seu processamento pelo sistema imune (MOYLE; TOTH, 2013).

Assim, é de fácil compreensão o fato de que os epítomos estão no centro da resposta imune humoral (WILSON, 2012). Entender detalhadamente suas propriedades é de suma importância para que sejam identificados de maneira rápida e eficiente. Com isso, é possível o desenvolvimento de novas terapias, tais como o desenvolvimento de vacinas, além de ser importante no diagnóstico, isolamento e caracterização de moléculas associadas com vários estágios da doença (PARREN et al., 2000).

Essa capacidade que o antígeno possui de provocar uma resposta imune no organismo onde está atuando é denominada imunogenicidade. Neste mecanismo,

várias moléculas são acionadas, afim de que o processamento e a apresentação dos antígenos ocorram diferencialmente, sendo reconhecidos por células T ou pelas células B (HUANG et al., 2007).

A maioria das células somáticas nucleadas apresentam moléculas MHC classe I, ligadas a epítomos, que são apresentadas para as células T citotóxicas induzindo apoptose nas células infectadas ou doentes. O número de aminoácidos nos epítomos apresentados pelas moléculas MHC classe I se encontra entre 8 e 11, principalmente apresentados a células T CD8+ ou citotóxicas (HUANG et al., 2007). As moléculas MHC classe II apresentam epítomos com tamanhos entre 13 a 17 aminoácidos para células T CD4+ ou Helper as que participam ativando células B (ROCHA; NEEFLES, 2008).

A região específica do anticorpo que se une a um epítomo denomina-se paratopo. Este pode estar formado por sequências muito variáveis de resíduos de aminoácidos que se encontram fixos nos esqueletos dos anticorpos (PARREN et al., 2000). Esta particularidade permite ao sistema imune reconhecer uma ampla quantidade de antígenos (SUBRAMANIAN; CHINNAPPAN, 2013).

Atualmente, ainda são estudadas as propriedades que influenciam a antigenicidade e imunogenicidade de um antígeno, assim como as diferentes interações entre anticorpo-antígeno (VIART et al., 2016), o que poderá contribuir para um melhor desenho experimental na produção de métodos imunoterapêuticos e moléculas para diagnóstico.

Alguns bancos de dados, como o *Protein Data Bank* (PDB) e o *Immune Epitope Data Bank* (IEDB), disponibilizam estruturas de complexos antígeno-anticorpo com informações úteis para os modelos computacionais. Mesmo com essa disponibilidade de dados, vários grupos de pesquisa têm adotado a análise de vários parâmetros e métodos matemáticos, considerando, por exemplo, a sequência do antígeno, as propriedades físico-químicas, a posição dos aminoácidos e conservação dos resíduos, dentre outros (CHEN et al., 2007; WEE et al., 2010).

No que se refere aos estudos *in silico* para os Papilomavírus, Kumar et al. (2005), utilizando a ferramenta IEDB, identificou epítomos imunogênicos relacionados à oncoproteína E5 do HPV16. Neste estudo, foi utilizado a sequência de referência K02718, disponível no NCBI. Foram identificados um total de 57 epítomos preditos para os alelos do MHC-I, com comprimento variável entre 9-12 resíduos de aminoácidos, onde a maioria pertenceu às classes HLA-A e HLA-B. Em

relação à análise conduzida para a identificação de epítomos relacionados aos alelos MHC-II, os autores identificaram um total de 257 epítomos pertencentes, principalmente, aos alelos HLA-DRB1*10:01; HLA-DRB1*14:01 e HLA-DRB1*16:01. Este estudo contribuiu para a identificação *in silico* de potentes epítomos que pode ser útil no entendimento das vias de ativação das células T e B, além de contribuir para o desenvolvimento de estratégias imunoterapêuticas.

A identificação de motivos de seqüência curta de proteínas patogênicas que são capazes de se ligar especificamente ao HLA é reconhecida como a etapa crucial do desenvolvimento de vacinas de peptídeos imunoterapêuticos. No entanto, é muito demorado e caro sintetizar todos os candidatos a peptídeos encontrados no genoma do patógeno e realizar o ensaio de ligação peptídica para identificar ligações HLAs de alta afinidade (ITO et al., 2002).

Alternativamente, a abordagem computacional fornece uma maneira promissora de rastrear cadeias de peptídeos em proteínas codificadas por vírus com alta afinidade para o HLA antes de realizar testes experimentais exaustivos, o que justifica a adoção da presente metodologia.

2.9 APLICABILIDADE DAS FERRAMENTAS APLICADAS NA ANÁLISE *IN SILICO* PARA A GERAÇÃO E INTERPRETAÇÃO DE DADOS BIOLÓGICOS

2.9.1 Fasta e Blast

Os métodos FASTA (*Fast Alignment*) e o BLAST (*Basic Local Alignment Search Tool*), correspondem a dois algoritmos heurísticos que são freqüentemente usados para busca em bancos de dados de seqüências de DNA e proteína. Um método é denominado heurístico por se basear numa estratégia de particionamento de uma sequência em partes menores de caracteres consecutivos. Esses métodos são rápidos e analisam regiões de alinhamentos potenciais entre duas sequências. (BAXEVANIS; OUELLETTE, 2001).

O algoritmo FASTA, desenvolvido por Pearson e Lipman, em 1985, de uma maneira geral, é um algoritmo utilizado para alinhamentos rápidos de pares de sequência de DNA e de proteína, tentando aproximar os resultados do melhor alinhamento local e os escores produzidos (BAXEVANIS; OUELLETTE, 2001). O software procura por identidade de padrões objetivando construir um alinhamento local baseado nestas identidades.

Adicionalmente, após o desenvolvimento da ferramenta BLAST, nos anos 1990, essa metodologia tornou-se a mais utilizada para a busca de sequências em bancos de dados, devido à sua velocidade de processamento e ao fato de que cada identidade é acompanhada de uma significância estatística (GUSFIELD, 1997). O BLAST reúne os principais programas denominados *blastp* (compara sequências de aminoácidos), *blastn* (compara duas sequências de nucleotídeos), *blastx* (compara uma sequência de nucleotídeo com uma sequência de proteína), *tblastn* (compara uma sequência de proteína com uma sequência de nucleotídeo) e *tblastx* (compara os *frames* traducionais de uma sequência de nucleotídeos com os *frames* traducionais de outra sequência de nucleotídeos do banco de dados).

2.9.2 Alinhamento múltiplo

O estudo das sequências gênicas é útil para analisar as relações evolutivas, bem como as restrições estruturais e/ou funcionais presentes no DNA, RNA ou na Proteína. Levando-se em conta estes fatos, os programas de alinhamento são empregados para identificar, por exemplo, sequências homólogas distantes, baseados nas suas semelhanças, isto significa que, sequências homólogas, que descendem de um ancestral comum, podem contribuir na construção da história evolutiva da sequência que está sendo analisada, e do organismo com um todo (ROCHA, 2000).

A utilização do alinhamento múltiplo reside, principalmente: (1) para o estudo da história evolutiva das espécies, (2) para relacionar funcionalmente e estruturalmente as sequências analisadas, podendo-se inferir novos dados e (3) para analisar quais as partes da sequência são similares e quais as partes são diferentes. Por isso, tem ampla utilização no estudo da filogenia.

2.9.3 Análise das sequências gênicas do HPV16 e HPV31

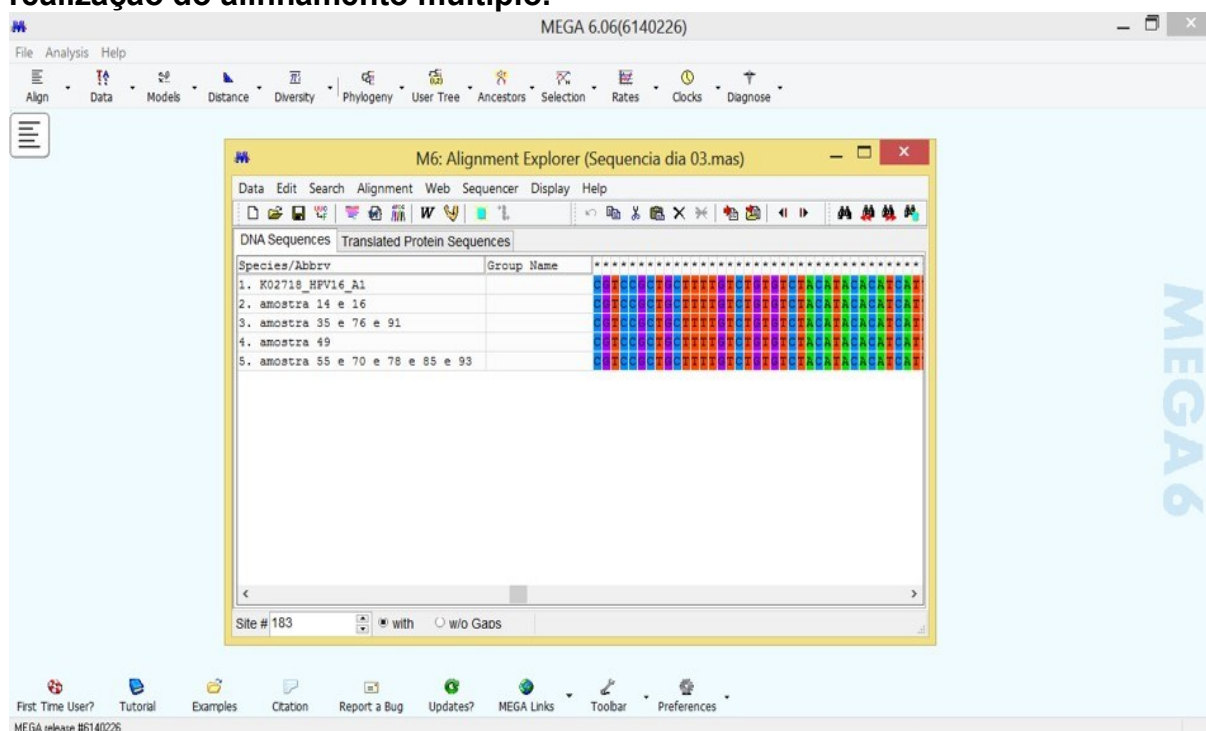
Para a análise dos cromatogramas, foram utilizadas as ferramentas Pregap4 (versão 1.5), responsável pela análise dos cromatogramas e geração dos dados de entrada para o Gap4 (versão 4.0), que será utilizado para a montagem do genoma. Ambas as ferramentas são pertencentes ao pacote de softwares Staden (STADEN, 1996). Com a montagem dos cromatogramas, obedeceu-se a um critério de qualidade PHRED $\geq$ 30, que trata de um algoritmo de análise da qualidade para cada

base da sequência, e esta “nota de qualidade”, em escala logarítmica, simboliza que a probabilidade de erro é de 1:1000, ou seja, 0,1% (EWING et al., 1998).

Em seguida, as sequências foram submetidas a um alinhamento o programa CLUSTALW, presente no pacote de ferramentas do software *Molecular Evolutionary Genetics Analysis* 6 (MEGA6), desenvolvido para análise comparativa de DNA e proteína, e útil para inferência de padrões evolutivos entre os genes e os genomas de diversas espécies (TAMURA et al., 2013).

O software de Análise Molecular de Genética Evolutiva (MEGA) foi desenvolvido com a finalidade de realizar análises comparativas de sequências de DNA e proteínas e, assim, poder inferir os padrões evolutivos de genes, genomas e espécies (Kumar et al. 1994; Tamura et al. 2011). A versão utilizada também permite a construção de árvores evolutivas moleculares, que são claramente úteis por tratar dos tempos de divergência sofrida pelas espécies, por exemplo (KUMAR; HEDGES 2011; WARD et al. 2013) (figura 6).

Figura 06 – Representação esquemática do software MEGA6 utilizado para a análise e identificação das sequências das variantes do oncogene E5 dos HPVs 16 e 31. O pacote MEGA6 inclui o programa ClustalW, que permite a realização do alinhamento múltiplo.



2.9.4 PSIPRED: *Protein Sequence Analysis Workbench*

Na análise de uma estrutura protéica, a presença de domínios similares em diversas seqüências, implica numa função bioquímica ou enovelamento estrutural semelhantes e que podem se tornar a base de investigações ou experimentos mais aprofundados, por exemplo, no que concerne a uma origem evolutiva comum. Sendo assim, o objetivo da predição de estrutura de proteína é tomar uma estrutura primária (seqüência) e, usando regras derivadas de estruturas conhecidas, predizer a estrutura secundária mais provável a ser adotada individualmente pelos resíduos.

Atualmente, várias abordagens computacionais têm sido adotadas para a predição da estrutura e função de uma proteína. Essas ferramentas contribuem para anotação de uma estrutura protéica hipotética. Nesse contexto, o uso de ferramentas computacionais que agregam diferentes métodos computacionais de análise, dão um maior suporte para a geração de modelos mais adequados (DORDEN; MAHADEVAN, 2015), contribuindo para a dedução das possíveis posições das alfa-hélices e das relações funcionais ou estruturais, por exemplo.

O software PSIPRED incorpora três diferentes métodos em sua interface, sendo representados pelo PSIPRED (JONES 1999a), GenTHREADER (JONES 1999b) e o MEMSAT 2. Estas ferramentas reunidas, tornam a análise mais robusta da estrutura secundária da proteína, tornando os resultados mais confiáveis. O método PSIPRED corresponde a uma ferramenta de alta acurácia para a predição da estrutura secundária de uma proteína, utilizando gráficos para a representação dos resíduos de aminoácidos da proteína. Outro método adotado, o MEMSAT 2 prevê a estrutura secundária e a topologia de proteínas integrais de membrana, reconhecendo modelos topológicos. O método GenTHREADER pode ser utilizado para a análise de individual de proteína, analisando as relações existentes entre as superfamílias de proteínas e suas relações ancestrais (figura 7) (McGUFFIN et al., 2000).

O uso desta ferramenta, como um importante veículo de análise da estrutura protéica, pode ser observado, na análise da estrutura secundária das isoformas das oncoproteínas b2a2 e b3a2, duas p210^{BCR-ABL}, relacionadas à translocação entre os cromossomos 9 e 22, que gera um isocromossomo, manifestando o quadro de Leucemia Mielódica Crônica (HAI et al., 2014). Os resultados obtidos, mostraram uma estrutura secundária com variações nos resíduos de aminoácidos entre as duas isoformas estudadas, e que estas isoformas podem interferir nos mecanismos de regulação da atividade tirosina quinase da proteína ABL. Dessa forma, ambas as

proteínas p210^{BCR-ABL} podem causar um efeito pleiotrópico em várias vias de transdução de sinais que afetam a sobrevivência celular, a progressão da doença e a estabilidade do genoma (HAI et al., 2014).

Estes resultados observados por Hai et al (2014), reforçam a importância na análise *in silico* e a necessidade de mais estudos sobre o papel biológico da oncoproteína E5, correlacionando-o com possíveis efeitos biológicos exercidos no tecido-alvo infectado, visto que mudanças conformacionais em sua estrutura, poderão sinalizar, para o surgimento de múltiplos epítomos imunogênicos, por exemplo.

Figura 07 – Layout do software PSIPRED utilizado para a análise e construção da estrutura secundária da oncoproteína E5 dos HPVs 16 e 31.

UCL Department Of Computer Science

Bioinformatics Group

Site Navigation

- Introduction
- People
- Projects
- Publications
- Web Servers
- Downloads
- Vacancies
- Contact
- Group Intranet

Server Navigation

- PSIPRED Server
- Server Overview
- Server Citation
- Help & Tutorials
- News
- History
- Software Download

The PSIPRED Protein Sequence Analysis Workbench

The PSIPRED Protein Sequence Analysis Workbench aggregates several UCL structure prediction methods into one location. Users can submit a protein sequence, perform the predictions of their choice and receive the results of the prediction via e-mail or the web.
For a summary of the available methods you can read [More...](#)

NOTE: users who need to run our methods on a large number of proteins should consider downloading our software using the menu on the left (Server Navigation -> Software Download).

The PSIPRED Team
Current Contributors David T. Jones, Daniel Buchan, Domenico Cozzetto & Kevin Bryson
Previous Contributors Tim Nugent, Federico Minneci, Anna Lohley, Sean Ward, Liam J. McGuffin

For queries regarding PSIPRED: psipred@cs.ucl.ac.uk

Input Sequence Filter

Choose Prediction Methods

☒ PSIPRED v3.3 (Predict Secondary Structure) ☐ pGenTHREADER (Profile Based Fold Recognition) ☐ BioSerf v2.0 (Automated Homology Modelling) ☐ FFPred 3 (Eukaryotic Function Prediction) ☐ MEMPACK (SVM Prediction of TM Topology and Helix Packing) ☐ DomSerf v2.0 (Automated Domain Modelling by Homology) Help...	☐ DISOPRED3 (Disorder Prediction) ☐ MEMSAT3 & MEMSAT-SVM (Membrane Helix Prediction) ☐ DomPred (Protein Domain Prediction) ☐ GenTHREADER (Rapid Fold Recognition) ☐ pDomTHREADER (Fold Domain Recognition)
---	---

Input Sequence (Single sequence or Multiple Sequence alignments; as raw sequence or fasta format)

Help...
If you wish to test these services follow this link to retrieve a test fasta sequence.

Submission Details

Email Address for job completion alert (optional)

Help...

Password (only required for licenced commercial e-mail addresses)

Help...

Short identifier for submission

Help...

2.9.5 – Immune Epitope Database and Analysis Resource (IEDB)

Os epítomos imunes são as estruturas moleculares reconhecidas por receptores imunes adaptativos: receptores de células T (TCRs), Receptores de células B (BCRs) e anticorpos. O propósito do IEDB é de catalogar todas as informações derivadas experimentalmente sobre epítomos imunes (VITA et al., 2010).

O IEDB corresponde a uma plataforma *on line* que inclui milhares de epítomos, incluindo peptídeos imunogênicos (GOMARA; HARO, 2007), e disponibiliza observações sobre como os epítomos foram identificados, o tipo de imunização, de antígeno, o tipo de ensaio e sobre algumas análises dos peptídeos. Juntamente aos epítomos identificados utilizando ensaios funcionais, o IEDB inclui dados sobre epítomos revisados e inferidos a partir das estruturas tridimensionais de complexos antígeno-anticorpo, disponíveis no PDB (BERMAN et al., 2000). Neste banco de dados temos elencados epítomos de células B e T caracterizados, bem como dados sobre o Complexo Principal de Histocompatibilidade (MHC). Também é possível localizar as estruturas moleculares reconhecidas por receptores imunes adaptativos e os contextos experimentais em que essas moléculas foram determinadas como sendo epítomos imunes. O banco de dados pode ser consultado para o estudo sobre a estrutura dos epítomos, organismo fonte, MHC restrição, tipo de ensaio ou organismo hospedeiro, entre outros critérios (figura 8) (VITA et al., 2010).

Um exemplo de aplicabilidade desta ferramenta nos estudos sobre Papilomavírus está representado pelo estudo conduzido por Kumar et al. (2015), onde a análise de epítomos imunogênicos foram conduzidos utilizando-se a amostra de referência K02718, que corresponde à sequência protótipo para o HPV16. O estudo procurou identificar epítomos para a oncoproteína 16E5. Foram identificados epítomos relacionados às moléculas do MHC-I e do MHC-II. Este estudo preliminar, pode contribuir para a geração de novos dados relativos às variantes dos oncogenes virais, em especial o E5.

2.9.6 MrBayes 3.2, jModelTest 2 e Tracer 1.5

Atualmente, vários métodos têm sido utilizados para a inferência de filogenias (tabela 4). O software MrBayes representa um pacote contendo um conjunto de softwares úteis para a realização de Inferência Filogenética. A inferência bayesiana utiliza uma função de Verossimilhança para produzir a árvore filogenética mais

provável para os dados. Uma de suas ferramentas, está representada pelo método Monte Carlo Cadeia de Markov (MCMC), que permite a análise de modelos complexos de árvores (RONQUIST et al., 2012). Em suas primeiras versões, o pacote MrBayes produzia modelos filogenéticos simples, tendo sido ampliado para a análise de uma gama de dados binários e o suporte para modelos filogenéticos mistos (RANNALA; YANG, 2008).

Figura 08 – Plataforma IEDB.

Summary Metrics

Peptidic Epitopes	450,617
Non-Peptidic Epitopes	2,669
T Cell Assays	339,032
B Cell Assays	415,704
MHC Ligand Assays	932,229
Epitope Source Organisms	3,668
Restricting MHC Alleles	771
References	19,525

Epitope Analysis Resource

T Cell Epitope Prediction

Scan an antigen sequence for amino acid patterns indicative of:

- MHC I Binding
- MHC II Binding
- MHC I Processing (Proteasome, TAP)
- MHC I Immunogenicity

B Cell Epitope Prediction

Predict linear B cell epitopes using:

- Antigen Sequence Properties

Predict discontinuous B cell epitopes using antigen structure via:

- Discotope
- Ellipro

Epitope Analysis Tools

Analyze epitope sets of:

- Population Coverage
- Conservation Across Antigens
- Clusters with Similar Sequences

As setas em negrito apontam para os links que direcionam para a predição dos epítomos de células T, a partir da inserção da sequência de aminoácidos da proteína em análise.

Tabela 04 – Comparação entre os tipos de métodos para inferência de filogenias.

Tipo	Método	Princípio	Programa
Métodos Quantitativos	UPGMA	Agrupa sequencialmente as OTUs com menor distância evolutiva entre si	Geneious MEGA MEGA
	Aproximação dos vizinhos	Busca a árvore com a menor soma total de ramos	Geneious HyPhy PAUP
	Máxima Parcimônia	Busca a filogenia com menor número de eventos evolutivos	MEGA Mesquite PAUP
	Máxima Verossimilhança	Busca a árvore com o valor de maior verossimilhança entre todas as filogenias construídas	PAML phyML MEGA
Métodos Qualitativos	Estatística Bayesiana	Amostra um número representativo de filogenias a partir do espaço amostral total de árvores e busca a mais provável	Mr. Bayes BEAST BAMBE

Extraído de Bioquímica: da biologia à flexibilidade molecular. 1ª edição. Sociedade Brasileira de Bioquímica e Biologia Molecular – SBBq. 2014.

A versão 3.2 representa uma das atualizações que contém, por exemplo, a construção de modelos evolutivos probabilísticos, surgimento de melhores algoritmos MCMC e diagnósticos de convergência evolutiva mais sofisticados, cálculo da razão dN/dS, melhora nos métodos para a escolha de modelos bayesianos, dentre outras funções (RONQUIST et al., 2012).

MrBayes usa o MCMC para aproximar as probabilidades posteriores das árvores. Na sua interface é possível pode alterar as premissas do modelo de substituição, remover e adicionar taxa e caracteres à análise, e é capaz de inferir estados ancestrais que acomodam a incerteza para a árvore filogenética e os parâmetros do modelo (RONQUIST et al., 2012).

Para obter o modelo evolutivo que mais se adequa ao conjunto de dados foi utilizado o programa jModelTest 2, que representa uma ferramenta para realizar a seleção estatística dos melhores modelos de substituição de nucleotídeos (DARRIBA et al., 2012). O referido programa também oferece, por exemplo, as estimativas de incertezas de seleção de modelo e a otimização das árvores (topologia e comprimento das ramificações).

jModelTest 2 utiliza o PhyML3 para obter estimativas de máxima verossimilhança do modelo. Utiliza modelos de substituição de nucleotídeos, testes

de razão de verossimilhança, critérios de informação Akaike (AIC) e Bayesiano (BIC) (GUINDON et al., 2010).

O software Tracer 1.5 é um programa para analisar os arquivos de rastreo gerados por execuções Bayesian MCMC. Trata-se de uma ferramenta gráfica para visualização e diagnóstico da saída do MCMC, podendo ler os arquivos de saída do MrBayes e BEAST (*Bayesian Evolutionary Analysis Sampling Trees*) (DRUMMOND; RAMBAUT, 2007).

Inferências filogenéticas são procedimentos de estimativas. Assim, existe uma série de algoritmos que podem ser aplicados e que são classificados em duas categorias: critérios qualitativos, baseados em otimização dos dados; e critérios quantitativos, baseados na distância entre os dados (FERNANDES-MATIOLI, 2001).

A aplicabilidade destas ferramentas tem sido avaliada em estudos de filogenia, visto o crescente número de novos genótipos virais, incluindo o estudo sobre as variantes de HPV.

Chen et al. (2011), analisou o DNA de amostras cervicais previamente caracterizadas para o tipo HPV, provenientes de grupos de pesquisa de vários países, como a Costa Rica, Taiwan, Tailândia, Ruanda, Burkina Faso e Zambia, para rastrear novas variantes dos HPVs 16 e dos genótipos relacionados HPV31, 33, 35, 52, 58 e 67. O genoma completo de 120 variantes relacionadas à família alfa-9 foi sequenciado para capturar a heterogeneidade viral máxima. A partir dos resultados do sequenciamento, a evolução viral foi caracterizada pela construção de árvores filogenéticas utilizando-se múltiplos algoritmos. O software MrBayes v3.1.2 foi utilizado com 10.000.000 ciclos para o algoritmo Markov Cadeia Monte Carlo (MCMC), para obter as árvores filogenéticas. Para a construção de árvores bayesianas, o software ModelTest v3.7 foi usado para identificar o melhor modelo evolutivo. Árvores de máxima parcimônia (MP) e *Neighbor Joining* (NJ) foram calculados por uma pesquisa heurística no PAUP*v4.0b10 e, para as análises máximas de parcimônia, as seqüências de nucleotídeos foram restritas à regiões informativas filogeneticamente. Os dados foram reamostrados 1.000 vezes de bootstrap.

Ainda neste estudo, a construção da árvore filogenética dos HPVs Alpha-9 foi construída utilizando-se o software MrBayes v.3.1.2, que inferiu o alinhamento global da sequência de cada genoma, linearizado a partir do ATG da ORF do gene E1. Os HPVs 34 e 73 foram considerados o outgroup (CHEN et al., 2011) (figura 2). E, para

o HPV31, as Árvores Bayesianas foram inferidas a partir do alinhamento global de sequências completas de nucleotídeos do genoma, e calculados o valor de credibilidade bayesiana usando o software MrBayes v3.1.2, a porcentagem de bootstrap pela máxima parcimônia (MP) e a porcentagem de bootstrap por *Neighbor Joining* (NJ), ambos se utilizando o software usando PAUP*v4.0b10 (figura 9).

Para o HPV16, a topologia da árvore também foi construída com base nas comparações individuais de cada genoma e uma árvore de máxima verossimilhança (ML) foi inferida a partir de um alinhamento global de seqüências completas de nucleotídeos do genoma do HPV16 usando o software RAxML HPC v7.2.8 e a *distância-p* foi usado para calcular as diferenças de seqüência de nucleotídeos pareadas para cada isolado, em comparação com todos os outros isolados, baseados nas seqüências completas de nucleotídeos do genoma, utilizando-se o software MEGA5 (figura 10).

2.9.7 Análise da pressão de seleção

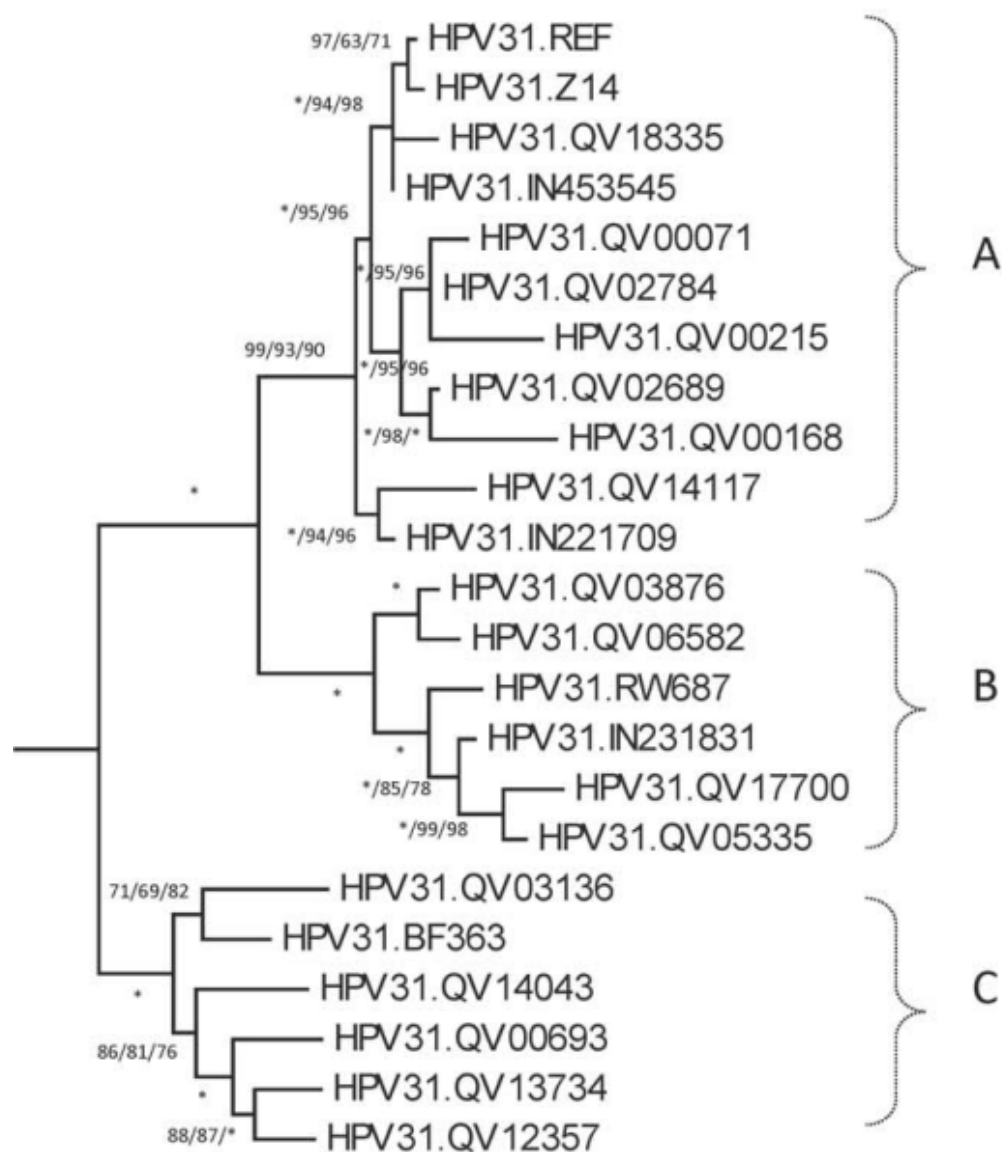
Diversos métodos filogenéticos para análise comparativa de sequências de DNA e proteínas estão se tornando cada vez mais importantes, devido ao rápido acúmulo de sequências gênicas. Por isso, torna-se comum a reconstrução da filogenia envolvendo uma grande quantidade de informações gênicas e esses métodos filogenéticos estão sendo amplamente utilizados para estimar as taxas evolutivas de genes e genomas, além de avaliar aspectos evolutivos e da seleção natural para interpretar dados genômicos (YANG, 2005).

Para a análise da pressão de seleção, foi utilizado o pacote de programas PAML (*Phylogenetic Analysis by Maximum Likelihood*), versão 4.9c para análises filogenéticas de seqüências de DNA ou proteínas usando máxima verossimilhança. Um dos seus softwares incorporados neste pacote, o programa CODEML, foi utilizado para obter as estimativas de máxima Verossimilhança, através da análise de substituição de códons (YANG, 2007).

A importância da ferramenta PAML está reside em seus sofisticados modelos de substituição, úteis para compreender o processo de evolução da sequência. Neste pacote podem ser realizadas várias análises, dentre elas: (1) Comparação e testes de árvores filogenéticas; (2) Estimativas de parâmetros em modelos de substituição; (3) Testes de razão de verossimilhança (LRTs) de hipóteses através de comparação de modelos estatísticos aninhados; (4) Estimativa de taxas de

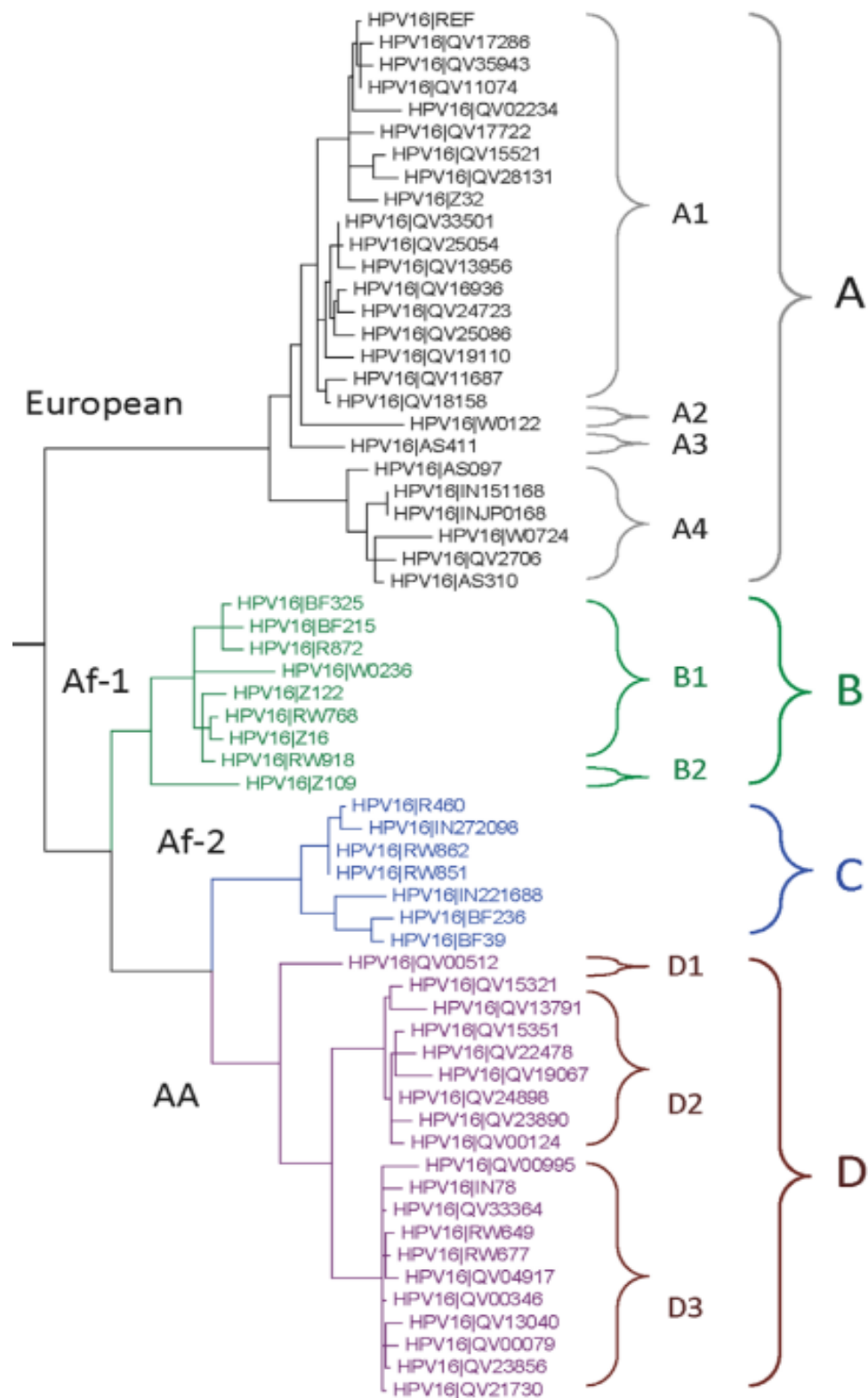
substituição sinônimas e não sinônimas; (5) Estimativa de matrizes empíricas de substituição de aminoácidos; (6) Estimativa dos tempos de divergência de espécies sob condições globais; (7) Reconstrução de seqüências ancestrais usando nucleotídeos, modelos de aminoácidos e códons; e (8) Geração de nucleotídeo, códon e seqüência de aminoácidos alinhados por simulação de Monte Carlo (YANG, 2007).

Figura 9 – Árvore topológica das variantes do HPV31 e comparações de pares de genomas completos individuais. Nesta construção, são distinguíveis as linhagens A, B e C e suas respectivas variantes.



Fonte: Chen et al. (2011).

Figura 10 – Topologia da árvore HPV16 e comparações pareadas de genomas completos individuais. Nesta construção, são distinguíveis as linhagens A, B, C e D, além de suas respectivas variantes.



Fonte: Burk et al. (2013).

2.9.8 Estrutura Terciária da Proteína

O método *Ab initio* ainda é um dos principais problemas não resolvidos em biologia computacional. Para minimizar este problema, o algoritmo QUARK foi desenvolvido. QUARK é um algoritmo de computador para predição de estrutura de proteína *ab initio* e dobramento de peptídeo de proteína, que visa construir o modelo de proteína 3D correto somente a partir da sequência de aminoácidos. Os modelos são construídos a partir de pequenos fragmentos (1-20 resíduos de comprimento) por simulação de Monte Carlo de troca de réplica sob o guia de um campo de força baseado em conhecimento de nível atômico (XU; ZHANG, 2012).

Como nenhuma informação de modelo global é usada na simulação de QUARK, o servidor é adequado para proteínas que não possuem modelos homólogos na biblioteca do PDB.

Outro método utilizado, emprega o uso do programa Robetta, que corresponde a um servidor de previsão de estrutura de proteínas, totalmente automatizado, que usa o método de inserção de fragmentos do Rosetta, combinando métodos de predição de estruturas baseados em modelo e *de novo* na tentativa de produzir modelos de alta qualidade que cubram cada resíduo de uma sequência submetida. Então, quando nenhum homólogo estrutural está disponível, os domínios são modelados com o protocolo Rosetta *de novo*, que permite que o comprimento total do domínio explore o espaço conformacional via inserção de fragmento, produzindo um grande conjunto a partir do qual os modelos finais são selecionados (KIM et al., 2004).

Para a verificação detalhada da estereoquímica da estrutura protéica, foi utilizado o pacote de programas PROCHECK, responsável por fornecer uma avaliação da qualidade geral da estrutura protéica, em comparação com estruturas bem refinadas da mesma resolução, e também destacam regiões que podem precisar de investigação adicional. Portanto, o pacote PROCHECK é útil para avaliar a qualidade das estruturas protéicas, das estruturas existentes e daquelas modeladas em estruturas conhecidas (LASKOWSKI et al., 1993).

E, para identificar o possível efeito das mutações na estrutura e função da proteína, o programa *Site Directed Mutator* (SDM) torna-se útil para prever o efeito que os SNPs (*Single Nucleotide Polimorphisms*) terão na estabilidade das proteínas. O software usa a frequência de substituições de aminoácidos dentro de uma família de proteínas para calcular o escore de estabilidade. Esse método apresenta alta

sensibilidade e, por isso, executa a tarefa de classificação de uma mutação como estabilizante ou desestabilizante, sendo considerada uma ferramenta útil em experimentos de mutagênese sítio-dirigida ou para prever se uma mutação terá impacto na estrutura protéica (WORTH et al., 2011).

3 RESULTADOS

3.1 ARTIGO I - DIVERSIDADE GENÉTICA DO ONCOGENE E5 DO PAPILOMAVÍRUS HUMANO 31 IDENTIFICADOS EM AMOSTRAS CLÍNICAS CIRCULANTES NO NORDESTE DO BRASIL.

Diversidade genética do oncogene E5 do Papilomavírus Humano 31 identificados em amostras clínicas circulantes no Nordeste do Brasil.

(Submetido à Revista: **Gynecologic Oncology**)

^aAntônio Humberto Pereira da Silva Júnior

^bAna Pavla Almeida Diniz Gurgel

^aBárbara Simas Chagas

^aRuany Cristyne de Oliveira Silva

^aRita de Cassia Pereira de Lima

^cMarcus Vinicius de Aragão Batista

^aAntonio Carlos de Freitas

^aFederal University of Pernambuco; Department of Genetics; Laboratory of Molecular Studies and Experimental Therapy (LEMTE); Pernambuco, Brazil.

^bDepartment of Engineering and Environment, Federal University of Paraiba, Paraiba – Brazil.

^cDepartment of Biology, Federal University of Sergipe, Sergipe, Brazil.

Running Title: Diversidade Genética no gene E5 do HPV31.

Corresponding author: Dr. Antonio Carlos de Freitas, Av. Prof. Moraes Rego, 1235, Cidade Universitária, 50670-901, Recife-PE, Brazil. Fax: +55 81 21268512. E-mail: acf_ufpe@yahoo.com.br.

Abstract

Intra-typical variations of the HPV31 E5 oncogene have been little studied in the Brazilian population. This oncogene is involved in the regulation processes of growth factors, cell proliferation and immune escape. In this study we have investigated the circulation of E5 oncogene variants from HPV31 in a population of Northeast Brazil and evaluated in silico the relevance of these polymorphisms. The 31E5 sequences were obtained from cervical samples and the polymorphisms identification was

performed by MEGA6 software while the *in silico* analysis was performed with the aid of the software PSIPRED, PALM and IEDB. Thirty-five samples were identified as being HPV31. In these samples, ten polymorphisms (five synonymous and five non-synonymous) were identified. P56A polymorphism was evaluated as being highly stabilizing, which can lead to the anomalous function of the protein. The phylogenetic analysis showed five isolates belonging to the A lineage, one to the B lineage and one to the C lineage. 31E5 oncogene is possibly under purified or negative selection. The 31E5 variants showed immunogenic epitopes mapped to molecules belonging to the HLA classes: -DRB1, -DRB5, -DPA, -DPB. The results will help on the comprehension of the geographic distribution of variants and the lineage fixation in different populations.

Keywords: HPV; Cervical Cancer; Polymorphisms on E5.

1 INTRODUÇÃO

O câncer cervical é considerado uma das principais causas de câncer em mulheres, onde estima-se cerca de 527.624 novos casos de pacientes diagnosticadas e cerca de 266.000 mortes anualmente notificadas (WHO, 2015). No Brasil, o câncer cervical é considerado a terceira causa mais comum de câncer, sendo esperado para o ano de 2016 aproximadamente 16.340 novos casos (INCA, 2016).

Diversos estudos epidemiológicos estabeleceram que as infecções persistentes provocadas pelos papilomavírus humano de alto risco (HR-HPV) correspondem ao principal fator responsável pelo desenvolvimento do câncer cervical (WHO, 2016). Os HPVs são pequenos vírus não-envelopados, com DNA fita dupla e um genoma com tamanho aproximado de 8 kb associado às histonas (Zur Hausen, 1996). Nele, distinguem-se duas regiões não traduzidas, denominadas Non Coding Region (NCR) e Long Control Region (LCR), além dos genes de expressão precoce (E1, E2, E4, E5, E6 e E7) e a região que contém os genes de expressão tardia (L1 e L2) (Bernard et al., 2010; Van Doorslaer, 2013). Os genes E6 e E7 e, mais recentemente o E5, são considerados oncogenes virais (de Freitas et al., 2017).

Sabe-se que a diversidade genética dos HPVs pode estar relacionada a um maior potencial de progressão da infecção e das lesões, a exemplo do que se observa com o papel das variantes A e B do HPV31, culminando com um ciclo

carcinogênico mais rápido (Xi et al., 2012; Burroni et al., 2013; Cornet et al., 2013; Chagas et al., 2013; Gurgel et al., 2013). No entanto, os dados da literatura trazem poucas informações a respeito da relevância clínica das variantes de HPV31 (Chagas et al., 2011, 2013).

Como o potencial de progressão da infecção e das lesões podem estar relacionados à diversidade genética do HPV, e como o HPV31 é classificado de alto risco oncogênico, análises sobre o papel funcional de variantes do HPV31 são relevantes. Além disso, existem poucos estudos sobre os dados da diversidade genética do oncogene E5 do HPV31.

O oncogene E5 codifica para uma proteína de baixo peso molecular, hidrofóbica e com número variável de aminoácidos (VENUTI et al., 2011). E5 apresenta como principais alvos as membranas celulares do aparato de Golgi, do retículo endoplasmático rugoso e das vesículas endossomais, sugerindo a sua participação no tráfego intracelular (Shapiro et al., 2000; Chang et al., 2001). Esse oncogene tem um amplo espectro de atuação, garantido pela sua capacidade de integração no genoma da célula hospedeira durante o processo de progressão maligna, embora sua ação seja raramente detectada nos estágios mais avançados dos tumores cervicais (VENUTI et al., 2011; DiMaio e PETTI, 2013; de Freitas et al., 2017). E5 participa do (1) estímulo à proliferação celular e expansão dos queratinócitos infectados, (2) redução da ação dos mecanismos apoptóticos, (3) regulação e degradação de receptores celulares, (4) inibição de proteínas que regulam o ciclo celular, (5) aumento do tamanho nuclear e do conteúdo de DNA, (6) poliploidia, (7) binucleação e instabilidade cromossômica, (8) escape imunológico, (9) redução da expressão de moléculas HLA acessíveis para epítomos de HPV, (10) supressão das proteína da via de estresse do Retículo Endoplasmático (RE), a ciclo-oxigenase-2 (COX-2), X-box binding protein (XBP-1) e a inositol-requiring enzyme-1a (IRE1 α), além da interferência em outras importantes funções biológicas (Maufort et al., 2010; Campo et al., 2010; MONFRÉ, 2011).

Desta forma, objetivo deste trabalho foi investigar variantes do oncogene E5 de HPV31 a partir de amostras cervicais, e avaliar *in silico* a possível repercussão biológica no processo infeccioso/carcinogênico.

2 MATERIAIS E MÉTODOS

2.1 CARACTERÍSTICAS DAS AMOSTRAS ESTUDADAS

Foram obtidas 107 amostras clínicas de escovados cervicais (*Cytobrush*), coletadas de pacientes atendidas pelo serviço de ginecologia do Hospital das Clínicas da Universidade Federal de Pernambuco (HC-UFPE). A idade das pacientes variou de 18 a 82 anos. Todas assinaram o termo de consentimento livre e esclarecido (TCLE) para a participação neste estudo, e os critérios de exclusão foram pacientes portadoras da imunodeficiência humana e recusa da assinatura do TCLE.

As amostras coletadas foram armazenadas em solução salina tamponada com fosfato (PBS) e armazenada a -20°C. A extração de DNA se deu utilizando-se o Kit GenominPrepBlood DNA isolation (AmershamBioscience), seguindo as instruções do fabricante e, em seguida, quantificadas utilizando-se um espectrofotômetro (NanoDrop).

Este trabalho foi aprovado pelo Comitê de Ética da Universidade Federal de Pernambuco (CAAE: 3606212.7.0000.5208).

2.2 DETECÇÃO E GENOTIPAGEM DO HPV

O DNA obtido foi submetido a um controle de qualidade e integridade, através da amplificação por PCR do gene WAF-1, que codifica a proteína p21 do genoma humano, como controle endógeno (Amaral et al., 2014). A detecção do DNA do HPV foi realizada por PCR, utilizando-se os primers degenerados MY09/MY11 (Manos et al., 1989), e tipificação para o HPV31 com o uso de PCR tipo-específica com os primers para o oncogene E6 (Chagas et al., 2011). Em seguida, a amplificação completa do oncogene E5 do HPV31, se deu com a utilização dos primers designados E5_HP31seqF (5'-GACATGTACAGATGGAAAAC-3') e E5_HP31R (5'-CGTTTAGTGCGTTTTGTAGAG-3').

2.3 SEQUENCIAMENTO DO ONCOGENE E5 DO HPV31

35 amostras HPV31 positivas foram enviadas para o sequenciamento, apenas o cromatograma de 07 amostras atendeu ao critério de qualidade PHRED $\geq$ 30, onde o produto de PCR de cada amostra foi sequenciado 2 vezes com os seguintes pares de primer: E5_HP31seqF (5'-GACATGTACAGATGGAAAAC-3') e E5_HP31R (5'-CGTTTAGTGCGTTTTGTAGAG-3') utilizando-se o ABI PRISM BigDye® Cycle Sequencing v3.1 Ready Reaction (Applied Biosystems).

2.4 ANÁLISE DAS SEQUÊNCIAS

As sequências de nucleotídeos foram analisadas utilizando-se o programa Gap4 (versão 4.0) e o Pregap4 (versão 1.5), do pacote Staden (STADEN, 1996). As sequências foram alinhadas com a sequência de referência da linhagem A1 (GenBank accession number J04353), utilizando a ferramenta Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1990). Também foi utilizado a sequência das amostras depositadas no *Papillomavirus Episteme* (PaVE), para a verificação de variações na sequência gênica do oncogene 31E5. Para isso, utilizou-se o programa MEGA6 e os alinhamentos foram realizados com o auxílio do algoritmo CLUSTALW (Tamura et al., 2011).

2.5 PREDIÇÃO DE EPÍTOPOS DE CÉLULAS T

Para a análise da sequência de aminoácidos e identificação de possíveis epítopos imunogênicos, foi utilizado o servidor online Immune Epitope Database and Analysis Resource (IEDB) (<http://www.iedb.org/>).

Os epítopos foram previstos utilizando os parâmetros padrões estabelecidos pelo servidor. Para o MHC-I, utilizamos o set de alelos mais frequentes na população mundial, comprimento da sequência do epítipo (que variou entre 9-10mer), e a classificação dos melhores epítopos de acordo com o percentile rank <1 e immunogenicity score >0.5.

Para a identificação de potenciais epítopos para as moléculas do MHC-II, novos critérios foram incluídos, como a análise do percentile value e o IC50 value (<50nM).

2.6 ANÁLISE FILOGENÉTICA

A análise filogenética foi realizada por meio do método da Inferência Bayesiana utilizando o programa MrBayes versão 3.2 (Ronquist et al., 2011). O modelo evolutivo que mais se adequa ao conjunto de dados foi avaliado utilizando o programa jModelTest 2 (Darriba et al., 2012). Aplicando o Critério de Informação Bayesiano (BIC), foi estabelecido o modelo TIM1+I, que foi utilizado. Quatro cadeias da Monte Carlo Cadeia de Markov (MCMC) foram corridas simultaneamente e amostradas a cada 500 gerações por um total de 10 milhões de gerações. Os escores de log-likelihood foram plotados contra as gerações usando o Tracer versão

1.5 (Rambaut and Drummond, 2007). Desta forma, as primeiras 5000 árvores amostradas de cada corrida, o que correspondeu a 25%, foram descartadas como burn-in. As árvores restantes foram usadas para construir uma árvore consenso usando a 50% majority-rule.

Três agrupamentos foram identificados como linhagens A, B e C, e os subgrupos foram identificados como sublinhagens A1, A2, B1, B2, C1, C2 e C3. Os genomas de referência utilizados para esta classificação foram: J04353 (A1), HQ537675 (A2), HQ537676 (B1), HQ537680 (B2), HQ537682 (C1), HQ537684 (C2) e HQ537685 (C3).

2.7 ANÁLISE DE SELEÇÃO POSITIVA

Para avaliar o tipo de pressão seletiva que está associada ao oncogene 31E5, foram utilizadas estimativas de Máxima Verossimilhança, utilizando modelos de substituição de códons através do programa CODEML, incorporado no pacote de programas PAML versão 4.9c (Yang, 2007). A detecção da seleção positiva foi realizada através do cálculo dos parâmetros de seis modelos de substituição de códons, M0, M1, M2, M3, M7 e M8. O Teste de Razão de Verossimilhança (LRT) foi utilizado para avaliar qual modelo melhor se adequa aos dados.

2.8 PREDIÇÃO DA ESTRUTURA SECUNDÁRIA E TERCIÁRIA DA ONCOPROTEÍNA E5 DE HPV31

Após a identificação das regiões polimórficas no oncogene E5 das amostras variantes do HPV31, foi utilizada a ferramenta PSIPRED para a predição de estrutura secundária da proteína, por se tratar de um método acurado e reunir diferentes métodos de análise em sua plataforma (PSIPRED, GenTHREADER e MEMSAT 2) (McGuffin et al., 2000).

Para a predição da estrutura terciária da oncoproteína 31E5 de referência foi utilizada a sequência proteica com o número de acesso AAA46954. Uma busca no banco Protein Data Bank (PDB) através da ferramenta blastp foi realizada para identificar uma possível estrutura molde, com o objetivo de realizar modelagem por homologia. Nenhuma similaridade significativa foi obtida com as estruturas do PDB, impossibilitando a utilização da modelagem comparativa. Desta forma, foi utilizada a modelagem da estrutura terciária de E5 de HPV31 através do método *ab initio*, utilizando o programa QUARK (Xu and Zhang, 2012) e o programa Robetta (Kim et

al., 2004). Além disso, o modelo estrutural foi predito para cada amostra que apresentou variação na sequência de aminoácidos. Foram obtidos 10 modelos diferentes para cada amostra para serem avaliados, que foram refinados por minimização de energia utilizando o programa ModRefiner (Dong and Zhang, 2011). Os modelos refinados foram avaliados pelos programas PROCHECK (Laskowski et al., 1993) através da avaliação de qualidade estereoquímica dos modelos a partir do gráfico de Ramachandran, Anolea (Melo and Feytmans, 1998), Gromos (van Gunsteren et al., 1996), QMEAN6 (Benkert et al., 2008), DFire (Zhou and Zhou, 2002), DSSP (Kabsch and Sander, 1983) e Promotif (Hutchinson and Thornton, 1996). De acordo com estes parâmetros, o melhor modelo foi escolhido para representar a estrutura das proteínas.

Para identificar o possível efeito das mutações na estrutura e função da oncoproteína 31E5 foi utilizado o método Site Directed Mutator (SDM), que prediz o efeito das mutações na estabilidade das proteínas (Worth et al., 2011).

2.9 CODON USAGE

Após a obtenção e identificação das sequências variantes do oncogene 31E5, foi utilizando o software *on line* Graphical Codon Usage Analyser, disponível em <http://gcua.schoedl.de/>, cujo objetivo foi identificar o uso de códons raros ou frequentes, utilizados pelas variantes, comparando com os códons preferenciais utilizados pelos mamíferos.

3 RESULTADOS

3.1 DETECÇÃO E TIPIFICAÇÃO DO HPV

A detecção e a tipificação foram feitas com 93 amostras (excluindo-se as amostras com resultados histopatológicos inconclusivos). O DNA do HPV esteve presente em 87,1% (81/93). O HPV31 foi detectado em 43,20% (35/81) de amostras HPV positivas.

Dentre as amostras HPV31 positivas avaliadas, foram elegíveis somente as sequencias com nota PHRED ≥ 30 , excluindo-se também aquelas que apresentaram inconsistências nos resultados do sequenciamento. As amostras selecionadas foram identificadas como HPV31E5_33PE, HPV31E5_56PE, HPV31E5_60PE, HPV31E5_62PE, HPV31E5_63PE, HPV31E5_64PE e HPV31E5_82PE.

A detecção do HPV31 e suas variantes de E5 se deu em amostras com diferentes resultados histopatológicos (tabela 1).

Tabela 01 – Resultado histopatológico das amostras de escovados cervicais, relacionadas à presença de polimorfismos em E5 de HPV31.

Amostra cervical	Resultado Histopatológico
HPV31E5_33PE*	NIC III
HPV31E5_56PE*	NIC I
HPV31E5_60PE	NIC I
HPV31E5_62PE	Sem lesão
HPV31E5_63PE*	NIC I
HPV31E5_64PE*	Sem lesão
HPV31E5_82PE	NIC I

* indica amostras contendo polimorfismos em E5.

3.2 VARIABILIDADE GENÉTICA DO ONCOGENE E5 DO HPV31

Foram identificamos um total de 10 substituições de nucleotídeos, das quais cinco substituições sinônimas, e cinco substituições não-sinônimas (tabela 2). Todos os polimorfismos identificados já estavam depositados na base de dados do NCBI e do PaVE.

As amostras mais polimórficas foram HPV31E5_33PE e HPV31E5_56PE, com um total de 06 e 05 substituições, respectivamente. A região com o maior número de amostras contendo polimorfismo ocorreu na posição 4053 (tabela 2).

Tabela 02 – Variações nucleotídicas do oncogene E5 do HPV31. Nucleotídeos conservados em relação a sequência de referência J04353 estão marcados com asterístico (*). L: Leucina; N: Asparagina; D: Ácido Aspártico; V: Valina; S: Serina; I: Isoleucina; P: Prolina; A: Alanina; F: Fenilalanina; Q: Glutamina.

Isolates	Positions of E5 nucleotide substitutions									
	3827	3828	3956	3957	3980	3981	4005	4052	4053	4064
Reference	A	A	G	A	T	C	G	T	T	A
HPV31E5_33PE	*	G	*	*	*	G	A	C	A	G
HPV31E5_56PE	G	*	*	G	A	*	*	C	A	*
HPV31E5_60PE	*	*	*	*	*	*	*	*	*	*
HPV31E5_62PE	*	*	*	*	*	*	*	*	*	*
HPV31E5_63PE	*	*	C	*	*	*	*	*	A	*
HPV31E5_64PE	*	*	C	*	*	*	*	*	A	*
HPV31E5_82PE	*	*	*	*	*	*	*	*	*	*
Position of altered códons	4	5	47	48	55	56	64	79	80	83
Reference	L	N	V	I	S	P	V	S	F	Q
	(CTA)	(AAT)	(GTG)	(ATT)	(TCT)	(CCA)	(GTT)	(TCT)	(TTT)	(CAA)
Altered amino acid	L	D	V	V	S	A	I	S	I	Q
	(CTG)	(GAT)	(GTC)	(GTT)	(TCA)	(GCA)	(ATT)	(TCC)	(ATT)	(CAG)

De acordo com o modelo de α -hélice estabelecido descrito por Saavedra et al. (2012), observamos que as substituições descritas neste estudo se localizam nas posições N5D, I48V e V64I, distribuídas ao longo da primeira, segunda e terceira α -hélice, respectivamente. Já a mutação P56A está localizada na região entre a segunda e terceira α -hélice, enquanto F80I, na região terminal da cadeia proteica.

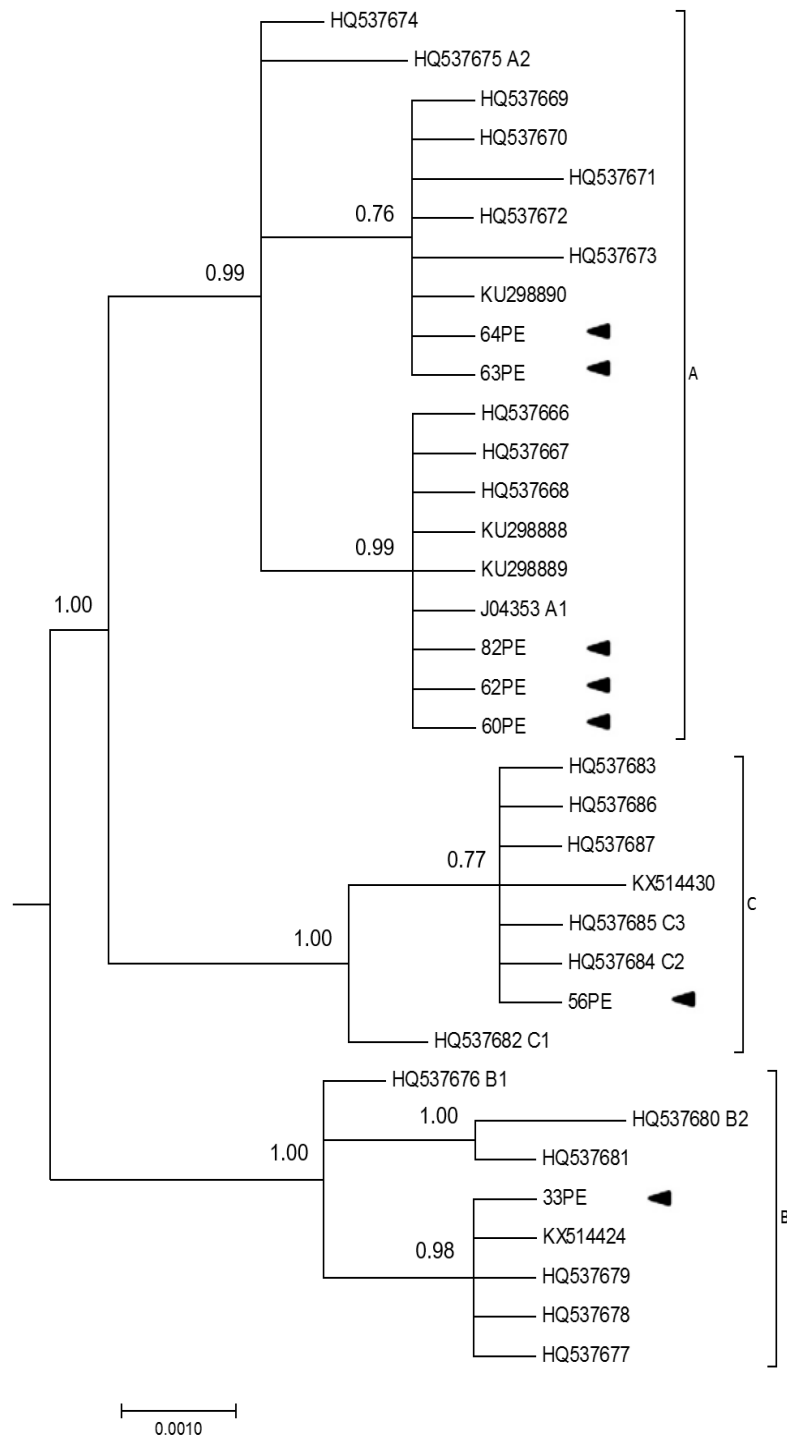
Os nossos resultados apontaram que os polimorfismos identificados em E5, apresentaram os melhores scores para a formação de epítomos imunogênicos com moléculas pertencentes ao MHC-I, nas amostras HPV31E5_33PE (59-68 CFCIYIVFIY - alelo HLA-A*26:01) e nas amostras HPV31E5_63PE (43-52 LLLIVILWVI - alelo HLA*02:01) e HPV31E5_64PE (43-52 LLLIVILWVI - alelo HLA*02:01).

Em relação às moléculas pertencentes ao MHC-II, os melhores scores foram obtidos pelas variantes HPV31E5_33PE (46-60 IVILWVIATSALRCF - alelo HLA-DRB5*01:01), HPV31E5_56PE (70-84 HLA-DRB1*07:01 - alelo PLFVIHTHASILSQQ), HPV31E5_63PE (67-81 IYIPLFVIHTHASIL HLA-DRB1*07:01; 68-82 PLFVIHTHASILS - alelo HLA-DRB1*07:01) e HPV31E5_64PE (67-81 IYIPLFVIHTHASIL - alelo HLA-DRB1*07:01; 68-82 PLFVIHTHASILS - alelo HLA-DRB1*07:01).

3.2 ANÁLISE FILOGENÉTICA

Para avaliar a diversidade mutacional do oncogene 31E5 foi realizada uma análise filogenética. A árvore filogenética obtida com o oncogene E5 evidencia um perfil de agrupamento característico das três linhagens de HPV31. A árvore apresenta agrupamentos bem suportados, possibilitando a classificação dos isolados em linhagens. Apenas os isolados 60PE, 62PE e 82PE foram agrupados de forma suportada na sublinhagem A1. A análise filogenética evidencia que cinco isolados são classificados como linhagem A, um isolado como B, e um isolado da linhagem C (Figura 01).

Figura 01 – Árvore filogenética baseado na Inferência Bayesiana obtida a partir da análise da sequência do oncogene 31E5. O suporte dos ramos foi determinado por probabilidades posteriores. As amostras HPV31E5_60PE, HPV31E5_62PE, HPV31E5_63PE, HPV31E5_64PE e HPV31E5_82PE pertencem à linhagem A; HPV31E5_33PE pertence à linhagem B; HPV31E5_56PE pertence à linhagem C.



3.3 ANÁLISE DE PRESSÃO SELETIVA

Na avaliação da pressão seletiva incidente sobre a oncoproteína 31E5, foi possível observar que esta proteína apresentou valores de ω entre 0 e 1, o que significa que o oncogene 31E5, nesta população, se encontra sob seleção. O teste de LRT para estimativas de pressão seletiva avalia se a adição de parâmetros livres em modelos aninhados aumenta significativamente a probabilidade de explicar os dados. Desta forma, para a oncoproteína 31E5, o modelo de substituição de códons mais adequado aos dados foi o modelo M0, apresentando $\omega = 0.35841$ (Tabelas 03 e 04).

Tabela 03 – Estimativas de pressão seletiva no oncogene E5 de HPV31 com base na probabilidade, calculada pelo programa PAML. O modelo que melhor se ajusta aos dados está em negrito.

ORF	Model	LnL	$\omega = dN/dS^*$	Parâmetros
E5	M0	-434.101372	0.35841	$\omega = 0.35841$
	M1	-434.101379	0.35840	$p_0 = 0.99999$ ($p_1 = 0.00001$) $\omega_0 = 0.35840$; $\omega_1 = 1.00000$
	M2	-434.101372	0.35841	$p_0 = 1.00000$; $p_1 = 0.00000$ ($p_2 = 0.00000$) $\omega_0 = 0.35841$; $\omega_1 = 1.00000$; $\omega_2 = 1.00000$
	M3	-434.101372	0.35841	$p_0 = 0.31058$; $p_1 = 0.24307$ ($p_2 = 0.44635$) $\omega_0 = 0.35841$; $\omega_1 = 0.35841$; $\omega_2 = 0.35841$
	M7	-434.103993	0.36510	$p = 56.94191$; $q = 99.00000$
	M8	-434.103559	0.35900	$p_0 = 0.99999$ ($p_1 = 0.00001$) $p = 55.45241$; $q = 99.00000$; $\omega_s = 1.00000$

*significa que os valores foram determinados a partir da média aritmética.

A oncoproteína 31E5 apresentou, de forma global, seleção negativa ou purificadora. Entretanto, foi verificada a possibilidade de sítios específicos da proteína estarem ou não sob seleção positiva ou diversificadora. A análise sítio-específica revelou que nenhum resíduo de aminoácido da oncoproteína 31E5 se encontrava sob seleção diversificadora (Tabela 05).

Tabela 04 – Resultados do teste LRT para estimativas de pressão seletiva, que verifica qual modelo se ajusta melhor aos dados.

ORF	Model 1	Model 2	LRT	Positively selected sites
E5	M0	M3	0.000000	-
	M1	M2	0.000014	None
	M1	M3	0.000014	None
	M7	M8	0.000868	None

3.4 ANÁLISE ESTRUTURAL DAS MUTAÇÕES DE 31E5: PERFIL DA ESTRUTURA SECUNDÁRIA E TERCIÁRIA

As substituições não-sinônimas, de acordo com o modelo descrito por Saavedra et al. (2012), foram identificadas no interior e entre as regiões de alfa-hélice da oncoproteína 31E5, com mudanças moderadas nas regiões de hélice, strand, coil and confidence of prediction, da sua estrutura secundária (figura 02). Estas mudanças estruturais moderadas também foram observadas em sua estrutura terciária (figura 03).

A maioria dos aminoácidos substituídos nestas regiões não-sinônimas mantiveram o caráter hidrofóbico, com exceção da substituição N (asparagina, aminoácido polar sem carga) → D (ácido aspártico, aminoácido polar com carga negativa), de natureza polar-hidrofílico. Com essas substituições, o percentual dos aminoácidos da oncoproteína 31E5, comparado à sequência de referência J04353, sofre uma discreta modificação no percentual dos seus resíduos de aminoácidos. As demais substituições identificadas são compostas por aminoácidos apolares-hidrofóbicos.

Em relação aos aminoácidos mais frequentes, temos que a sequência J04353 é rica em Leucina (17 resíduos de aminoácidos - Leu, L), Valina (13 resíduos de aminoácidos - Val, V) e Isoleucina (11 resíduos de aminoácidos - Ile, I). Mesmo sendo a amostra mais polimórfica, HPV31E5_33PE apresentou o mesmo percentual destes aminoácidos em sua composição.

Quanto à conformação tridimensional da oncoproteína 31E5, vários critérios de avaliação foram utilizados para confirmar que os modelos obtidos correspondem a uma estrutura que remete a uma conformação natural de estabilidade. De acordo com todos os parâmetros analisados, o melhor modelo da oncoproteína 31E5 apresentou 97,5% dos resíduos nas regiões mais favoráveis no gráfico de Ramachandran, mostrando que foi possível obter um modelo bastante satisfatório. Todos os outros métodos descritos anteriormente, e que foram utilizados para a avaliação de qualidade, confirmam a confiabilidade do modelo (dados não mostrados).

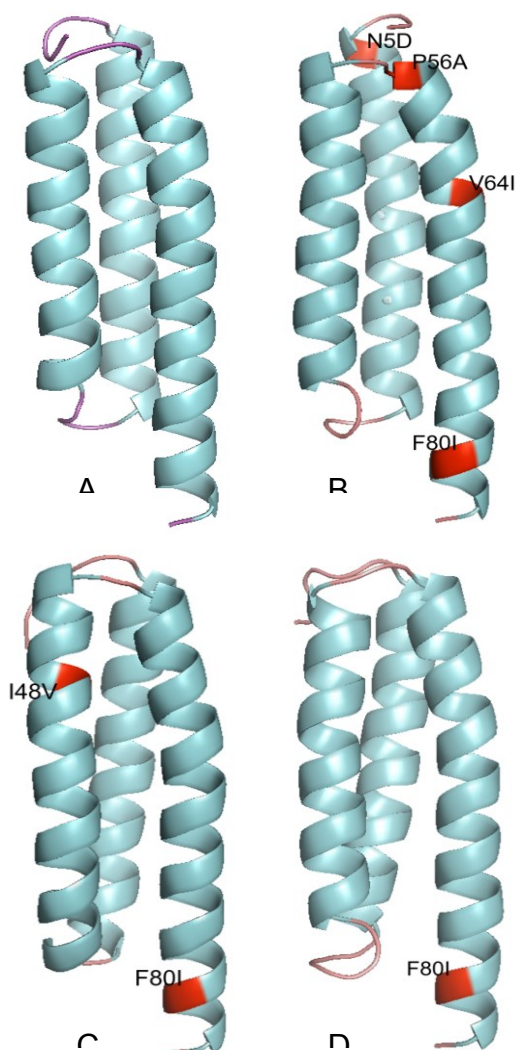
A oncoproteína 31E5 possui 84 resíduos de aminoácidos, formada por três alfa-hélices ligadas por pequenas alças, denominadas hélices A, B e C, estabelecendo um grande domínio transmembrana. Esta oncoproteína é formada basicamente por resíduos de aminoácidos hidrofóbicos compondo as regiões de hélice. As cinco mutações não-sinônimas encontradas nas variantes descritas neste estudo estão localizadas em regiões de alfa-hélice. N5D está localizada no início da hélice A, I48V no final da hélice B, P56A no início da hélice C, V64I na parte central da hélice C, e F80I no final da hélice C (Figura 02).

Na avaliação do impacto destas mutações sobre a estabilidade da oncoproteína, observamos que as mutações N5D e V64I apresentaram valores preditos de $\Delta\Delta G$ entre 0 e 1, indicando ser mutações ligeiramente estabilizantes, não estando assim associadas com nenhuma alteração funcional da proteína. A mutação I48V apresentou valor predito de $\Delta\Delta G = -0.59$, indicando ser uma mutação ligeiramente desestabilizante, mas não estando assim associada com nenhuma alteração funcional da proteína. A alteração P56A apresentou $\Delta\Delta G = 3.25$, indicando ser altamente estabilizante, podendo levar a um mau funcionamento da proteína. Já a alteração F80I apresentou $\Delta\Delta G = -0.08$, indicando ser uma alteração neutra, não estando assim associada com alterações funcionais da proteína (Tabela 05).

3.5 Variabilidade e Utilização de Códon Preferenciais

Uma possível ação regulatória dos polimorfismos do oncogene 31E5, identificados neste estudo (Tabela 2), poderia estar relacionada com algum mecanismo de adaptabilidade destas linhagens à célula hospedeira. No entanto, não foi observado diferenças significativas entre o códon utilizado nas regiões polimórficas das variantes estudadas com os utilizados preferencialmente em mamíferos.

Figura 03 – Modelo tridimensional da oncoproteína 31E5 estabelecida a partir do método *ab initio*.



A estrutura evidencia as três alfa-hélices da proteína transmembrana da sequência de J04353 com a identificação dos locais de ocorrência das mutações. A: simboliza a predição do modelo para a oncoproteína 31E5, da sequência de referência J04353; B: polimorfismos identificados no isolado 33PE; C: polimorfismos identificados no isolado 56PE; D: polimorfismos identificados nos isolados 63 e 64.

Tabela 05 – Impacto das mutações na estabilidade da oncoproteína E5 das amostras variantes do HPV31 avaliadas neste estudo.

Mutation	Wild Residue	Residue Position	Mutant Residue	Predicted $\Delta\Delta G$	Outcome
1	N	5	D	0.70	ligeiramente estabilizante e não associado ao mau funcionamento da proteína.
2	I	48	V	-0.59	ligeiramente estabilizante e não associado ao mau funcionamento da proteína.
3	P	56	A	3.25	altamente estabilizante e pode causar mau funcionamento da proteína.
4	V	64	I	0.62	ligeiramente estabilizante e não associado ao mau funcionamento da proteína.
5	F	80	I	-0.08	neutro e não associado com o mau funcionamento da proteína.

4. Discussão

Este estudo objetivou investigar a variabilidade genética do oncogene E5 de HPV31, buscando inferir as possíveis repercussões biológicas destas variantes na biologia da infecção a partir de uma abordagem *in silico*.

Os nossos resultados detectaram cinco variações não-sinônimas em E5 HPV31, sendo elas N5D, I48V, P56A, V64I e F80I, distribuídas em quatro amostras clínicas. Estas variações já haviam sido identificadas e estão depositadas no *Papilloma virus genome database* (PaVE). A amostra designada HPV31E5_33PE foi a que apresentou o maior número de polimorfismos. Em nosso estudo avaliamos se os polimorfismos tiveram algum impacto no codon usage e consequentemente uma interferência na regulação gênica. O resultado não indicou alterações significativas nos percentuais de utilização de códons preferenciais. Isto sugere que estes polimorfismos não interferem na regulação gênica pelo mecanismo de uso de códons raros. No entanto, não podemos descartar que estes polimorfismos possam interferir na tradução por meio de alterações na estrutura do mRNA.

Atualmente, 16E5 tem sido extensamente estudada (De Freitas et al. 2017). E5 é uma proteína transmembrana de múltipla passagem, cuja modelagem molecular, pode servir de referência para uma análise comparativa entre variantes já catalogadas e as novas, e auxiliar no entendimento do papel exercido pelas variantes do oncogene E5 nos diferentes graus das lesões (Nath et al., 2006; Wetherill et al., 2012; Schiffman et al., 2010, 2013). Embora E5 de HPV16 e HPV31 sejam distintas, observa-se que a característica de hidrofobicidade é mantida, pela natureza química dos seus aminoácidos.

As análises da composição química da oncoproteína 31E5, das amostras estudadas, nos permitiu observar que, a exemplo do que ocorre com a oncoproteína 16E5, temos a manutenção estrutural das três hélices transmembranas, corroborando com os dados da literatura (Alonso & Reed, 2002; Yang et al., 2003). Na análise estrutural da oncoproteína 31E5, os polimorfismos N5D, I48V, P56A, F80I e V64I encontram-se inseridos dentro destas regiões de transmembrana. Apesar de não se observar grandes alterações estruturais nas proteínas variantes encontradas, a mutação P56A foi predita como sendo altamente estabilizante e, desta forma, poderia causar uma alteração no funcionamento da proteína. Para confirmar esse efeito desta mutação, estudos funcionais *in vitro* são necessários.

No sítio mutacional P56A, temos uma das hélices transmembranares mais importantes de E5, por estar relacionada à localização celular da oncoproteína (Cortese et al., 2010), possibilitando a inserção de E5 em várias membranas celulares, como no aparato de Golgi, retículo endoplasmático, endossomos e membrana plasmática (Disbrow et al., 2005; Lewis et al., 2008). Nath et al. (2006) idealizou a construção de algumas variantes do oncogene E5 de HPV16, transfectando-o em células NIH-3T3. Dentre os resultados obtidos, foram observados (1) mudanças no perfil do ciclo celular, onde algumas construções de variantes de 16E5 em cooperação com EGF (*Epidermal Growth Factor*), (2) redução da porcentagem de células na fase G_0G_1 , (3) aumento do período corresponde à fase G2-M com consequente aumento do crescimento celular e (4) mudanças na expressão da proteína p21. Em HPV31, essa região transmembranar de E5 tem forte relação com o crescimento independente de ancoragem, como observado em estudos com células HaCat (Barbaresi et al., 2010). Até o momento, não foi possível avaliar se estes polimorfismos identificados em 31E5 podem interferir na escolha das regiões de inserção desta oncoproteína nas amostras estudadas, ou outras modificações fisiológicas. A troca dos aminoácidos em diferentes regiões 31E5, podem ser um importante alvo de estudo e biomarcador do câncer cervical, já que podem sinalizar um maior potencial de persistência e de progressão das infecções, do mesmo modo como vem sendo observado as análises conduzidas com 16E5 (Hsieh et al., 2000; Nath et al., 2006).

Estas variações intratípicas podem ser uma importante ferramenta para a reconstrução de dados filogenéticos e para o entendimento sobre os mecanismos que conduzem à propagação dos HPVs em diversas populações (Souza et al., 2009). Sendo assim, a análise da heterogeneidade genética do oncogene 31E5 realizada neste estudo, permitiu observar variantes gênicas incluídas nas três linhagens definidas para o HPV31 (A, B e/ou C), circulantes no Estado de Pernambuco, e que não foram identificadas para esta localidade (Gurgel et al., 2015). Os estudos para a análise da heterogeneidade encontrada no genoma do HPV31 até o momento abrangem principalmente as regiões LCR, L1, E6 e E7 (Chagas et al., 2011, 2013; Gurgel et al., 2015). No entanto, os polimorfismos encontrados em E5 também podem contribuir para a oncogenicidade da respectiva variante, uma vez que E5 atua na regulação de proteínas celulares (EGF-R, Bcl-2, Bax, Fas e Calnexina), que em conjunto ajudam com a atividade transformante

mediada pelas oncoproteínas E6 e E7 (DiMaio & PETTI, 2013; Grabowska & Rimer, 2012).

No contexto da imunidade inata, a retenção de moléculas de MHC I e II no interior de células dendríticas ajuda na evasão da resposta imune e sobrevivência celular (Song et al., 2015; Aldarouish and Wang, 2016). Este fato, associado com o downregulation do complexo HLA e bloqueio da ativação dos linfócitos T citotóxicos, interfere na eficiência da resposta imune (Ashrafi et al., 2005; Campo et al., 2010). A análise desta retenção no interior do RE e do aparato de Golgi mostrou que ela pode ocorrer mediante a interação entre a porção hidrofóbica das hélices transmembrana da oncoproteína 16E5 (Cortese et al. 2010). Em nosso estudo, o percentual de interação da sequência 31E5 gerado pelo servidor IEDB, permitiu elencar possíveis epítomos imunogênicos com as moléculas constituintes do MHC-I e MHC-II, podendo representar uma possível relação de supressão das proteínas constituintes do sistema imunológico. O maior percentual de interação relacionado com a formação de epítomos imunogênicos, foi observado com as moléculas HLA-A*02:01, do MHC-I.

É importante ressaltar que o potencial patogênico das variantes de HPV podem sofrer influência de diversos fatores, como as complexas interações dos vírus com a genética dos seus hospedeiros, bem como a prevalência de variantes de HPV circulantes em diferentes populações e regiões geográficas (Xi *et al.*, 2006; Souza *et al.*, 2009). Essa heterogeneidade na etiologia do câncer cervical, tem demonstrado que a principal via de variação genética vem surgindo ao longo do acúmulo de alterações nas sequências de nucleotídeos. Esse processo é conhecido como fixação de linhagem, e o acúmulo destas variações contribuem imensamente para a diversificação e especiação dos papilomavírus humano (Burk *et al.* 2013). Entretanto, no estudo aqui apresentado, a oncoproteína 31E5 mostrou, de forma global, seleção negativa ou purificadora, e nenhuma mutação encontrada estava relacionada com a seleção diversificadora.

Até onde é do nosso conhecimento, esta é a primeira análise sobre a variabilidade genética do oncogene 31E5 circulante na Região Nordeste do Brasil. O estudo destas variantes implica na ampliação conhecimento do potencial de infectividade de variantes de HPV31, além de contribuir para melhorar nosso conhecimento acerca da sua constituição genética e o seu potencial carcinogênico. A análise funcional *in vitro* desta interação será útil para a determinação da qualidade de E5 como alvo terapêutico.

Agradecimentos

Esta pesquisa foi apoiada pela Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE, APQ-0562-2.02/14 e APQ-0748-2.02/17), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, 401305/2014-7 e 447978/2014-4), e Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Conflito de interesse

Os autores declaram não haver conflito de interesses.

Referências

1. Aldarouish M, Wang C. Trends and advances in tumor immunology and lung cancer immunotherapy. *J Exp Clin Cancer Res*. 2016;35:157.
2. Alonso, A. & Reed, J. (2002). Modelling of the human papillomavirus type 16 E5 protein. *Biochim Biophys Acta*. Vol. 1601, No. 1, pp. (9-18).
3. S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, *J. Mol. Biol.* 215 (1990) 403-410.
4. Amaral, C. M. et al. MDM2 polymorphism associated with the development of cervical lesions in women infected with Human papillomavirus and using of oral contraceptives. *Infectious Agents and Cancer*, v. 9, n. 1, p. 24, 18 jul. 2014.
5. Ashrafi GH, Haghshenas MR, Marchetti B, O ' Brien PM, Campo MS. E5 protein of human papillomavirus type 16 selectively downregulates surface HLA class I. *Int J Cancer*. 2005;113:276 – 83.
6. Barbaresi, S., Cortese, M.S., Quinn, J., Ashrafi, G.H., Graham, S.V. & Campo, M.S. (2010). Effects of human papillomavirus type 16 E5 deletion mutants on epithelial morphology: functional characterization of each transmembrane domain. *J Gen Virol*. Vol. 91, No. 2, pp. (521-530).
7. Benkert, P., Tosatto, S.C.E. and Schomburg, D. (2008). "QMEAN: A comprehensive scoring function for model quality assessment. "Proteins: Structure, Function, and Bioinformatics, 71(1):261-277.
8. Bernard, H.-U. et al. Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology*, v. 401, n. 1, p. 70–79, 25 maio 2010.

9. Borzacchiello G, Roperto F, Campo MS and VENUTI A: 1st international workshop on papillomavirus E5 oncogene - a report. *Virology* 408: 135-137, 2010.
10. Burk, R. D.; Harari, A.; Chen, Z. Human papillomavirus genome variants. *Virology*, v.445, n. 1, p. 232-243, 2013. ISSN 0042-6822.
11. Burrioni, E. et al. Codon 72 polymorphism of p53 and HPV type 16 E6 variants as risk factors for patients with squamous epithelial lesion of the uterine cervix. *Journal of medical virology*, v. 85, n. 1, p. 83–90, 2013.
12. Campo MS, Graham SV, Cortese MS, Ashrafi GH, Araibi EH, Dornan ES, et al. HPV-16 E5 down-regulates expression of surface HLA class I and reduces recognition by CD8 T cells. *Virology*. 2010;407:137 – 42.
13. Campo MS, Graham SV, Cortese MS, Ashrafi GH, Araibi EH, Dornan ES, et al. HPV-16 E5 down-regulates expression of surface HLA class I and reduces recognition by CD8 T cells. *Virology* 2010;407:137-42.
14. Chagas, B. S. et al. New variants of E6 and E7 oncogenes of human papillomavirus type 31 identified in Northeastern Brazil. *Gynecologic Oncology*, v. 123, n. 2, p. 284–288, nov. 2011.
15. Chagas, B. S. et al. Novel E6 and E7 oncogenes variants of human papillomavirus type 31 in Brazilian women with abnormal cervical cytology. *Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases*, v. 16, p. 13–18, jun. 2013.
16. Chang JL, Tsao YP, Liu DW, Huang SJ, Lee WH and Chen SL: The expression of HPV-16 E5 protein in squamous neoplastic changes in the uterine cervix. *J Biomed Sci* 8: 206-213, 2001.
17. Cornet, I. et al. Human papillomavirus type 16 E6 variants in France and risk of viral persistence. *Infectious Agents and Cancer*, v. 8, n. 1, p. 4, 23 jan. 2013.
18. Cortese MS, Ashrafi GH, Campo MS. All 4 di-leucine motifs in the first hydrophobic domain of the E5 oncoprotein of human papillomavirus type 16 are essential for surface MHC class I downregulation activity and E5 endomembrane localization. *Int J Cancer*. 2010;126:1675 – 82.
19. D. Xu, Y. Zhang, Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. *Proteins*, 2012, 80, 1715-1735.
20. Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* 9(8), 772.

21. de Araujo Souza PS, Sichero L, Maciag PC. HPV variants and HLA polymorphisms: the role of variability on the risk of cervical cancer. 2009.
22. de Freitas, A.C.F, de Oliveira T.H.A., Barros Junior, M.R., VENUTI, A. hrHPV E5 oncoprotein: immune evasion and related immunotherapies *Journal of Experimental & Clinical Cancer Research* (2017) 36:71
23. DiMaio D, PETTI LM. The E5 proteins. *Virology*. 2013;445:99 – 114.
24. Disbrow, G.L., Hanover, J.A. & Schlegel, R. (2005). Endoplasmic reticulum-localized human papillomavirus type 16 E5 protein alters endosomal pH but not trans-Golgi pH. *J Virol*. Vol. 79, No. 9, pp. (5839-5846).
25. Dong Xu and Yang Zhang. Improving the Physical Realism and Structural Accuracy of Protein Models by a Two-step Atomic-level Energy Minimization. *Biophysical Journal*, vol 101, 2525-2534 (2011).
26. Grabowska AK, Riemer AB. The invisible enemy - how human papillomaviruses avoid recognition and clearance by the host immune system. *Open Virol J*. 2012;6:249 – 56.
27. Gurgel, A. P. A. D. et al. Prevalence and genetic variability in capsid L1 gene of rare human papillomaviruses (HPV) found in cervical lesions of women from North-East Brazil. *BioMed research international*, v. 2013, p. 546354, 2013.
28. Hsieh, C.H., Tsao, Y.P., Wang, C.H., Han, C.P., Chang, J.L., Lee, J.Y., Chen, S.L. Sequence variants and functional analysis of human papillomavirus type 16 E5 gene in clinical specimens. *Arch Virol* (2000) 145: 2273–2284.
29. Hutchinson, E. G. and J. M. Thornton (1996). "PROMOTIF--a program to identify and analyze structural motifs in proteins." *Protein Sci* 5(2): 212-20.
30. Kabsch W, Sander C (1983). "Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features". *Biopolymers*. 22 (12): 2577–637.
31. Kim DE, Chivian D, Baker D. (2004) Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res*. 32 Suppl 2:W526-31.
32. Laskowski R A, MacArthur M W, Moss D S & Thornton J M (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.*, 26, 283-291.
33. Lewis, C., Baro, M.F., Marques, M., Grüner, M., Alonso, A. & Bravo, I.G. (2008). The first hydrophobic region of the HPV16 E5 protein determines protein cellular

- location and facilitates anchorage-independent growth. *Viol J.* Vol. 5, No. 1, pp. (30).
34. Manos, M.M., Ting, Y., Wright, D.K., Lewis, A.J., Broker, T.R., Wolinsky, S.M., 1989. The use of polymerase chain reaction amplification for the detection of genital human papillomavirus. *Cancer Cells* 7, 209–214.
 35. Maufort JP, Shai A, Pitot HC, Lambert PF. A role for HPV16 E5 in cervical carcinogenesis. *Cancer Res* 2010;70:2924-31.
 36. McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. *Bioinformatics*. 2000;16(4):404-5.
 37. Melo, F. and Feytmans, E. (1998) "Assessing Protein Structures with a Non-local Atomic Interaction Energy". *Journal of Molecular Biology* 277, 1141-1152.
 38. MONFRÉ, E. R. M. Avaliação dos níveis de citocinas e HLA-G solúvel em linhagens celulares tumorais de colo uterino tratadas com alcalóides de *Pterogyne nitens*. 2011.
 39. Nath, R. et al. Analyses of variant human papillomavirus type-16 E5 proteins for their ability to induce mitogenesis of murine fibroblasts. *Cancer cell international*, v. 6, n. 1, p.19, 2006. ISSN 1475-2867.
 40. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol*. 2012 May;61(3):539-42.
 41. Schapiro F, Sparkowski J, Adduci A, Supryniewicz F, Schlegel R, Grinstein S. Golgi alkalization by the papillomavirus E5 oncoprotein. *J Cell Biol* 2000;148:305-15.
 42. Schiffman, M.; Wentzensen, N. From human papillomavirus to cervical cancer. *Obstetrics & Gynecology*, v. 116, n. 1, p. 177-185, 2010. ISSN 0029-7844.
 43. Schiffman M, Wentzensen N. Human papillomavirus infection and the multistage carcinogenesis of cervical cancer. *Cancer Epidemiology Biomarkers & Prevention*. 2013;22(4):553-60.
 44. Schiffman M, Khan MJ, Solomon D, Herrero R, Wacholder S, Hildesheim A, et al. A study of the impact of adding HPV types to cervical cancer screening and triage tests. *Journal of the National Cancer Institute*. 2005;97(2):147-50.
 45. Song D, Li H, Li H, Dai J. Effect of human papillomavirus infection on the immune system and its role in the course of cervical cancer. *Oncol Lett*. 2015;10:600 – 6.

46. Staden, R. The Staden sequence analysis package. *Molecular biotechnology*, v. 5, n.3, p.233–241, 1996.
47. Tamura, K. et al. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution*. 2011.
48. van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; vdf Hochschulverlag AG an der ETH Zürich and BIOMOS b.v.: Zürich, Groningen, 1996.
49. Van Doorslaer, K. Evolution of the Papillomaviridae. *Virology, Special Issue: The Papillomavirus Episteme*. v. 445, n. 1–2, p. 11–20, out. 2013.
50. VENUTI, A. et al. Papillomavirus E5: the smallest oncoprotein with many functions. *Molecular cancer*, v. 10, n. 1, p. 1, 2011. ISSN 1476-4598.
51. Wetherill, I. F.; Ross, R.; MacDonald, A. HPV E5: an enigmatic oncoprotein. *Small DNA Tumour Viruses*, v. 55, 2012.
52. Worth CL, Preissner R, Blundell TL. (2011) SDM - a server for predicting effects of mutations on protein stability and malfunction. *Nucleic Acids Res.* 39:W215-22.
53. Xi L, Schiffman M, Koutsky L, et al. Association of human papillomavirus type 31 variants with risk of cervical intraepithelial neoplasia grades 2-3. *Int J Cancer* 2012.
54. Xi LF, Kiviat NB, Hildesheim A, Galloway DA, Wheeler CM, Ho J, et al. Human papillomavirus type 16 and 18 variants: race-related distribution and persistence. *Journal of the National Cancer Institute*. 2006;98(15):1045-52.
55. Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol* 24: 1586–1591.
56. Yang, D.H., Wildeman, A.G. & Sharom, F.J. (2003). Overexpression, purification, and structural analysis of the hydrophobic E5 protein from human papillomavirus type 16. *Protein Expr Purif.* Vol. 30, No. 1, pp. (1-10).
57. Zhou, H., and Zhou, Y. (2002). "Distance-scaled, finite ideal-gas reference state improves structure-derived potentials of mean force for structure selection and stability prediction. " *Protein Sci.* 11:2714-2726.
58. Zur Hausen, H. Papillomavirus infections--a major cause of human cancers. *Biochimica Et Biophysica Acta*, v. 1288, n. 2, p. F55–78, 9 out. 1996.

59. WHO/ICO Information Center on HPV and Cervical Cancer (HPV Information Centre) (2015) Human papillomavirus and related cancers in Brazil.

3.2 ARTIGO 2 - IN SILICO EVALUATION OF HUMAN PAPILLOMAVIRUS TYPE 16 E5 ONCOGENE FROM CERVICAL SAMPLES OF NORTHEAST BRAZILIAN WOMAN

In silico evaluation of Human Papillomavirus type 16 E5 oncogene from cervical samples of Northeast Brazilian woman.

(Submetido à Revista: **Infection, Genetics and Evolution**)

^aAntônio Humberto Pereira da Silva Júnior

^aBárbara Simas Chagas

^bAna Pavla Almeida Diniz Gurgel

^aRuany Cristyne de Oliveira Silva

^aMarconi Rêgo Barros Júnior

^cMarcus Vinicius de Aragão Batista

^aAntonio Carlos de Freitas

^aFederal University of Pernambuco; Department of Genetics; Laboratory of Molecular Studies and Experimental Therapy (LEMTE); Pernambuco, Brazil.

^bDepartment of Engineering and Environment, Federal University of Paraíba, Paraíba – Brazil.

^cDepartment of Biology, Federal University of Sergipe, Sergipe, Brazil.

Short Title: Genetic Diversity in the E5 gene of HPV16.

Corresponding author: Dr. Antonio Carlos de Freitas, Av. Prof. Moraes Rego, 1235, Cidade Universitária, 50670-901, Recife-PE, Brazil. Fax: +55 81 2126-8512. E-mail: acf_ufpe@yahoo.com.br.

Abstract

Infections caused by human papillomavirus are the main cause for the onset and development of cervical cancer precursor lesions. The HPV16 genotype and its variants are the main viral type of high oncogenic risk, which are widely distributed in

Brazil, and in the evaluation of the genetic heterogeneity distribution of its viral oncogenes, the E5 oncogene has not yet been evaluated. The 16E5 oncogene is one of the major oncogenes related to carcinogenesis, expressing itself in the early stages of infection and deregulating innumerable proteins that act on cell cycle regulation, apoptosis, and blockage of immune system molecules. In this study we investigated the variability of circulating 16E5 oncogene in the Northeast of Brazil and evaluated *in silico* the repercussions of these polymorphisms. The E5 sequence was obtained from cervical samples through the primer specific PCR followed by sequencing. Polymorphism identification was made with the help of the MEGA6 software and the phylogenetic tree was constructed utilizing the Bayesian Inference method. The secondary structure was analyzed using the PSIPRED software and the estimated selection pressure was obtained using the PAML software. In order to map the immunogenic epitopes the IEDB platform was used. Amongst the 93 HPV positive samples, 24 were identified as being HPV16. The phylogenetic analysis showed the isolates belonging to the A and D lineages. A total of seven polymorphisms were identified, three of which were synonymous and four non-synonymous. The I44L polymorphism was evaluated as being increase stability of the E5 protein and the I65V polymorphism related to the reduced stability to the E5 oncoprotein, and these polymorphisms are not associated with the anomalous function of the protein. HPV16 E5 protein presented, overall, negative or purifying selection. Residues 48 L ($p = 0.037$) and 65 I ($p = 0.007$) were under diversification selection. A large diversity of immunogenic epitopes were mapped with the HLA-A and HLA-B molecules of MHC-I and HLA-DRB1, -DRB5, -DPA, -DPB, -DQA1 and MHC-II DQB1. The results obtained will aid in the understanding of the mechanisms of fixation of these strains in the population, as well as in the distribution of the variants and in the elaboration of new immunotherapeutic strategies.

Keywords: HPV16; Cervical Cancer; Polymorphisms on E5.

1 INTRODUÇÃO

As infecções causadas pelos Papilomavírus humano (HPVs) tem sido o principal fator determinante para o surgimento do câncer cervical, o qual é considerado uma das principais causas de mortalidade entre as mulheres ao redor do mundo. Recentemente, foram previstos um total de 527,624 novos casos

diagnosticados, tendo sido notificados aproximadamente 266,000 mortes oriundas desta neoplasia (Ferlay et al., 2015). No Brasil são estimados cerca de 16,370 novos casos para o biênio 2018-2019, cuja taxa de incidência na população brasileira é de 15.43/100,000 mulheres, representando a terceira causa mais comum de neoplasia e, na região Nordeste do Brasil, a incidência é de 20.47/100,000 mulheres, valor que se encontra acima da média nacional (INCA, 2018).

O genótipo HPV16 é um dos tipos virais com ampla distribuição geográfica ao redor do mundo (de Sanjose et al., 2010), e este achado tem sido confirmado em estudos na região nordeste do Brasil (Chagas et al., 2011; 2013; Gurgel et al., 2015). As lesões provocadas e a persistência no epitélio infectado por este genótipo pode contribuir para os estágios mais avançados da infecção, culminando com o estabelecimento do câncer cervical (Freitas et al., 2017), aliado à capacidade de evasão do sistema imune (Chang et al., 2001; Ashrafi et al., 2006; Freitas et al., 2017).

Atualmente, o genótipo HPV16 tem sido classificado em quatro linhagens e dez sublinhagens: linhagem A, incluindo as sublinhagens A1 (European), A2 (European), A3 (E linhagem) e A4 (Asian); linhagem B, incluindo as sublinhagens B1 African-1 (Afr1a) e B2 African-1 (Afr1b); linhagem C denominada African-2 (Afr2a); e linhagem D incluindo as sublinhagens D1 North American (NA)1, D2 Asian-American (AA)2 e D3 Asian-American (AA)1 (Burk et al., 2013). Para esta classificação, estes grupos obtiveram mais de 95% de homologia da região LCR (Burk et al., 2013; Cornet et al., 2012; Tornesello et al., 2011; Mendoza et al., 2013). No entanto, outras regiões podem ser consideradas para fins de classificação, como a análise do gene L1, E6 e E7, além da utilização do genoma completo (de Villiers et al., 2004; Bernard et al., 2010; Cornet et al., 2012; Burk et al., 2013; Chen et al., 2013).

Diante desta diversidade, muitos estudos têm procurado identificar e discutir o papel das variantes genéticas do HPV16 (Xi et al., 2011; VENUTI et al., 2011; Chagas et al., 2011, 2013; Kavati, 2012; Anaya-Saavedra et al., 2013; Schiffman e Wentzensen, 2013), bem como identificar a distribuição das variantes dos oncogenes virais, a exemplo dos oncogenes E6 e E7 e das regiões LCR e do gene L1 do HPV16 (Chagas et al., 2011; 2013; Gurgel et al., 2015), visto os diferentes potenciais carcinogênicos apresentados pelas variantes virais (Cornet et al., 2013; Kim et al., 2013; Burk et al., 2013). No entanto, o estudo da identificação e a

distribuição das variantes do oncogene E5 do HPV16 ainda não foi avaliado na população brasileira.

A oncoproteína E5 do HPV16 (16E5) está relacionada com diversos mecanismos celulares nas fases iniciais da infecção e que contribuem com a carcinogênese (VENUTI et al., 2011). Estruturalmente, 16E5 caracteriza-se por uma proteína hidrofóbica, de baixo peso molecular, podendo estar presente inserida em membranas biológicas, como o retículo endoplasmático (ER), no aparato de Golgi (GA), membrana plasmática e membrana nuclear (Borzacchiello et al., 2010; VENUTI et al., 2011; DiMaio e PETTI, 2013).

Em relação ao seu papel biológico, a oncoproteína 16E5 atua desregulando vários mediadores celulares, a exemplo dos fatores de crescimento EGF e dos seus receptores EGFR (DiMaio e PETTI, 2013). É capaz de atuar na inibição da apoptose, induzida pelo Ligante do Fator de Necrose Tumoral (TNFL), e modular genes relacionados à adesão e motilidade celular (Grabowska e Riemer, 2012). No sistema imunológico, a oncoproteína 16E5 atua na modulação da resposta inata e adaptativa, nos processos inflamatórios, bem como na síntese de citocinas, além de diminuir a expressão de moléculas da Classe I e II do Complexo Principal de Histocompatibilidade e dos Interferons (VENUTI et al., 2011; MONFRÉ, 2011; Grabowska e Riemer, 2012; Freitas et al., 2017). Em relação às moléculas MHC-I e MHC-II, alguns estudos mostraram que a oncoproteína E5 do HPV16 atua na downregulation do MHC-I (HLA-I em humanos), ligando-se a sua cadeia pesada e impedindo a sua expressão na superfície celular, evitando a detecção de células infectadas e consequente eliminação; e no bloqueio de moléculas do MHC-II, reduzindo a resposta imune (Gruener et al., 2007; Miura et al., 2010).

16E5 confere um maior potencial carcinogênico quando atua em colaboração com os oncogenes E6 e E7, promovendo a indução da sinalização mitogênica e regulação de importantes moléculas, tais como o Receptor do Fator de Crescimento Epitelial (EGF-R) e demais moléculas envolvidas no controle da diferenciação celular, sobrevivência e crescimento, tais como Bcl-2, Bax, Fas e Calnexina (Bravo e Alonso, 2004; Schiffman et al., 2005; Grabowska e Riemer, 2012; DiMaio e PETTI, 2013).

Nesse contexto, o estudo da heterogeneidade genética do oncogene 16E5 se torna particularmente interessante devido ao seu amplo espectro de atuação, tornando-se um importante veículo para a compreensão da carcinogênese. Dessa

forma, este trabalho teve como objetivo identificar a circulação de variantes do oncogene E5 do HPV16 na região Nordeste do Brasil, e a partir destes dados, analisar *in silico* as possíveis repercussões biológicas destas variantes no curso do processo infeccioso.

2 MATERIAIS E MÉTODOS

2.1 CARACTERÍSTICAS DAS AMOSTRAS ESTUDADAS

Foram analisadas 93 amostras clínicas de escovados cervicais (*Cytobrush*), coletadas de pacientes atendidas pelo serviço de ginecologia do Hospital das Clínicas da Universidade Federal de Pernambuco (HC-UFPE). A idade das pacientes variou de 18 a 82 anos. Todas assinaram o termo de consentimento livre e esclarecido (TCLE) para a participação neste estudo, e os critérios de exclusão foram (i) pacientes portadoras da imunodeficiência humana e (ii) recusa da assinatura do TCLE.

As amostras coletadas foram armazenadas em solução salina tamponada com fosfato (PBS) e armazenada a -20°C. A extração de DNA se deu utilizando-se o Kit GenominPrepBlood DNA isolation (AmershanBioscience), seguindo as instruções do fabricante e, em seguida, quantificadas utilizando-se um espectrofotômetro (NanoDrop).

Este trabalho foi aprovado pelo Comitê de Ética da Universidade Federal de Pernambuco (CAAE: 3606212.7.0000.5208).

2.2 DETECÇÃO E GENOTIPAGEM DO HPV

O DNA obtido foi submetido a um controle de qualidade e integridade, através da amplificação por PCR, do gene WAF-1, que codifica a proteína p21 do genoma humano, como controle endógeno (Amaral et al., 2014). Em seguida, a detecção do DNA do HPV foi realizada por PCR, utilizando-se os primers degenerados MY09/MY11 (Manos et al., 1989), e a tipificação para o HPV16 foi realizada utilizando-se uma PCR tipo-específica com os primers para o oncogene E6 (Chagas et al., 2011). Em seguida, a amplificação do gene E5 do HPV16, se deu com a utilização dos primers específicos, designados como E5_HP16seqF (5'-GCATTGGACAGGACATAATG-3') e E5_HP16seqR (5'-AGAACGTTTGTGTCGCATTG-3').

2.3 SEQUENCIAMENTO DO ONCOGENE E5 DO HPV31

Um total de 24 amostras HPV16 positivas foram enviadas para o sequenciamento. A análise do cromatograma atendeu ao critério de qualidade PHRED ≥ 30 , onde o produto de PCR de cada amostra foi sequenciado 2 vezes com os seguintes pares de primers E5_HP16seqF (5'-GCATTGGACAGGACATAATG-3') e E5_HP16seqR (5'-AGAACGTTTGTGTCGCATTG-3'), utilizando-se o ABI PRISM BigDye® Cycle Sequencing v3.1 Ready Reaction (Applied Biosystems).

2.4 ANÁLISE DAS SEQUÊNCIAS

As sequências de nucleotídeos foram analisadas utilizando-se o programa Gap4 (versão 4.0) e o Pregap4 (versão 1.5), do pacote Staden (Staden, 1996). As sequências foram alinhadas com a sequência de referência da linhagem A1 (GenBank accession number K02718), utilizando a ferramenta Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1990). Também foi utilizado a sequência das amostras depositadas no *Papillomavirus Episteme* (PaVE), para a verificação de variações na sequência gênica do oncogene E5 do HPV16. Para isso, utilizou-se o programa MEGA6 e os alinhamentos foram realizados com o auxílio do algoritmo CLUSTALW (Tamura et al., 2011).

2.5 PREDIÇÃO DE EPÍTOPOS

Para a análise da sequência de aminoácidos e identificação de possíveis epítopos imunogênicos foi utilizado o servidor online Immune Epitope Database and Analysis Resource (IEDB) (<http://www.iedb.org/>).

Os epítopos foram previstos utilizando os parâmetros padrões estabelecidos pelo servidor. Para o MHC-I, foi utilizado o set de alelos mais frequentes na população mundial, comprimento da sequência do epítipo (que variou entre 9-10mer) e a classificação dos melhores epítopos de acordo com o percentile rank <1 e immunogenicity score >0.5 .

Para a identificação de potenciais epítopos para as moléculas do MHC-II, novos critérios foram incluídos, como a análise do percentile value e o IC50 value ($<50\text{nM}$).

2.6 ANÁLISE FILOGENÉTICA

A análise filogenética das sequências do oncogene E5 dos isolados de HPV16 foi realizada a partir da construção de árvores filogenéticas por meio do método de Neighbor-Joining utilizando o programa MEGA7 (Kumar et al., 2016). O modelo evolutivo que mais se adequa ao conjunto de dados foi avaliado utilizando o programa jModelTest 2.1.7 (Darriba et al., 2012). Aplicando o Critério de Informação Bayesiano (BIC), foi estabelecido o modelo TrN+I, que foi utilizado. Para o suporte dos ramos, foram utilizadas 1000 réplicas de bootstrap não-paramétrico.

Quatro agrupamentos foram identificados como linhagens A, B, C e D. Os genomas de referência utilizados para esta classificação foram: K02718 (A1), AF536179 (A2), HQ644236 (A3), AF534061 (A4), AF536180 (B1), HQ644298 (B2), AF472509 (C1), HQ644257 (D1), AY686579 (D2) e AF402678 (D3).

2.7 ANÁLISE DE SELEÇÃO POSITIVA

Para avaliar o tipo de pressão seletiva que está associada ao oncogene E5 de HPV16, foram utilizadas estimativas de Máxima Verossimilhança, utilizando modelos de substituição de códons através do programa CODEML, incorporado no pacote de programas PAML versão 4.9g (Yang, 2007). A predição foi realizada através do cálculo dos parâmetros de seis modelos de substituição de códons, denominados M0, M1, M2, M3, M7 e M8. O Teste de Razão de Verossimilhança (LRT) foi utilizado para avaliar qual modelo melhor se adequa aos dados.

2.8 PREDIÇÃO DA ESTRUTURA SECUNDÁRIA E TERCIÁRIA DA ONCOPROTEÍNA E5 HPV31

Após a identificação das regiões polimórficas no oncogene E5 das amostras variantes do HPV16, foi utilizada a ferramenta PSIPRED para a predição de estrutura secundária da proteína, por se tratar de um método acurado e reunir diferentes métodos de análise em sua plataforma (PSIPRED, GenTHREADER e MEMSAT 2) (McGuffin et al., 2000).

Para a predição da estrutura terciária da oncoproteína E5 de HPV16, foi utilizada a sequência proteica com o número de acesso AAA46938. Uma busca no banco Protein Data Bank (PDB) através da ferramenta blastp foi realizada para identificar uma estrutura molde determinada experimentalmente, com o objetivo de realizar modelagem por homologia. Como E5 é uma proteína transmembrana, há uma considerável dificuldade para cristalizar a mesma, dificultando o processo de

determinação de sua estrutura experimentalmente. Consequentemente, nenhuma similaridade significativa foi obtida com as estruturas do PDB, impossibilitando a utilização da modelagem comparativa.

Desta forma, foi utilizada a modelagem da estrutura terciária de E5 de HPV16 através do método *ab initio*, utilizando o programa QUARK (Xu and Zhang, 2012). Além disso, o modelo estrutural foi predito para cada amostra que apresentou variação na sequência de aminoácidos. Foram obtidos 10 modelos diferentes para cada amostra para serem avaliados, que foram refinados por minimização de energia utilizando o programa ModRefiner (Dong and Zhang, 2011). Os modelos refinados foram avaliados pelo programa PROCHECK (Laskowski et al., 1993) através da avaliação de qualidade estereoquímica dos modelos a partir do gráfico de Ramachandran.

Para identificar o possível efeito das mutações na estrutura e função da oncoproteína 16E5 foi utilizado o programa Site Directed Mutator (SDM), que prediz o efeito das mutações na estabilidade das proteínas (Worth et al., 2011).

2.9 CODON USAGE

Após a obtenção e identificação das sequências variantes do oncogene E5 do HPV16, foi utilizando o software *on line* Graphical Codon Usage Analyser 2.0, disponível em <http://gcua.schoedl.de/>, cujo objetivo foi identificar o uso de códons raros ou frequentes, utilizados pelas variantes, comparando com os códons preferenciais utilizados pelos mamíferos.

3 RESULTADOS

3.1 DETECÇÃO E TIPIFICAÇÃO DO HPV

A partir da análise das 93 amostras HPV positivas, um total de 24 amostras foram genotipadas como HPV16 positivas e a observação de variantes do oncogene E5 foram observadas em 19 isolados (79,16%), sendo as amostras designadas como HPV16E5_06PE, HPV16E5_11PE, HPV16E5_14PE, HPV16E5_16PE, HPV16E5_25PE, HPV16E5_30PE, HPV16E5_33PE, HPV16E5_35PE, HPV16E5_45PE, HPV16E5_49PE, HPV16E5_55PE, HPV16E5_70PE, HPV16E5_71PE, HPV16E5_76PE, HPV16E5_78PE, HPV16E5_79PE, HPV16E5_85PE, HPV16E5_91PE, HPV16E5_93PE, e os seus respectivos resultados histopatológicos apresentados na tabela 1.

Tabela 01 – Resultado histopatológico das amostras de escovados cervicais, relacionadas à presença de polimorfismos em E5 de HPV16.

Amostra cervical	Resultado Histopatológico
HPV16E5_06PE	Sem lesão
HPV16E5_11PE	CIN I
HPV16E5_14PE	Carcinoma invasivo
HPV16E5_16PE	Carcinoma invasivo
HPV16E5_25PE	Sem lesão
HPV16E5_30PE	Carcinoma invasivo
HPV16E5_33PE	CIN III
HPV16E5_35PE	CIN II
HPV16E5_45PE	CIN III
HPV16E5_49PE	CIN I
HPV16E5_55PE	Carcinoma invasivo
HPV16E5_70PE	CIN II
HPV16E5_71PE	CIN III
HPV16E5_76PE	CIN I
HPV16E5_78PE	Sem lesão
HPV16E5_79PE	CIN III
HPV16E5_85PE	CIN III
HPV16E5_91PE	CIN I
HPV16E5_93PE	CIN III

3.2 VARIABILIDADE GENÉTICA DO ONCOGENE HPV16 E5

As variantes do oncogene E5 do HPV16 circulantes na população mundial são classificadas, atualmente, nas linhagens A, B, C e D e nas suas respectivas sublinhagens A1 (European, E), A2 (European, E), A3 (European, E), A4 (Asian, E (As)), B1 (African-1, Afr1a), B2 (African-1, Afr1b), C (African-2, Afr2a), D1 (North American (NA)1), D2 (Asian-American (AA)2) e D3 (Asian-American (AA)1), sendo a sublinhagem A1 denominada a sequência de referência. As análises das variações do oncogene E5 do HPV16 depositadas nos bancos de dados, revelaram a existência de vários sítios polimórficos (tabela 2).

A partir desta investigação *in silico* das variantes do oncogene E5 circulantes em diversas populações mundiais, procuramos identificar a presença/ausência destes polimorfismos na população brasileira, ou até mesmo, buscar a identificação de novas regiões polimórficas deste oncogene.

Após a identificação das amostras clínicas HPV16 positivas, amplificação e sequenciamento do oncogene E5, identificamos um total de 07 sítios polimórficos, dentre as quais 04 foram descritos como substituições não-sinônimas (A3979C; T3988A; C3991G; A4042G) e 03 delas, como substituições sinônimas (C3991T; G4017A; T4089C). Os polimorfismos não-sinônimos refletiram mudanças na sequência dos aminoácidos, indicado pelos códons ATA3979CTA, TTA3988ATA, CTA3991TTA e ATA4042GTA, respectivamente (tabela 3).

Tabela 02 – Variações nucleotídicas do oncogene E5 do HPV16. A tabela mostra as variações presentes nas sublinhagens do HPV16, depositadas no banco de dados do NCBI. Nucleotídeos conservados em relação a sequência de referência K02718 estão marcados com asterístico (*).

Isolates	Positions of E5 nucleotide substitutions												Variant lineage	Sublineage
	3	3	3	3	3	4	4	4	4	4	4	4		
	8	8	9	9	9	0	0	0	0	0	0	0		
	5	6	6	7	9	1	3	4	4	4	5	7		
	8	8	7	9	1	7	4	2	3	9	9	7		
Reference K02718	T	G	A	A	C	G	T	A	T	T	A	A	A	A1
AF536179	*	*	*	C	*	*	G	*	*	G	*	*	*	A2
HQ644236	*	*	*	C	G	*	*	G	*	*	*	*	*	A3
AF534061	*	*	*	C	*	*	G	G	G	*	*	T	*	A4
AF536180	C	A	*	C	T	*	*	T	*	*	*	*	C	B1
HQ644298	C	A	*	C	T	*	*	T	*	*	*	*	C	B2
AF472509	C	*	*	C	G	*	*	G	*	*	C	*	C	C1
HQ644257	C	*	*	C	T	A	*	G	*	*	*	*	C	D1
AY686579	C	*	G	C	G	A	*	G	*	*	*	*	C	D2
AF402678	C	*	*	C	G	A	*	G	*	*	*	*	C	D3

Tabela 03 – Variações nucleotídicas do oncogene E5 do HPV16.

Isolates	Positions of E5 nucleotide substitutions						Variant lineage
	3979	3988	3991	4017	4042	4089	
Reference	A	T	C	G	A	T	
HPV16E5_06PE	*	*	*	*	*	*	
HPV16E5_11PE	*	*	*	*	*	*	
HPV16E5_14PE	C	*	G	A	G		-
HPV16E5_16PE	C	*	G	A	G	*	-
HPV16E5_25PE	*	*	*	*	*	*	
HPV16E5_30PE	*	*	*	*	*	*	
HPV16E5_33PE	*	*	*	*	*	*	
HPV16E5_35PE	C	*	*	*	G	*	-
HPV16E5_45PE	*	*	*	*	*	*	
HPV16E5_49PE	C	A	T	A	G	C	D1
HPV16E5_55PE	C	*	G	A	G	C	D3
HPV16E5_70PE	C	*	G	A	G	C	D3
HPV16E5_71PE	*	*	*	*	*	*	
HPV16E5_76PE	C	*	*	*	G	*	-
HPV16E5_78PE	C	*	G	A	G	C	D3
HPV16E5_79PE	*	*	*	*	*	*	
HPV16E5_85PE	C	*	G	A	G	C	D3
HPV16E5_91PE	C	*	*	*	G	*	-
HPV16E5_93PE	C	*	G	A	G	C	D2
Position of altered códons	44	47	48	56	65	80	
Reference	I	L	L	A	I	F	
Altered amino acid	L	I	L/V	A	V	F	

A partir desta investigação *in silico* das variantes do oncogene E5 circulantes em diversas populações mundiais, procuramos identificar a presença/ausência

destes polimorfismos na população brasileira, ou até mesmo, buscar a identificação de novas regiões polimórficas deste oncogene.

Após a identificação das amostras clínicas HPV16 positivas, amplificação e sequenciamento do oncogene E5, identificamos um total de 07 sítios polimórficos, dentre as quais 04 foram descritos como substituições não-sinônimas (A3979C; T3988A; C3991G; A4042G) e 03 delas, como substituições sinônimas (C3991T; G4017A; T4089C). Os polimorfismos não-sinônimos refletiram mudanças na sequência dos aminoácidos, indicado pelos códons ATA3979CTA, TTA3988ATA, CTA3991TTA e ATA4042GTA, respectivamente (tabela 3).

De acordo com os resultados obtidos e confrontando-os com os resultados histopatológicos das amostras estudadas, foi possível identificar a presença de variantes da oncoproteína E5 nos diferentes estágios das lesões cervicais. A amostra contendo o maior número de sítios polimórficos está representado pelo isolado HPV16E5_49PE (CINI).

Em relação a predição dos epítomos para os alelos MHC-I e MHC-II, está evidente que a diversidade dos alelos MHC é variável entre os inúmeros grupos étnicos, o que também é percebido pela resposta imune, de acordo com a diversidade genotípica encontrada entre as variantes de HPV16.

De acordo com o mapeamento dos epítomos realizado mediante o uso dos padrões estabelecidos na plataforma IEDB, foi observado que a sequência da oncoproteína 16E5 da amostra de referência K02718, formou epítomos imunogênicos com moléculas pertencentes ao MHC-I (HLA-A e HLA-B) e MHC-II (HLA-DR). Para o MHC-I, o principal método escolhido pela plataforma foi o método Consensus (ann/smm) e os principais epítomos estão mapeados com as moléculas HLA-A*26:01 (FIVYIIFVY; posição 60-68), HLA-A*02:01 (LLIRPLLLSV; posição 27-36), que obtiveram o Percentile rank de 0.17. O mesmo método Consensus (ann/smm) adotado pelo software IEDB, permitiu elencar outros epítomos, tais como HLA-A*02:03, HLA-A*02:06, HLA-A*68:01, HLA-A*68:02, HLA-A*33:01, HLA-A*23:01. Em relação aos alelos HLA-B, o melhor Percentile rank foi de 0.41, mediante a análise do método Consensus (ann/smm), representado pelo alelo HLA-B*15:01 (ASAFRCFIVY; posição 54-63); no entanto, outros alelos também foram anotados, tais como o HLA-B*51:01 e o HLA-B*57:01. Não foi observado formação de epítomos imunogênicos com moléculas HLA-C. O escore de imunogenicidade obtido para a amostra de referência foi de 1.85434.

Esta análise preliminar dos principais epítomos mapeados junto à sequência de referência da oncoproteína 16E5 (K02718), permitiu observar se as mudanças estruturais das variantes desta proteína, nos sítios mutacionais, modificou a distribuição e a formação dos epítomos. Utilizando-se o método Consensus (ann/smm), foi observado que os polimorfismos I44L e I65V presentes nas amostras 14PE, 16PE, 35PE, 55PE, 70PE, 76PE, 78PE, 85PE, 91PE e 93PE estão inseridos nos mesmos epítomos imunogênicos anotados para a amostra de referência K02718. Em relação à amostra 49PE, considerada a mais polimórfica, o principal epítomo foi mapeado junto ao alelo HLA-A*02:01 (LLIRPLLSV; posição 27-36); no entanto, considerado apenas os sítios mutacionais não-sinônimos I44L, L47I, L48V e I65V, estas regiões formaram epítomos anotados com vários alelos, dentre os quais HLA-A*23:01, HLA-A*26:01, HLA-A*24:02, HLA-A*68:02, HLA-B*15:01, HLA-B*44:02, HLA-B*51:01, HLA-B*53:01 e HLA-B*57:01.

Na análise dos epítomos mapeados com as moléculas pertencentes ao MHC-II (HLA-DR, -DP and -DQ), para a amostra de referência K02718, observamos que o método mais utilizado foi o Consensus (comb.lib./smm/nn). Os valores de referência do IEDB para a seleção dos epítomos foram percentile rank < 1 e IC50(nM) <50, tendo sido considerados os parâmetros do SMM align IC50 (nM) and NN align IC50 (nM). Sendo assim, observamos que os principais epítomos formados pela amostra de referência K02718 foram representados pelos alelos HLA-DPA1*01:03/DPB1*02:01 (FIVYIIFVYIPLFLI; posição 60-74), HLA-DPA1*03:01/DPB1*04:02 (STYTSLIILVLLLWI; posição 37-51), HLA-DRB1*07:01 (YIPLFLIHTHARFLI, posição 68-82), HLA-DQA1*01:02/DQB1*06:02 (TNLDTASTTLLACFL; posição 2-16), HLA-DQA1*01:02/DQB1*06:02 (MTNLDASTTLLACF, posição 01-15), HLA-DPA1*02:01/DPB1*01:01 (RCFIVYIIFVYIPLF, posição 58-72), HLA-DRB1*11:01 (LIILVLLLWITAASA, posição 42-56), HLA-DRB5*01:01 (VYIIFVYIPLFLIHT, posição 62-76), HLA-DRB1*01:01 (ILVLLLWITAASAFAFR, posição 44-58) e HLA-DRB1*01:01 (IILVLLLWITAASAF, posição 43-57).

As variantes 14PE, 16PE, 35PE, 55PE, 70PE, 76PE, 78PE, 85PE, 91PE e 93PE, que possuem polimorfismo I44L e I65V em comum, mostraram bons epítomos mapeados com os alelos HLA-DPA1*01:03/DPB1*02:01, HLA-DRB1*07:01, HLA-DPA1*02:01/DPB1*01:01, HLA-DQA1*01:02/DQB1*06:02, HLA-DRB1*01:01 e HLA-DPA1*03:01/DPB1*04:02, pelo método Consensus (comb.lib./smm/nn), de acordo

com o SMM align IC50 (nM) and NN align IC50 (nM). Estes mesmos epítomos foram mapeados para a variante 49PE, mesmo sendo o genótipo mais polimórfico circulante nesta população. Todas as moléculas oriundas destes alelos formaram epítomos com as oncoproteínas 16E5 variantes, estando os sítios polimórficos inseridos na sequência do epítomo, como observado para a sequência de referência K02718.

3.3 ANÁLISE FILOGENÉTICA

A árvore filogenética obtida com o gene E5 evidencia um perfil de agrupamento característico das três linhagens de HPV16 (Figura 01). A árvore apresenta agrupamentos que possibilitam a classificação dos isolados em linhagens. Os isolados HPV16E5_06PE, HPV16E5_11PE, HPV16E5_25PE, HPV16E5_30PE, HPV16E5_33PE, HPV16E5_45PE, HPV16E5_71PE, HPV16E5_79PE foram agrupados de forma suportada na sublinhagem A1 (Figura 04). A análise filogenética evidencia que 11 isolados são classificados como linhagem A, e oito isolados como linhagem D (Figura 01).

3.4 ANÁLISE DE PRESSÃO SELETIVA

Na avaliação da pressão seletiva incidente sobre a proteína E5 de HPV16, foi possível observar que esta proteína apresentou valores de ω entre 0 e 1, o que significa que o gene E5 de HPV16, nesta população, se encontra sob seleção. O teste de LRT estimou que os modelos de substituição de códons mais adequados aos dados foram os modelos M2 e M3, apresentando $\omega = 0.91480$ (Tabela 4).

Figura 01 – Árvore filogenética baseado no método de Neighbor-Joining obtida a partir da análise da sequência do gene E5 de HPV16. O suporte dos ramos foi determinado por 1000 réplicas de bootstrap, sendo mostrados apenas os valores maiores que 50%.

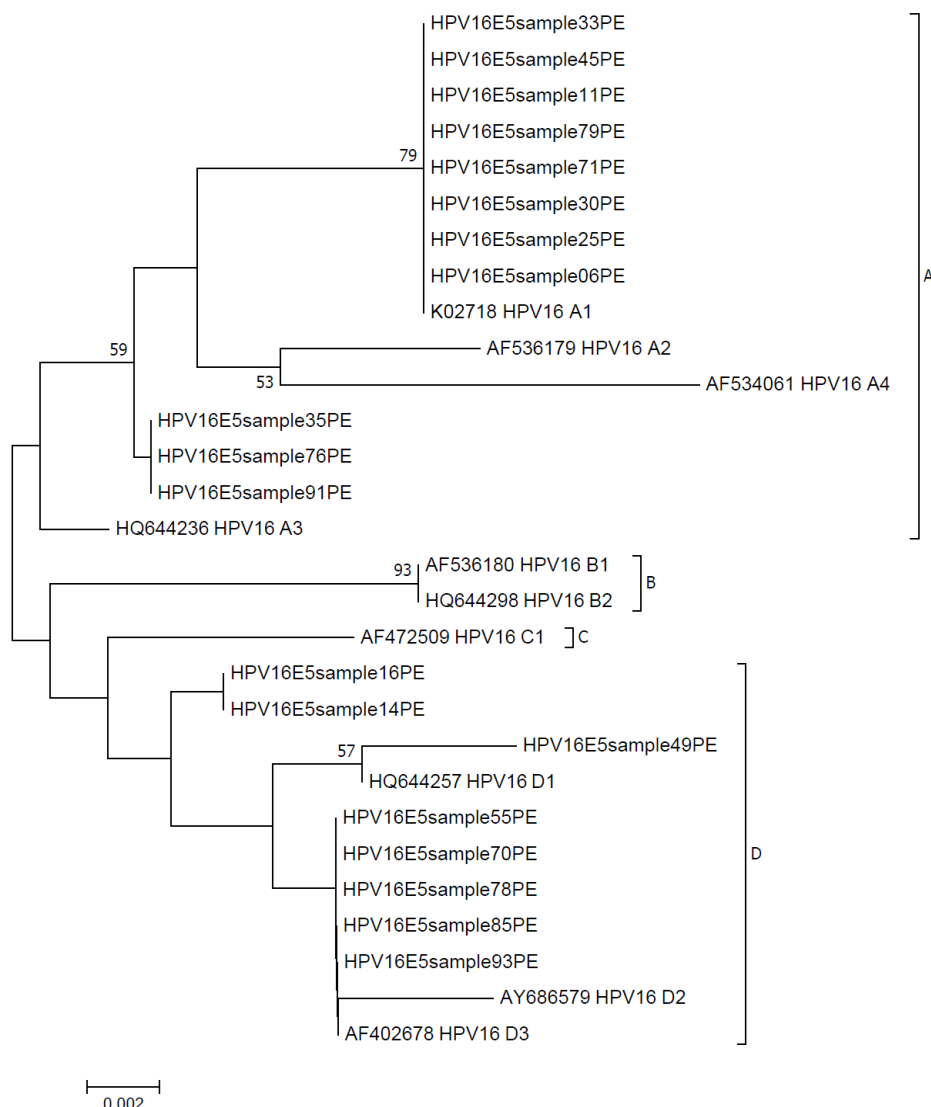


Tabela 04 – Estimativas de pressão seletiva no gene HPV16 E5 com base na máxima verossimilhança.

ORF	Model	lnL	$\omega = dN/dS^*$	Parameters
E5	M0	-475.636423	0.81436	$\omega = 0.81436$
	M1	-472.450423	0.33674	$p_0 = 0.66326$ ($p_1 = 0.33674$) $\omega_0 = 0.00000$; $\omega_1 = 1.00000$
	M2	-466.217010	0.91480	$p_0 = 0.96934$; $p_1 = 0.00000$ ($p_2 = 0.03066$) $\omega_0 = 0.35685$; $\omega_1 = 1.00000$; $\omega_2 = 18.55892$
	M3	-466.217010	0.91480	$p_0 = 0.04767$; $p_1 = 0.92168$ ($p_2 = 0.03066$) $\omega_0 = 0.35685$; $\omega_1 = 0.35685$; $\omega_2 = 18.55893$
	M7	-472.486769	0.30000	$p = 0.00500$; $q = 0.01155$
	M8	-466.218500	0.91560	$p_0 = 0.96934$ ($p_1 = 0.03066$) $p = 55.05023$; $q = 99.00000$; $\omega_s = 18.56975$

O modelo que melhor se ajusta aos dados está em negrito. * significa que os valores foram determinados a partir da média aritmética.

A proteína E5 de HPV16 apresentou, de forma global, seleção negativa ou purificadora. Entretanto, foi verificada a possibilidade de sítios específicos da proteína estarem ou não sob seleção positiva ou diversificadora. A análise sítio-específica revelou que os resíduos de aminoácido 48 L ($p = 0.037$) e 65 I ($p = 0.007$) da proteína E5 de HPV16 se encontrava sob seleção diversificadora, de acordo com os modelos M2, M3 e M8, e foram significativos estatisticamente de acordo com o teste LRT. Os resíduos 44 I ($p = 0.495$), 47 L ($p = 0.493$) foram demonstrados como estando sob seleção positiva apenas pelo modelo M8, mas não foi significativo estatisticamente.

3.5 ANÁLISE ESTRUTURAL DAS MUTAÇÕES 31E5: PERFIL DA ESTRUTURA SECUNDÁRIA E TERCIÁRIA

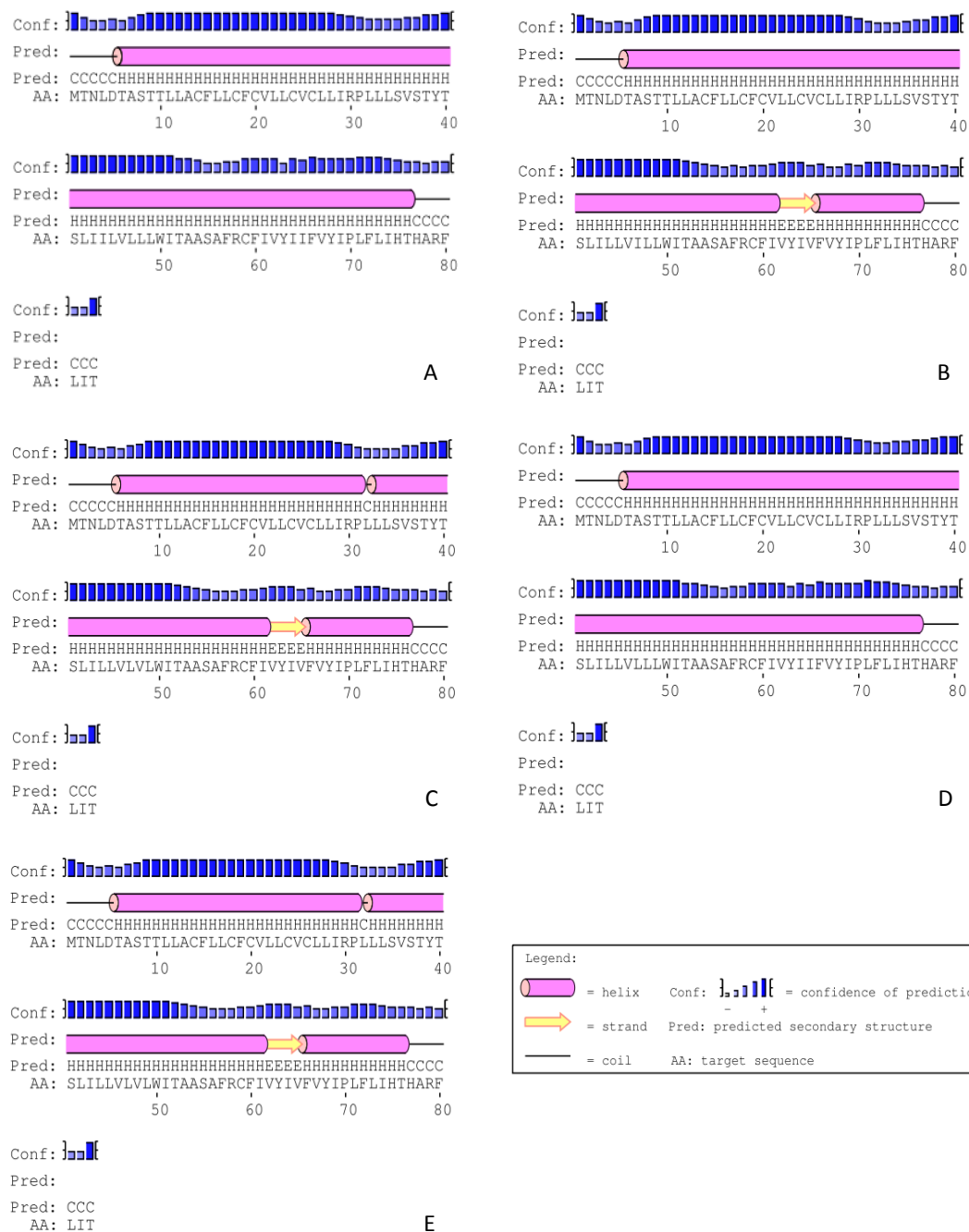
Em HPV16, a oncoproteína E5 é uma proteína transmembrana do tipo α -hélice, de caráter hidrofóbica e constituída por um total de 83 aminoácidos. De acordo com Saavedra et al. (2012), na estrutura secundária do oncogene E5 do HPV16, representada pela sequência de referência K02718, são distinguíveis três regiões importantes para o funcionamento desta oncoproteína: (1) região localizada entre os aminoácidos 1 e 30, responsável pela localização celular e ancoragem independente de crescimento, (2) região localizada entre os aminoácidos 41 e 54 e (3) região localizada entre os aminoácidos 54 ao 78, ambas contendo domínios de ligação da subunidade da bomba de prótons ATPase-16 kDa, responsável pela acidificação dos endossomos.

Muitos dos polimorfismos identificados nas amostras variantes do oncogene 16E5 estão inseridos nas regiões das hélices transmembranares, como observado na estrutura secundária das variantes proteicas estudadas. As amostras com maiores modificações estruturais estão representadas pelas amostras HPV16E5_14PE, HPV16E5_16PE, HPV16E5_49PE, HPV16E5_55PE, HPV16E5_70PE, HPV16E5_78PE, HPV16E5_85PE e HPV16E5_93PE (Figura 3).

Nas amostras variantes 14PE, 16PE, 55PE, 70PE, 78PE, 85PE e 93PE os aminoácidos constituintes da estrutura secundária apresentaram-se da seguinte maneira: amino acids 01 to 05 – coil (C); amino acids 06 to 31 – helix (H); amino acid 32 (C); amino acids 33 to 61 (H); amino acids 62 to 65 – strand (E); amino acids 66 to 76 (H); and amino acids 77 to 83 (C). Da mesma maneira, a amostra mais

polimórfica 49PE, apresentou as seguintes mudanças na estrutura secundária: amino acids 01 to 05 (C); amino acids 06 to 61 (H); amino acid 62 to 65 (E); amino acids 66 to 76 (H); and amino acids 77 to 83 (C) (Figura 02).

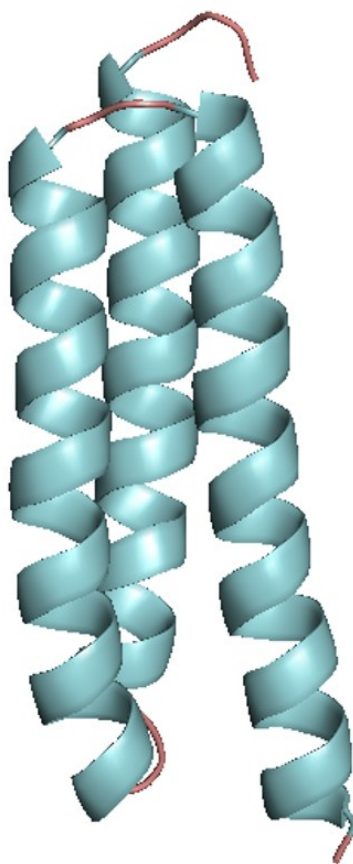
Figura 02 – Predição da estrutura secundária da oncoproteína 16E5. A: predição da amostra de referência K02718; B: predição do isolado 49PE; C: predição dos isolados 14PE e 16PE; D: predição dos isolados 35PE, 76PE e 91PE; E: predição dos isolados 55PE, 70PE, 78PE, 85PE e 93PE. Os resultados gráficos foram otidos a partir do servidor PSIPRED.



Como apenas as amostras 49PE e 14PE apresentaram mutações não-sinônimas, somente elas foram comparadas com a estrutura da proteína de referência. O melhor modelo da proteína E5 de HPV16 de referência apresentou 98.7% dos resíduos nas regiões mais favoráveis no gráfico de Ramachandran, mostrando que foi possível obter um modelo bastante satisfatório. A amostra 14PE apresentou 91.9% dos resíduos nas regiões mais favoráveis no gráfico de Ramachandran. A amostra 49PE apresentou 93.2% dos resíduos nas regiões mais favoráveis no gráfico de Ramachandran. Desta forma, as análises mostram que foi possível obter modelos satisfatórios para a avaliação estrutural das mutações.

A proteína E5 de HPV16 possui 83 resíduos de aminoácidos, formada por três alfa-hélices ligadas por pequenas alças, estabelecendo um grande domínio transmembrana. Esta proteína é formada basicamente por resíduos de aminoácidos hidrofóbicos compondo as regiões de hélice (Figura 03).

Figura 03 – Modelo tridimensional da proteína E5 de HPV16 estabelecida a partir do método *ab initio*. A estrutura evidencia os três domínios de alfa-hélice.



As mutações não-sinônimas encontradas nas variantes descritas neste estudo estão localizadas em regiões de alfa-hélice. Apesar das mutações observadas, o alinhamento das estruturas da proteína E5 de referência com as duas amostras que apresentaram as mutações não-sinônimas evidencia que não houveram grandes alterações na estrutura da proteína (Figura 04).

Na avaliação do impacto destas mutações sobre a estabilidade da proteína, observamos que a mutação I44L apresentou valor predito de $\Delta\Delta G$ de 0.37, indicando ser uma mutação que causa um pequeno aumento da estabilidade da proteína. As mutações L47I, L48V e I65V apresentaram valores preditos de $\Delta\Delta G$ negativos, indicando ser mutações ligeiramente desestabilizantes (Tabela 5). Entretanto, apesar destas mutações estarem relacionadas com alterações estruturais da proteína E5, nenhuma delas foi predita estar associada com alterações funcionais da proteína.

Figura 04 – Alinhamento estrutural da proteína E5 de HPV16 de referência (marrom), HPV16E5_14PE (azul) e HPV16E5_49PE (rosa).

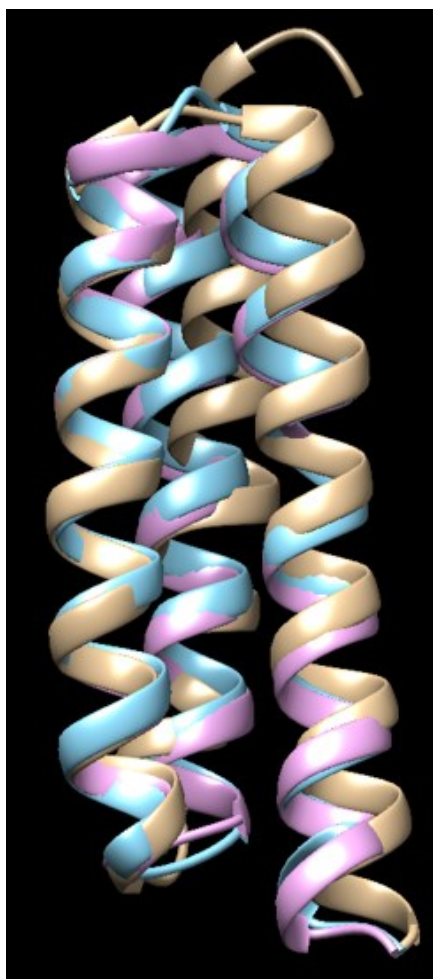


Tabela 5 – Impacto das mutações na estrutura e/ou função da proteína E5 de HPV16.

Mutation	Wild Residue	Residue Position	Mutant Residue	Predicted $\Delta\Delta G$	Outcome
1	I	44	L	0.37	Maior estabilidade
2	L	47	I	-0.48	Estabilidade reduzida
3	L	48	V	-0.67	Estabilidade reduzida
4	I	65	V	-0.35	Estabilidade reduzida

3.6 VARIABILIDADE E USOS DE CÓDONS PREFERENCIAIS

As mutações não-sinônimas I44L, L47I, L48V e I65V foram analisadas, utilizando-se os padrões do Graphical Codon Usage Analyser 2.0, para averiguar se as mudanças dos códons nestas posições refletiram-se num maior percentual de utilização de códons preferenciais em *Homo sapiens* (tabela 6). Os resultados obtidos não mostraram significância na utilização de códons preferenciais nas regiões contendo as substituições não-sinônimas, não apresentando vantagens adaptativas no curso da infecção pelas variantes 16E5.

Tabela 06 – Análise dos sítios polimórficos e das mudanças na sequência de utilização de códons preferenciais nas variantes do oncogene E5 do HPV16.

Isolate	Position of non-synonymous polymorphisms	Codon Usage	Percentage of use (%)	Relative adaptiveness (%)
K02718	44	ATA	17	36
	47	TTA	8	20
	48	CTA	7	18
	65	ATA	17	36
14PE and 16PE	44	CTA	7	18
	47	TTA	8	20
	48	GTA	12	26
	65	GTA	12	26
49PE	44	CTA	7	18
	47	ATA	17	36
	48	TTA	8	20
	65	GTA	12	26
35PE, 76PE and 91PE	44	CTA	7	18
	47	TTA	8	20
	48	CTA	7	18
	65	ATA	17	36
55PE, 70PE, 78PE, 85PE and 93PE	44	CTA	7	18
	47	TTA	8	20
	48	GTA	12	26
	65	GTA	12	26

A tabela mostra a mudança dos códons nas regiões não-sinônimas e a comparação do percentual de utilização dos códons no oncogene E5 do HPV16, da amostra de referência K02718, com as variantes identificadas neste estudo. *Dados gerados pelo Graphical Codon Usage Analyser 2.0.

4 DISCUSSÃO

No Brasil, o HPV16 é um dos tipos virais com ampla distribuição nas diversas regiões do País, em especial na região Nordeste (Chagas et al., 2011, 2013; Gurgel et al., 2015). Inúmeras variantes do HPV16 têm sido identificadas em diferentes localidades e grupos étnicos ao redor do mundo, e exibe diferentes potenciais oncogênicos, indicando que algumas linhagens específicas podem afetar a persistência da infecção pelo HPV e a progressão das lesões precursoras do câncer cervical (Li et al., 2011; Sun et al., 2013). As variantes do HPV16 diferem na prevalência e nas propriedades bioquímicas e biológicas, com implicações incertas na etiologia do câncer do colo do útero (Pleasa et al., 2014). No entanto, a distribuição de variantes do oncogene 16E5 no Brasil, ainda não foi avaliado.

Os resultados apresentados neste estudo apontam para a circulação de variantes do oncogene E5, pertencentes às linhagens A (linhagem europeia) e D (linhagem não-européia) do HPV16, no Nordeste do Brasil. Diversos estudos têm mostrado que as variantes do HPV16 apresentam diferentes efeitos no que tange à sua patogenicidade, sendo observado mudanças na taxa de replicação celular com ampla multiplicação, além de modificações na transcrição de alguns genes, como aqueles relacionados à apoptose e, este cenário, contribui com a eficiência, persistência e a progressão da infecção viral (Bernard et al., 2010; Schiffman e Wentzensen, 2010; Chagas et al., 2011, 2013; de Freitas et al., 2012; Burk et al., 2013). Por exemplo, em HPV16, tem-se observado que os polimorfismos presentes nas linhagens não-européias (B, C e D), as tornam com uma maior capacidade de desenvolver o seu potencial carcinogênico, em relação aos polimorfismos identificados em variantes da linhagem europeia (A) (Schiffman et al., 2010; Smith et al., 2011).

Os polimorfismos não-sinônimos I44L, L47I, L48V e I65V identificados neste estudo, simbolizaram mudanças estruturais na sequência das oncoproteínas variantes de 16E5. Nath et al. (2006), avaliaram o impacto das seguintes variantes 16E5: Leu⁴⁴Val⁶⁵, Thr³⁷Leu⁴⁴Val⁶⁵, Thr⁶⁴ e Tyr⁸Leu⁴⁴Val⁶⁵, cujos produtos de amplificação foram clonados em *pcDNA3.1Myc-His* e, posteriormente, avaliaram os resultados destas construções em células NIH-3T3 transfectadas. Os resultados observados foram um aumento nos níveis de expressão da construção Tyr⁸Leu⁴⁴Val⁶⁵ em relação a Thr³⁷Leu⁴⁴Val⁶⁵ e Leu⁴⁴Val⁶⁵, e baixos níveis de transcritos gerados pela variante Thr⁶⁴.

Os nossos resultados apontaram para a ampla circulação das variantes Leu⁴⁴ e Val⁶⁵ na população do Nordeste brasileiro, o que também se observa no Reino Unido. O polimorfismo I44L foi predito como uma mutação que está relacionada ao aumento da estabilidade da proteína, com o $\Delta\Delta G$ predito igual a 0.37, enquanto I65V está relacionado com a perda da estabilidade com os seu $\Delta\Delta G$ predito igual a -0.35. Estes resultados são reflexos de análises isoladas dos polimorfismos. No entanto, as variantes apresentadas neste estudo apresentam múltiplos polimorfismos, o que pode modificar o comportamento celular. A variante Leu⁴⁴Val⁶⁵ foi a responsável em provocar a maior redução na fase G₀-G₁ e, nas etapas subsequentes do ciclo celular, um aumento da fase G₂-M, sugerindo que estas variantes estejam relacionadas ao crescimento celular (Nath et al., 2006).

Embora a mutação I65V apresentar um valor $\Delta\Delta G$ negativo, a sua presença não foi associada com o mau funcionamento da proteína. Dessa forma, as análises *in silico* permitiram avaliar que as mudanças estruturais na oncoproteína 16E5, ocupam sítios relacionados com dois domínios funcionais da oncoproteína 16E5, onde o primeiro domínio compreende a região relacionada ao voltage gating motif e, o segundo, ao sítio da bomba 16 kDa ATPase, relacionado ao processo de acidificação dos endossomos. Estas observações estão de acordo com o observado por Ullman et al. (1994) e Rodriguez et al. (2000), respectivamente. I65V e I44L também foram relacionados à repressão de p21 (Nath et al., 2006). Em relação às mudanças observadas na região correspondente ao domínio voltage gating motif, a mutação I65V esteve presente em 57,89% das variantes estudadas, e que são pertencentes às linhagens A e D.

Apesar dos resultados apresentados por Nath et al. (2006), ainda não dispomos de relatos sobre o papel exercido pelos polimorfismos em 16E5 isoladamente, nem tampouco de construções gênicas contendo múltiplas substituições de nucleotídeos numa mesma variante, como observado na variante 49PE, o que justifica a análise preliminar *in silico* apresentada.

Estruturalmente, 16E5 corresponde a uma proteína α -hélice transmembrana (Bible et al., 2000; Nath et al., 2006; Saavedra et al., 2012). As substituições não-sinônimas I44L, L47I, L48V e I65V estão inseridas na segunda e terceira hélices da oncoproteína 16E5 e estas mudanças podem afetar o papel funcional das conexinas presentes nas membranas celulares a qual E5 se insere para contribuir com o estabelecimento da infecção pelo HPV16 (Bible et al., 2000). As variações inseridas

entre os resíduos de aminoácidos 41-54 e a presença de polimorfismos nesta região pode estar relacionada ao maior potencial de infecção pela variante.

Plesa et al. (2014), realizaram uma análise dos polimorfismos circulantes na Romênia e as mutações mais prevalentes foram T3904A, T3910C, G3919C, T3928C, A3978C, T3988C, T3991G, T4033C, A4041G/C, onde os respectivos códons mutantes passaram a ser representados pelos seguintes resíduos de aminoácidos: F19I, V21A, C24S, L27P, I44L, L47S, L48A, V62A, I65V/L. Semelhante a este estudo, identificamos os polimorfismos I44L e I65V circulantes no Brasil. Além disso, foi observado uma variabilidade nos resíduos 47 e 48 desta proteína. Enquanto na população da Romênia, os polimorfismos encontrados foram L47S e L48A, no Brasil, as variantes apresentaram o polimorfismo L47I e L48V. No entanto, nenhum estudo avaliou o impacto destas variações.

O sequenciamento do oncogene 16E5 e a análise da sua heterogeneidade também foi avaliado em amostras de pacientes com neoplasia intraepitelial grau III (CIN III) e carcinoma cervical na Suécia, onde um total de 09 variações foram identificadas, dentre as quais cinco delas não-sinônimas, e as mais comuns ocorreram na posição 3979 e 4042 (Hu et al., 2001). Este resultado reflete a circulação das variantes nos resíduos 44 e 65 da oncoproteína E5 na Suécia e, como já apresentado, esta variação também esteve amplamente distribuída na população do Reino Unido. A análise filogenética permitiu elencar a distribuição das variantes entre as linhagens A e D do HPV16 sendo, portanto, as principais encontradas circulantes no Nordeste do Brasil, o que simboliza uma distribuição geográfica mediante o processo de migração populacional.

Do ponto de vista filogenético, através da análise molecular das variações intratípicas, observamos a formação de dois clados bem distintos, onde os isolados estão distribuídos nas linhagens A e D. Além disso, comparamos as taxas de substituições sinônimas e não-sinônimas, cujos resultados mostraram que os modelos M2 e M3 representam o modelo de substituição de códons mais adequado, estando o oncogene 16E5 sob seleção negativa ou purificadora. Adicionalmente, verificamos que a análise sítio-específica dos resíduos de aminoácidos 48L e 65I, encontravam-se sob seleção diversificadora, e os resíduos 44I e 47L, sob seleção positiva. Chen et al. (2005), conduziu o sequenciamento do genoma completo de doze isolados de HPV16 derivado de células esfoliadas cervicais, analisou a evolução molecular e a pressão seletiva das variantes, utilizando estimativas de

Máxima Verossimilhança pelos modelos de substituição de códons tendo sido observado que o oncogene 16E5 encontra-se sob seleção positiva e que a resposta a esta pressão de seleção pode contribuir para o potencial carcinogênico das variantes do HPV16.

Todas estas variações intratípicas se tornaram objeto de estudo para a avaliação dos epítomos imunogênicos formados pelas variantes 16E5. Este fato se deve, principalmente, por saber que esta oncoproteína tem sido considerada um alvo de estudo para o desenvolvimento de estratégias imunoterapêuticas e para o desenvolvimento de vacinas (Cordeiro et al., 2014). Sendo assim, uma análise *in silico* prévia contribui para o delineamento de novas estratégias imunoterapêuticas.

Kumar et al. (2015), idealizou um estudo *in silico* objetivando a identificação de epítomos da oncoproteína E5 do HPV16, utilizando a sequência de referência K02718 do HPV16. Para o mapeamento dos epítomos, foi utilizado a plataforma IEDB, cujos os parâmetros para o MHC-I foram percentile value < 1.0 e, para o MHC-II, além do percentile value <1.0, foram considerados os parâmetros de IC50 value (<50nM). Considerando os principais alelos distribuídos na população indiana, o servidor IEDB listou, para o MHC-I, os principais epítomos relacionados aos alelos HLA-A*24:02, HLA-A*26:01 e HLA-A*14:02 e, para o MHC-II, os principais epítomos mapeados em HLA-DRB1*10:01, HLA-DRB1*14:01 e HLA-DRB1*16:01. Vale salientar que os epítomos listados estão em concordância apenas para a amostra de referência K02718 e considerando os alelos com ampla distribuição na população indiana.

Na análise das variantes de 16E5 apresentadas neste estudo, as descrições de novos alelos relacionados à formação de epítomos podem ser somadas aos já descritos, uma vez que foram adotados os mesmos parâmetros estabelecidos pela plataforma IEDB. Além disso, neste trabalho, também consideramos os epítomos constituídos a partir da análise *in silico* das variantes do oncogene E5. 16E5 por se tratar de uma oncoproteína que atua nas fases iniciais do ciclo infeccioso, os resultados obtidos justificam o desenvolvimento de novas estratégias imunoterapêuticas.

Agradecimentos

Esta pesquisa foi apoiada pela Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE, APQ-0562-2.02/14 e APQ-0748-2.02/17),

Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, 401305/2014-7 e 447978/2014-4), e Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Conflito de interesse

Os autores declaram não ter conflito de interesses.

Referências

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J. Mol. Biol.* 215, 403-410.
- Amaral, C.M., Cetkovská, K., Gurgel, A.P., Cardoso, M.V., Chagas, B.S., Paiva Júnior, S.S., de Lima, R. de C., Silva-Neto, J.C., Silva, L.A., Muniz, M.T., Balbino, V.Q., de Freitas, A.C., 2014. MDM2 polymorphism associated with the development of cervical lesions in women infected with Human papillomavirus and using of oral contraceptives. *Infectious Agents and Cancer.* 9, 24.
- Ashrafi, G.H., Haghshenas, M.R., Marchetti, B., O'Brien, P.M., Campo, M.S., 2005. E5 protein of human papillomavirus type 16 selectively downregulates surface HLA class I. *Int. J. Cancer.* 113, 276-83.
- Benkert, P., Tosatto, S.C.E., Schomburg, D., 2008. "QMEAN: A comprehensive scoring function for model quality assessment. *Proteins: Structure, Function, and Bioinformatics.* 71, 261-77.
- Bernard, H-U., Burk, R.D., Chen, Z., van Doorslaer, K., zur Hausen, H., de Villiers, E-M., 2010. Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology.* 401, 70-9.
- Bible, J.M., Mant, C., Best, J.M., Kell, B., Starkey, W.G., Raju, K.S., et al., 2000. Cervical lesions are associated with human papillomavirus type 16 intratypic variants that have high transcriptional activity and increased usage of common mammalian códons. *Journal of General Virology.* 81, 1517-1527.
- Bravo, I.G., Alonso, A., 2004. Mucosal human papillomavirus encode four diferente E5 proteins whose chemistry and phylogeny correlate with malignant or benign growth. *Journal of Virology.* 13613-13626.
- Burk, R.D., Harari, A., Chen, Z., 2013. Human papillomavirus genome variants. *Virology.* 445, 232-43.

- Chagas, B.S., Batista, M.V.A., Guimarães, V., Balbino, V.Q., Crovella, S., Freitas, A.C., 2011. New variants of E6 and E7 oncogenes of human papillomavirus type 31 identified in Northeastern Brazil. *Gynecologic Oncology*. 123, 284-88.
- Chagas, B.S., Batista, M.V.D.A., Crovella, S., Gurgel, A.P.A.D., Serra, I.G.S.S., Silva-Neto, J.C., Amaral, C.M.M., Balbino, V.Q., Muniz, M.T., de Freitas, A.C., 2013. Novel E6 and E7 oncogenes variants of human papillomavirus type 31 in Brazilian women with abnormal cervical cytology. *Infection, Genetics and Evolution*. 16, 13-18.
- Chang, J.L., Tsao, Y.P., Liu, D.W., Huang, S.J., Lee, W.H., Chen, S.L., 2001. The expression of HPV-16 E5 protein in squamous neoplastic changes in the uterine cervix. *J. Biomed. Sci.* 8, 206-13.
- Chen, Z., Schiffman, M., Herrero, R., DeSalle, R., Anastos, K., Segondy, M., Sahasrabuddhe, V.V., Hsing, A.W., Burk, R.D., 2013. Evolution and Taxonomic Classification of Alphapapillomavirus 7 Complete Genomes: HPV18, HPV39, HPV45, HPV59, HPV68 and HPV70. *PLoS ONE*. 8, e72565.
- Cordeiro, M.N., Paolini, F., Massa, S., Curzio, G., Illiano, E., Duarte Silva, A.J., et al., 2014. Antitumor effects of genetic vaccines against HPV major oncogenes. *Hum Vaccin Immunother.* e34303.
- Cornet, I., Gheit, T., Franceschi, S., Vignat, J., Burk, R.D., Sylla, B.S., Tommasino, M., Clifford, G.M., IARC HPV Variant Study Group, 2012. Human Papillomavirus Type 16 Genetic Variants: Phylogeny and Classification Based on E6 and LCR. *J Virol*. 86, 6855 – 61.
- Cornet, I., Gheit, T., Clifford, G.M., Combes, J-D., Dalstein, V., Franceschi, S., Tommasino, M., Clavel, C., 2013. Human papillomavirus type 16 E6 variants in France and risk of viral persistence. *Infectious Agents and Cancer*. 8, 4.
- Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*. 9, 772.
- de Freitas, A.C., Gurgel, A.P.A.D., Chagas, B.S., Coimbra, E.C., Amaral, C.M.M., 2012. Susceptibility to cervical cancer: an overview. *Gynecologic Oncology*. 304-11.
- de Freitas, A.C.F., de Oliveira, T.H.A., Barros Junior, M.R., VENUTI, A., 2017. hrHPV E5 oncoprotein: immune evasion and related immunotherapies. *Journal of Experimental & Clinical Cancer Research*. 36, 71.
- de Sanjose, S., Quint, W.G., Alemany, L., Geraets, D.T., Klaustermeier, J.E., Lloveras, B., et al., 2010. Human papillomavirus genotype attribution in invasive

- cervical cancer: a retrospective cross-sectional worldwide study. *Lancet Oncol.* 11, 1048–1056.
- de Villiers, E.M., Fauquet, C., Broker, T.R., Bernard, H-U., zur Hausen, H., 2004. Classification of papillomaviruses. *Virology.* 324, 17-27.
- DiMaio, D., PETTI, L.M., 2013. The E5 proteins. *Virology.* 445, 99-114.
- Ferlay, J., Soerjomataram, I., Ervik, M., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D.M., Forman, D., Bray, F., 2015. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer.* 136, 359-86.
- Grabowska, A.K., Riemer, A.B., 2012. The invisible enemy - how human papillomaviruses avoid recognition and clearance by the host immune system. *Open Virol. J.* 6, 249-56.
- Gruener, M., Bravo, I.G., Momburg, F., Alonso, A., Tomakidi, P., 2007. The E5 protein of the human papillomavirus type 16 down-regulates HLA-I surface expression in calnexin-expressing but not in calnexin-deficient cells. *Virol. J.* 4, 116.
- Gurgel, A.P., Chagas, B.S., do Amaral, C.M.M., Nascimento, K.C., Leal, L.R., Silva Neto, J. da C., Cartaxo Muniz, M.T., de Freitas, A.C., 2015. Prevalence of papillomavirus variants and genetic diversity in the L1 gene and long control region of HPV16, HPV31 and HPV58 found in North-East Brazil. *BioMed Res. Int.* 2015, 130828.
- Hu, X., Pang, T., Guo, Z., Pontén, J., Nistér, M., Afink, G.B., 2001. Oncogene lineages of human papillomavirus type 16 E6, E7 and E5 in preinvasive and invasive cervical squamous cell carcinoma. *J. Pathol.* 195, 307-311.
- Hutchinson, E. G., Thornton, J.M., 1996. PROMOTIF--a program to identify and analyze structural motifs in proteins. *Protein Sci.* 5, 212-20.
- Instituto Nacional de Câncer José Alencar Gomes da Silva, Ministério da Saúde, 2018. Estimativas 2018: Incidência de câncer no Brasil. Available at <http://www.inca.gov.br>. (accessed 27 May 2018).
- Kabsch, W., Sander, C., 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers.* 22, 2577–637.
- Kim, D.E., Chivian, D., Baker, D., 2004. Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res.* 32, W526-31.

- Kumar, A., Yadav, I.S., Hussian, S., Das, B.C., Bharadwaj, M., 2015. Identification of immunotherapeutic epitope of E5 protein of human papillomavirus-16: an in silico approach. *Biologicals xxx*. 1-5.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* 26, 283-91.
- Li, N., Franceschi, S., Howell-Jones, R., Snijders, P.J., Clifford, G.M., 2011. Human papillomavirus type distribution in 30,848 invasive cervical cancers worldwide: Variation by geographical region, histological type and year of publication. *Int J Cancer*. 128, 927–935.
- Manos, M.M., Ting, Y., Wright, D.K., Lewis, A.J., Broker, T.R., Wolinsky, S.M., 1989. The use of polymerase chain reaction amplification for the detection of genital human papillomavirus. *Cancer Cells*. 7, 209-14.
- McGuffin, L.J., Bryson, K., Jones, D.T., 2000. The PSIPRED protein structure prediction server. *Bioinformatics*. 16(4), 404-5.
- Melo, F., Feytmans, E., 1998. Assessing Protein Structures with a Non-local Atomic Interaction Energy. *Journal of Molecular Biology*. 277, 1141-52.
- Mendoza, L., Picconi, M.A., Mirazo, S., Mongelós, P., Giménez, G., Basiletti, J., Arbiza, J., 2013. Distribution of HPV-16 variants among isolates from Paraguayan women with different grades of cervical lesion. *Int J Gynecol Obstet*. 122, 44 – 7.
- Miura, S., Kawana, K., Schust, D.J., Fujii T., Yokoyama, T., Iwasawa, Y., Nagamatsu, T., Adachi, K., Tomio, A., Tomio, K., Kojima, S., Yasugi, T., Kozuma, S., Taketani, Y., 2010. CD1d, a sentinel molecule bridging innate and adaptive immunity, is downregulated by the human papillomavirus (HPV) E5 protein: a possible mechanism for immune evasion by HPV. *J. Virol*. 84, 11614 – 11623.
- Nath, R., Mant, C.A., Kell, B., Cason, J., Bible, J.M., 2006. Analyses of variant human papillomavirus type-16 E5 proteins for their ability to induce mitogenesis of murine fibroblasts. *Cancer cell international*. 6, 19.
- Pleasa, A., Anton, G., Iancu, I.V., Diaconu C. C., Huica, I., Stanescu, A. D., Socolov D., Nistor, E., Popa, E. Stoian, M., Botezatu, A., 2014. Molecular variants of human papilloma virus 16 E2, E4, E5, E6 and E7 genes associated with cervical neoplasia in Romanian patients. *Arch Virol*. 159, 3305–3320.
- Rodriguez, M.I., Finbow, M.E., Alonso, A., 2000. Binding of human papillomavirus 16 E5 to the 16 kDa subunit c (proteolipid) of the vacuolar H⁺-ATPase can be

dissociated from the E5-mediated epidermal growth factor receptor overactivation. *Oncogene*. 19, 3727-3732.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539-42.

Saavedra, A.P., Torres, T.P., Amparán, L.C., Gama, M.M., Jaimes, A.M.G., Guadarrama, F.E., Xicotencatl, L.G., 2012. Molecular bases of human papillomavirus pathogenesis in the development of cervical cancer, human papillomavirus and related diseases – From Bench to Bedside – Research aspects, ISBN: 978-953-307-855-7, InTech.

Schiffman, M., Wentzensen, N., 2010. From human papillomavirus to cervical cancer. *Obstetrics & Gynecology*. 116, 177-85.

Schiffman, M., Wentzensen, N., 2013. Human papillomavirus infection and the multistage carcinogenesis of cervical cancer. *Cancer Epidemiology Biomarkers & Prevention*. 22, 553-60.

Smith, B., Chen, Z., Reimers, L., van Doorslaer, K., Schiffman, M., Desalle, R., Herrero, R., Yu, K., Wacholder, S., Wang, T., Burk, R.D., 2011. Sequence imputation of HPV16 genomes for genetic association studies. *PLoS One* 6, e21375.

Staden, R., 1996. The Staden sequence analysis package. *Molecular biotechnology*. 5, 233-41.

Sun, Z., Lu, Z., Liu, J., Wang, G., Zhou, W., Yang, L., Liu, C., Wang, B., Ruan, Q., 2013. Genetic variations of E6 and long control region of human papillomavirus type 16 from patients with cervical lesion in Liaoning, China. *BMC Cancer*. 13, 459.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution*. 28, 2731-9.

Tornesello, M.L., Losito, S., Benincasa, G., Fulciniti, F., Botti, G., Greggi, S., Buonaguro, L., Buonaguro, F.M., 2011. Human papillomavirus (HPV) genotypes and HPV16 variants and risk of adenocarcinoma and squamous cell carcinoma of the cervix. *Gynecol Oncol*. 121, 32 – 42.

- Ullman, C.G., Haris, P.I., Kell, B., Cason, J., Jewers, R.J., Best, J.M., Emery, V.C., Perkins, S.J., 1994. Hypothetical structure of the membrane-associated E5 oncoprotein of human papillomavirus type 16. *Biochem Soc Transact.* 22, 439S.
- van Gunsteren, W.F., Billeter, S.R., Eising, A.A., Hünenberger, P.H., Krüger, P., Mark, A.E., Scott, W.R.P., Tironi, I.G., 1996. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; vdf Hochschulverlag AG an der ETH Zürich and BIOMOS b.v.: Zürich, Groningen. 1-1042.
- VENUTI, A., Paolini, F., Nasir, L., Corteggio, A., Roperto, S., Campo, M.S., Borzacchiello, G., 2011. Papillomavirus E5: the smallest oncoprotein with many functions. *Molecular cancer.* 10, 1.
- Worth, C.L., Preissner, R., Blundell, T.L., 2011. SDM - a server for predicting effects of mutations on protein stability and malfunction. *Nucleic Acids Res.* 39, W215-22.
- Xu, D., Zhang, Y., 2011. Improving the Physical Realism and Structural Accuracy of Protein Models by a Two-step Atomic-level Energy Minimization. *Biophysical Journal.* 101, 2525-34.
- Xu, D., Zhang, Y., 2012. Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. *Proteins.* 80, 1715-1735.
- Yang, Z., 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24, 1586-91.
- Zhou, H., Zhou, Y., 2002. Distance-scaled, finite ideal-gas reference state improves structure-derived potentials of mean force for structure selection and stability prediction. *Protein Sci.* 11, 2714-26.

4 DISCUSSÃO GERAL

Muitos estudos têm se concentrado na análise genética da região LCR e na ORF dos genes E6 e E7 e na variação da sequência entre os tipos de HPV (CHAGAS et al., 2013; GURGEL et al., 2015). Outras investigações sobre a variabilidade genética do gene L1 e da ORF do gene E2 do HPV16, tem mostrado que essas regiões representam uma potencial fonte de variabilidade. No caso dos estudos conduzidos sobre a estrutura do gene L1, variações das sequências entre os tipos de HPV tem representado uma importante via de entendimento da estruturação do capsídeo viral, reconhecimento imune, neutralização viral e no desenvolvimento de estratégias vacinais (BERNARD et al., 2006; SICHERO; VILLA, 2006; SICHERO et al., 2007).

Em relação aos estudos das ORFs dos genes E6 e E7, variações nucleotídicas podem potencializar a ação de algumas variantes de HPV16, culminando, por exemplo, com a inativação das proteínas p53 e pRb, o que aumenta os riscos de desenvolvimento de uma neoplasia cervical (KAST et al., 1994; BURK et al., 2013). Com isso, alguns estudos epidemiológicos têm mostrado uma associação entre a progressão das lesões de alto grau que são precursoras dos tumores invasivos, com o acúmulo de variantes de HPV16 (LONDESBOROUGH et al., 1996; ZEHBE et al., 2001).

No cenário atual, existem poucos estudos que tratem das variações intratípicas encontradas na ORF do oncogene E5 do HPV16. Com isso, a identificação das variações genéticas neste gene, tornam-se limitadas e, consequentemente, o seu papel biológico no mecanismo de iniciação e de promoção da infecção.

No oncogene E5 do HPV16, a variação dos nucleotídeos, normalmente, encontram-se dentro do segmento mediano das regiões 5' e 3'. Os resultados de Eriksson et al. (1999), identificaram quatorze substituições de nucleotídeos no oncogene E5, sendo seis do tipo sinônimas e oito classificadas como não-sinônimas. As substituições observadas no gene E5 mantém, na maioria das vezes, a região de ocorrência da alteração genética, conservada. Dessa forma, muitas alterações sinônimas contribuem para a manutenção dos resíduos de aminoácidos que compõe a proteína. Um dos achados que subsidiam este dado, está na ausência de substituições entre os aminoácidos 11 e 24, que representam uma extensão dos

resíduos hidrofóbicos da região da hélice transmembrana da proteína, sendo, esta região, considerada conservada entre os diversos tipos de papilomavírus (MYERS et al., 1995).

Em outras regiões hidrofóbicas desta proteína, a substituição de aminoácidos também mantém a natureza conservada da sequência protéica entre os tipos de HPVs. No geral, as substituições encontradas no oncogene E5 são correlacionadas com um mecanismo de pressão de seleção, mesmo sabendo-se que a sua permanência ao longo do processo carcinogênico, não se dá nas etapas mais avançadas das lesões (VENUTI et al., 2011).

Esta variação na sequência da proteína pode afetar o potencial carcinogênico dos vírus, modificando a atividade transformante da célula, mediado pela oncoproteína E5, através da sua interação com os EGFRs (CONRAD et al., 1993; STRAIGHT et al., 1993).

Em relação ao HPV31, os estudos sobre a variabilidade genética, apontam para a linhagem C como a mais persistente em relação às linhagens A e B e, curiosamente, as linhagens A e B estão comumente associadas ao desenvolvimento das neoplasias intraepiteliais cervicais, grau 3 (NIC3) (SCHIFFMAN; WENTZENSEN, 2010). No entanto, a avaliação das variantes encontradas no oncogene E5 ainda não foram objeto de estudo.

No Estado de Pernambuco, vários esforços foram feitos para se avaliar a incidência dos tipos de HPVs, para o desenvolvimento de estratégias de elaboração de vacinas profiláticas. Neste estudo, foi observado uma alta incidência do HPV31, embora estes dados apresentem variação quanto à geografia (CHAGAS et al., 2011; CHAGAS et al., 2013).

Em relação aos estudos sobre variabilidade genética do HPV31, os genes mais bem estudados são o E6 e o E7 e estas alterações nucleotídicas estão fortemente associadas em regiões de epítomos de células T e células B (SILVA et al., 2009).

Outro estudo envolvendo a variabilidade genética do HPV31 circulantes no Estado de Pernambuco, foi conduzido para a região LCR e para o gene L1. Foram revelados substituições de nucleotídeos sinônimas e não-sinônimas, porém, ainda não foi desenvolvido a análise funcional destes polimorfismos (GURGEL et al., 2015).

Com essa variedade de sítios polimórficos indentificados nos HPVs 16 e 31, este trabalho se propôs a investigar as possíveis variações encontradas no oncogene E5 dos HPVs 16 e 31, fato que ainda não foi relatado na população brasileira, nem tampouco na região nordeste.

Este estudo mostrou a circulação de variantes do oncogene E5 nos HPVs 16 e 31. Essa observação se deu a partir da coleta de amostras cervicais (*cytobrush*), provenientes do HC-UFPE. Após a identificação das regiões polimórficas, procedemos com a análise *in silico* destas variantes. Propusemos o mapeamento de epítomos imunogênicos para cada variante identificada, além da análise estrutural da oncoproteína E5 (estrutura secundária e terciária). Os dados obtidos a partir das sequências gênicas foram processados afim de se obter a construção da árvore filogenética, bem como a análise da pressão de seleção destas variantes.

Em relação ao mapeamento dos epítomos imunogênicos, foram observadas várias regiões da proteína que apresentam interações com as moléculas do MHC-I e MHC-II, o que pode simbolizar uma forma de atuação do sistema imunológico afim de evitar uma evasão do vírus.

A predição computacional de regiões de ligação com moléculas do MHC classe I, vem sendo amplamente utilizada na identificação de epítomos imunogênicos e no desenvolvimento de estratégias vacinais (GROOT et al., 2001; MOISE et al., 2009; NIELSEN et al., 2010; LARSEN et al., 2010). Para isso, várias ferramentas e pesquisas em bioinformática tem permitido a predição de sítios de ligação de peptídeos às moléculas pertencentes à classe I e II do MHC, a exemplo do software *on line* IEDB.

A relevância destes achados, que envolve um estudo *in silico* da oncoproteína E5, está no aumento das informações acerca do comportamento do sistema imunológico dos seres humanos em relação à biologia da infecção viral, nos estágios iniciais ou mais avançados das lesões, que podem culminar com o desenvolvimento do câncer cervical, através da evasão viral do sistema imunológico.

Estas interações podem refletir o potencial de ação da oncoproteína E5 sobre o sistema imunológico, e que, paralelamente, contribui para um maior poder de atuação das oncoproteínas E6 e E7, principalmente, no que diz respeito, à proliferação celular e descontrole do ciclo, além da inibição dos mecanismos relacionados à apoptose (ZUR HAUSEN, 1996; TAN et al., 2012). Além disso, a oncoproteína E5 é expressa logo após a infecção, conferindo ao HPV um

mecanismo de escape imunológico e o início da carcinogênese (VENUTI et al., 2011).

A oncoproteína E5, nas células infectadas, tem demonstrado o potencial de alterar a apresentação dos antígenos celulares, reduzindo a expressão de moléculas do antígeno leucocitário humano (HLA), de Classe I, e também está relacionado em alterar a expressão de moléculas do MHC classe II na superfície celular (ASHRAFI et al., 2005; ZHANG et al., 2003).

A regulação do MHC classe I por E5, para impedir o reconhecimento das células infectadas pelos linfócitos T citotóxicos, ajuda a diminuir a sua expressão na superfície celular, deixando as células suscetíveis ao reconhecimento e ataque pelas células Natural Killers (NK). Estas células possuem receptores que reconhecem moléculas clássicas (inclusive HLA-C), e não clássicas (HLA-E) de HLA classe I. Quando essas moléculas não estão presentes e não são reconhecidas ocorre a lise celular mediada por NK. Por este motivo, a oncoproteína E5 consegue inibir a expressão na superfície celular de HLA-A e HLA-B, mas não altera a expressão de HLA-C e HLA-E, evitando a morte celular tanto pela via de linfócitos T citotóxicos quanto pela via das células NK (CAMPO et al., 2010).

As variantes identificadas neste estudo, tanto para o HPV16 quanto para o HPV31, apresentaram epítomos com moléculas pertencentes, principalmente, à HLA-A e HLA-B, o que pode simbolizar um potente alvo de estudo e desenvolvimento de agentes imunoterapêuticos.

Em relação à predição da estrutura protéica, temos que a disponibilidade de sequências do genoma completo de vários organismos aumentou grandiosamente, devido ao surgimento dos métodos de sequenciamento. Por exemplo, houve um aumento de aproximadamente 21% no número de nucleotídeos virais sequenciados e depositados no GenBank, e um aumento anual global de 43,6%, de depósitos de sequência de nucleotídeos de vírus, em todo o mundo, no ano de 2014 (BENSON et al., 2015).

Muitas das proteínas, geradas a partir dos genomas sequenciados, são anotadas como proteínas hipotéticas e que não tem uma função bem caracterizada. Então, uma melhor compreensão do genoma e das proteínas por ele sintetizadas, torna-se extremamente útil para o entendimento do possível papel biológico exercido por essas proteínas (IJAQ et al., 2015).

Neste estudo, as sequências que foram analisadas da proteína E5, foram derivadas da análise do sequenciamento do gene E5, provenientes de amostras biológicas, e posteriormente, analisadas no software PSIPRED, com o objetivo de investigar a estrutura secundária da proteína, bem como as variações encontradas nas sequências das variantes, em relação à sequência protótipo da oncoproteína E5 dos HPV16 e 31, depositadas no NCBI. O software PSIPRED disponibiliza uma acurácia dos resultados obtidos com a análise das sequências das variantes estudadas, permitindo gerar dados mais robustos sobre a estrutura e a topologia destas proteínas transmembranares.

Na análise filogenética das variantes de HPV16 deste estudo, foram evidenciados uma distribuição de isolados pertencentes às A (linhagem europeia) e D (linhagem não-européia), sendo esta última descrita com um maior potencial oncogênico nos estudos de Schiffman et al. (2010) e Smith et al. (2011). Em relação à árvore filogenética elaborada para o HPV31, as variantes foram agrupadas nas linhagens A, B e C. Esta diversidade de linhagens identificadas, indica uma ampla circulação das variantes dos HPVs 16 e 31 no Estado de Pernambuco.

Além dessa ampla circulação, é importante perceber que as análises *in silico* mostraram que as mutações I44L e P56A presentes na oncoproteína E5, nos HPVs 16 e 31, respectivamente, indicaram uma alta estabilidade da proteína, sendo necessário a idealização de ensaios funcionais para a compreensão deste achado.

Além disso, em relação à natureza química dos aminoácidos, mesmo com as substituições sinônimas e não-sinônimas presentes no oncogene E5 de ambos os HPVs estudados, em nada alterou o caráter hidrofóbico da proteína.

Para o HPV16, a análise comparativa das taxas de substituições sinônimas e não-sinônimas, mostraram que os modelos M2 e M3 representam o modelo de substituição de códons mais adequado, estando o oncogene 16E5 sob seleção negativa ou purificadora. Para o HPV31, a oncoproteína apresentou-se, de forma global, seleção negativa ou purificadora, cujo modelo M0 foi o mais adequado.

Dessa maneira, as várias abordagens computacionais, para a predição da estrutura e função das proteínas, variam de acordo com o método empregado, e possuem uma grande importância para a montagem de um banco de dados que ajudam a prever a estrutura hipotética de uma proteína. Os custos laboratoriais podem ser reduzidos, com consequente obtenção de resultados mais rápidos e

eficazes. No entanto, essa abordagem computacional deve ser avaliada via estudos funcionais, para efeito comprobatório dos resultados obtidos *in silico*.

5 CONCLUSÕES

O presente estudo identificou a circulação de variantes do oncogene E5 dos HPVs 16 e 31, e esta variabilidade genética está relacionada aos diferentes potenciais de infectividade de cada variação encontrada no genótipo do vírus.

As análises *in silico* foram fundamentais para a identificação de regiões relacionadas à formação de epítomos imunogênicos. As variações encontradas na sequência de aminoácidos da oncoproteína E5 dos HPVs 16 e 31, apresentaram regiões de epítomos com o MHC classe I e MHC classe II.

As moléculas do complexo HLA-A estão mais relacionadas com a ação da oncoproteína 16E5 e 31E5. No entanto, outras moléculas não-clássicas também podem ser alvos moleculares de E5, comprometendo o sistema imunológico do paciente.

As substituições de aminoácidos presentes na oncoproteína 16E5 e 31E5, conduziram para discretas modificações na estrutura secundária e terciária e, em alguns casos, levando à uma maior estabilidade da proteína. As alterações modificaram a estrutura da hélice, da cauda e das alças da proteína E5.

Embora as alterações estruturais observadas na oncoproteína 16E5 e 31E5 sejam discretas, estas mudanças podem refletir numa interação celular diferenciada entre a oncoproteína E5 e os seus alvos.

Com a análise filogenética das variantes do oncogene E5 dos HPVs 16 e 31, percebemos que na população brasileira já circulam as variantes de alto risco oncogênico. Neste estudo tivemos a variante D do oncogene 16E5 e as variantes A e B do oncogene 31E5, relacionados à elevada taxa de desenvolvimento do câncer cervical.

Baseado na análise do oncogene E5, foi observado que as variantes podem se encontrar sob diferentes meios de seleção.

Embora não se tenha o depósito do modelo cristalizado da oncoproteína 16E5 e 31E5 no banco de dados PDB, foi possível realizar a modelagem pelo método *ab initio* das variantes da oncoproteína E5 dos HPVs 16 e 31.

As mutações avaliadas podem apresentar diferentes efeitos sob a oncoproteína E5. Em geral, observamos que o impacto destas mutações tornou a oncoproteína ligeiramente estabilizante, altamente estabilizante ou ligeiramente desestabilizante.

A presença de variantes do oncogene E5 circulantes na população mostra a necessidade de realização de uma triagem molecular nas amostras das pacientes com diferentes resultados histopatológicos, visto vez que o oncogene E5 está relacionado com as etapas iniciais do ciclo de infecção, o que auxiliaria na compreensão do processo infeccioso.

O estudo dessas variantes implica na ampliação do conhecimento sobre o potencial de infecção das variantes dos HPVs 16 e 31, além de contribuir para melhorar nosso conhecimento sobre sua construção genética e seu potencial carcinogênico.

As mutações identificadas precisam ser avaliadas com o auxílio de estudos funcionais que levem a uma melhor interpretação do impacto dessas mutações no curso do processo carcinogênico, uma vez que os testes *in silico* demonstraram alterações conformacionais na estrutura protéica das variantes estudadas, bem como a formação de epitopos imunogênicos, que pode ser útil para o desenvolvimento de estratégias imunoterapêuticas.

Esta é a primeira análise sobre a variabilidade genética do oncogene 31E5 circulante na Região Nordeste do Brasil.

REFERÊNCIAS

- ANAYA-SAAVEDRA, Gabriela et al. HPV oral lesions in HIV-infected patients: the impact of long-term HAART. **Journal Of Oral Pathology & Medicine**, [S.l.], v. 42, n. 6, p.443-449, dez. 2013.
- ANGHEBEM-OLIVEIRA, Mauren Isfer; MERLIN, Júlio César. A proteína p16 é um novo marcador para progressão neoplásica do colo uterino? **RBAC**, [S.l.], vol. 42, n. 3, p. 181-185, 2010.
- ASHRAFI, G. H. et al. E5 protein of human papillomavirus type 16 selectively downregulates surface HLA class I. **International Journal of Cancer**. v. 113, n.2, p.276-283, Jan. 2005.
- BAXEVANIS, A. D.; OUELLETTE, B. F. **Bioinformatics: a practical guide to the analysis of genes and proteins**. [S.l.]: John Wiley & Sons, 2001.
- BELLEUDI, F. et al. HPV16 E5 affects the KGFR/FGFR2b-mediated epithelial growth through alteration of the receptor expression, signaling and endocytic traffic. **Oncogene**, [S.l.], v. 30, n. 50, p.4963-4976, maio 2011.
- BENSON, D. A. et al. GenBank. **Nucleic acids research**, v. 43, n. Database issue, p. D30-35, jan. 2015.
- BERGVALL, Monika; MELENDY, Thomas; ARCHAMBAULT, Jacques. The E1 proteins. **Virology**, [S.l.], v. 445, n. 1-2, p.35-56, out. 2013.
- BERMAN, H. M. et al. The Protein Data Bank. **Nucleic Acids Res.**, [S.l.], v. 28, n. 1, p. 235-242, jan. 2000.
- BERNARD, Hans-ulrich et al. Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. **Virology**, [S.l.], v. 401, n. 1, p.70-79, maio 2010.
- BERNARD, Hans-ulrich. The clinical importance of the nomenclature, evolution and taxonomy of human papillomaviruses. **Journal Of Clinical Virology**, [S.l.], v. 32, supl. 1, p.1-6, mar. 2005.
- BERUMEN, J. et al. Asian-American variants of human papillomavirus 16 and risk for cervical cancer: a case-control study. **J. Natl. Cancer Inst.**, [S.l.], v. 93, n. 17, p. 1325-1330, set. 2001.
- BONTKES, H. J. et al. HPV 16 infection and progression of cervical intra-epithelial neoplasia: analysis of HLA polymorphism and HPV 16 E6 sequence variants. **Int J Cancer**, [S.l.], v. 78, n. 2, p. 166-171, out. 1998.
- BOSCH, F. X. et al. The causal relation between human papillomavirus and cervical câncer. **J Clin Pathol.**, [S.l.], v. 55, n. 4, p. 244-265, abr. 2002.

BOUVARD, Véronique et al. A review of human carcinogens—Part B: biological agents. **The Lancet Oncology**, [S.l.], v. 10, n. 4, p.321-322, abr. 2009.

BRAVO, I. G.; ALONSO, A. Mucosal Human Papillomaviruses Encode Four Different E5 Proteins Whose Chemistry and Phylogeny Correlate with Malignant or Benign Growth. **Journal Of Virology**, [S.l.], v. 78, n. 24, p.13613-13626, nov. 2004.

BROERE, F. et al. **Principles of immunopharmacology**. [S.l.]: Springer, 2011.

BURD, E. M. Human papillomavirus and cervical cancer. **Clinical Microbiology Reviews**, [S.l.], v. 16, n. 1, p. 1-17, 2003.

BURK, Robert D.; HARARI, Ariana; CHEN, Zigui. Human papillomavirus genome variants. **Virology**, [S.l.], v. 445, n. 1-2, p.232-243, out. 2013.

CÂMARA, G. N. N. L. et al. Os papilomavírus humanos – HPV: Carcinogênese e imunogênese. **Universitas Ciências da Saúde**, [S.l.], v. 1, n. 1, p. 159-168, mar. 2008.

CAMPO, M. S. et al. HPV-16 E5 down-regulates expression of surface HLA class I and reduces recognition by CD8 T cells. **Virology**. v. 407, n.1, p. 137-142, nov. 2010.

CHAGAS, Bárbara S. et al. New variants of E6 and E7 oncogenes of human papillomavirus type 31 identified in Northeastern Brazil. **Gynecologic Oncology**, [S.l.], v. 123, n. 2, p.284-288, nov. 2011.

CHAGAS, Bárbara Simas et al. Novel E6 and E7 oncogenes variants of human papillomavirus type 31 in Brazilian women with abnormal cervical cytology. **Infection, Genetics And Evolution**, [S.l.], v. 16, n. 0, p.13-18, jun. 2013.

CHANG, J. L. et al. The expression of HPV-16 E5 protein in squamous neoplastic changes in the uterine cervix. *Journal Biomedical Science*, [S.l.], v. 8, n. 2, p. 206-213, mar.-abr. 2001.

CHEN, J. et al. Prediction of linear B-cell epitopes using amino acid pair antigenicity scale. **Amino Acids**, [S.l.], v. 33, n. 3, p.423-428, jan. 2007.

CHEN, Zigui et al. Evolution and Taxonomic Classification of Human Papillomavirus 16 (HPV16) - Related Variant Genomes: HPV31, HPV33, HPV35, HPV52, HPV58 and HPV67. **Plos One**, [S.l.], v. 6, n. 5, p. e20183, maio 2011.

CLIFFORD, G. M. et al. Comparison of HPV type distribution in high-grade cervical lesions and cervical cancer: a meta-analysis. **British Journal Of Cancer**, [S.l.], v. 89, n. 1, p.101-105, jul. 2003.

CONESA-ZAMORA, Pablo. Immune responses against virus and tumor in cervical carcinogenesis: Treatment strategies for avoiding the HPV-induced immune escape. **Gynecologic Oncology**, [S.l.], v. 131, n. 2, p.480-488, nov. 2013.

CONRAD, M.; BUBB, V.; SCHLEGEL, R. The human papillomavirus type 6 and 16 E5 proteins are membrane-associated proteins which associate with the 16-kilodalton pore-forming protein. **J. Virol.**, [S.I.], v. 67, n. 10, p. 6170-8, out. 1993.

CULLEN, A. P. et al. Analysis of the physical state of different human papillomavirus DNAs in intraepithelial and invasive cervical neoplasm. **J. Virol.**, [S.I.], v. 62, n. 2, p. 606-612, fev. 1991.

DA SILVA, M. B. et al. HPV31 and HPV33 incidence in cervical samples from women in Recife, Brazil. **Genetics and Molecular Research**, v. 8, n. 4, p. 1437-1443, 2009.

DARRIBA, Diego et al. JModelTest 2: more models, new heuristics and parallel computing. **Nature Methods**, [S.I.], v. 9, n. 8, p.772-772, ago. 2012.

DE GROOT, A. S. et al. From genome to vaccine: in silico predictions, ex vivo verification. **Vaccine**, v. 19, n. 31, p. 4385-4395, ago. 2001.

DEGLI-ESPOSTI, Mariapia A.; SMYTH, Mark J. Close encounters of different kinds: dendritic cells and NK cells take centre stage. **Nature Reviews Immunology**, [S.I.], v. 5, n. 2, p. 112, fev. 2005.

DIMAIO, Daniel; PETTI, Lisa M.. The E5 proteins. **Virology**, [S.I.], v. 445, n. 1-2, p.99-114, out. 2013.

DINIZ, M. O. et al. Immune Responses and Therapeutic Antitumor Effects of an Experimental DNA Vaccine Encoding Human Papillomavirus Type 16 Oncoproteins Genetically Fused to Herpesvirus Glycoprotein D. **Clinical And Vaccine Immunology**, [S.I.], v. 17, n. 10, p.1576-1583, ago. 2010.

DOORBAR, J. The E4 protein; structure, function and patterns of expression. **Virology**, v. 445, n. 1, p. 80-98, out. 2012.

DOORBAR, John. The E4 protein; structure, function and patterns of expression. **Virology**, [S.I.], v. 445, n. 1-2, p.80-98, out. 2013.

DORDEN, Shane; MAHADEVAN, Padmanabhan. Functional prediction of hypothetical proteins in human adenoviruses. **Bioinformation**, [S.I.], v. 11, n. 10, p.466-473, out. 2015.

DRUMMOND, Alexei J; RAMBAUT, Andrew. BEAST: Bayesian evolutionary analysis by sampling trees. **Bmc Evolutionary Biology**, [S.I.], v. 7, n. 1, p.214-221, nov. 2007.

ERIKSSON, A. et al. Human papillomavirus type 16 variant lineages characterized by nucleotide sequence analysis of the E5 coding segment and the E2 hinge region. **Journal of General Virology**. v. 80, p.595–600. 1999.

ETHERINGTON, I.J. et al. Histologic and Immunologic Associations of an HPV16 Variant in LoSIL Smears. **Gynecologic Oncology**, [S.I.], v. 72, n. 1, p.56-59, jan. 1999.

EWING, E. et al. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. **Genome Res.**, [S.I.], v. 8, n. 3, p. 175-185, 1998.

FEHRMANN, F.; KLUMPP, D. J.; LAIMINS, L. A. Human papillomavirus type 31 E5 protein supports cell cycle progression and activates late viral functions upon epithelial differentiation. **J. Virol.**, [S.I.], v. 75, n. 5, p. 2819-2831, mar. 2003.

FERLAY, Jacques et al. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. **International Journal Of Cancer**, [S.I.], v. 136, n. 5, p.359-386, out. 2015.

FIFE, Brian T.; BLUESTONE, Jeffrey A.. Control of peripheral T-cell tolerance and autoimmunity via the CTLA-4 and PD-1 pathways. **Immunological Reviews**, [S.I.], v. 224, n. 1, p.166-182, ago. 2008.

FREITAS, Antonio Carlos de et al. HrHPV E5 oncoprotein: immune evasion and related immunotherapies. **Journal Of Experimental & Clinical Cancer Research**, [S.I.], v. 36, n. 1, p.1-15, 25 maio 2017.

FREITAS, Antonio Carlos de et al. Susceptibility to cervical cancer: An overview. **Gynecologic Oncology**, [S.I.], v. 126, n. 2, p.304-311, ago. 2012.

GARCÍA-VALLVÉ, Santiago; ALONSO, Ángel; BRAVO, Ignacio G.. Papillomaviruses: different genes have different histories. **Trends In Microbiology**, [S.I.], v. 13, n. 11, p.514-521, nov. 2005.

GENTHER, S. M. et al. Quantitative role of the human papillomavirus type 16 E5 gene during the productive stage of the viral life cycle. **J. Virol.**, [S.I.], v. 75, n. 5, p. 2832-2842, mar. 2003.

GHITTONI, Raffaella et al. The biological properties of E6 and E7 oncoproteins from human papillomaviruses. **Virus Genes**, [S.I.], v. 40, n. 1, p.1-13, out. 2010.

GÓMORA, M. J.; HARO, I. Synthetic peptides for the immunodiagnosis of human diseases. **Curr. Med. Chem.**, [S.I.], v. 14, n. 5, p. 531-546, 2007.

GUINDON, Stéphane et al. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. **Systematic Biology**, [S.I.], v. 59, n. 3, p.307-321, mar. 2010.

GURGEL, A. P. A. D. **Análise de variabilidade genética do gene L1 e da região longa de controle (LCR) dos Papilomavírus Humano (HPVs) circulantes da Região Nordeste do Brasil**. 2015. 139 f. Tese (Doutorado) - Curso de Inovação Tecnológica, Universidade Federal do Pernambuco, Recife, 2015.

GUSFIELD, Dan. **Algorithms on strings, trees and sequences**: computer science and computational biology. [S.l.]: Cambridge university press, 1997.

HAI, Abdul et al. Differences in structural elements of Bcr-Abl oncoprotein isoforms in Chronic Myelogenous Leukemia. **Bioinformation**, [S.l.], v. 10, n. 3, p.108-114, mar. 2014.

HASAN, U. A. et al. TLR9 expression and function is abolished by the cervical cancer-associated human papillomavirus type 16. **J. Immunol.**, [S.l.], v. 178, n. 5, p. 3186-3197, mar. 2007.

HELLER, Corina et al. Identification of Key Amino Acid Residues That Determine the Ability of High Risk HPV16-E7 to Dysregulate Major Histocompatibility Complex Class I Expression. **Journal Of Biological Chemistry**, [S.l.], v. 286, n. 13, p.10983-10997, fev. 2011.

HUANG, Jian; HONDA, Wataru; KANEHISA, Minoru. Predicting B cell epitope residues with network topology based amino acid indices. **Genome Informatics**, [S.l.], v. 19, p. 40-49, 2007.

HUBERT, Pascale et al. E-cadherin-dependent adhesion of dendritic and Langerhans cells to keratinocytes is defective in cervical human papillomavirus-associated (pre) neoplastic lesions. **The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland**, [S.l.], v. 206, n. 3, p. 346-355, jul. 2005.

HUMMEL, Mary; LIM, Hock B.; LAIMINS, Laimonis A. Human papillomavirus type 31b late gene expression is regulated through protein kinase C-mediated changes in RNA processing. **Journal of virology**, [S.l.], v. 69, n. 6, p. 3381-3388, jun. 1995.

IJAQ, J. et al. Annotation and curation of uncharacterized proteins-challenges. **Frontiers in genetics**, v. 6, p. 119, mar. 2015.

INCA. Instituto Nacional do Câncer. **Estimativas 2018**: incidência de câncer no Brasil. Ministério da Saúde, 2018. Disponível em: www.inca.gov.br. Acesso em: 27 maio 2018.

ITO, Takashi et al. Roles for the two-hybrid system in exploration of the yeast protein interactome. **Molecular & Cellular Proteomics**, [S.l.], v. 1, n. 8, p. 561-566, ago. 2002.

JONES, D. T. GenTHREADER: an efficient and reliable protein fold recognition method for genomic sequences. **Journal of molecular biology**, v. 287, n. 4, p. 797-815, abr. 1999a.

JONES, D. T. Protein secondary structure prediction based on position-specific scoring matrices. **Journal of molecular biology**, v. 292, n. 2, p. 195-202, set. 1999b.

JOURA, Elmar A. et al. A 9-Valent HPV Vaccine against Infection and Intraepithelial Neoplasia in Women. **New England Journal Of Medicine**, [S.l.], v. 372, n. 8, p.711-723, fev. 2015.

KAST, W. M. et al. Role of HLA-A motifs in identification of potential CTL epitopes in human papillomavirus type 16 E6 and E7 proteins. **The Journal of Immunology**, v. 152, n. 8, p. 3904-3912, abr. 1994.

KAUFMAN, R. H.; ADAM, E.; VONKA, V. Human papillomavirus infection and cervical carcinoma. **Clin. Obstet. Gynecol.**, [S.l.], v. 43, n. 2, p. 363-80, jun. 2000.

KAVATI, A. E. **Interação de oncoproteínas virais E6 e E7 de HPV16/18 com alvos celulares potenciais para o desenvolvimento de estratégias terapêuticas**. 2012. 108 f. Dissertação (Mestrado) - Curso de Biotecnologia, Universidade de São Paulo, São Paulo, 2012.

KIM, D. E.; CHIVIAN, D.; BAKER, D. Protein structure prediction and analysis using the Robetta server. **Nucleic Acids Research**, [S.l.], v. 32, p.526-531, jul. 2004.

KRAWCZYK, E. et al. The Human Papillomavirus Type 16 E5 Oncoprotein Translocates Calpactin I to the Perinuclear Region. **Journal of Virology**, [S.l.], v. 85, n. 21, p.10968-10975, ago. 2011.

KREIDER, John W. et al. Morphological transformation in vivo of human uterine cervix with papillomavirus from condylomata acuminata. **Nature**, [S.l.], v. 317, n. 6038, p. 639-641, out. 1985.

KUMAR, Anoop et al. Identification of immunotherapeutic epitope of E5 protein of human papillomavirus-16: An in silico approach. **Biologicals**, [S.l.], v. 43, n. 5, p.344-348, set. 2015.

KUMAR, S.; HEDGES, S. B.. TimeTree2: species divergence times on the iPhone. **Bioinformatics**, [S.l.], v. 27, n. 14, p.2023-2024, maio 2011.

LARSEN, M. V. et al. Identification of CD8+ T cell epitopes in the West Nile virus polyprotein by reverse-immunology using NetCTL. **PLoS One**, v. 5, n. 9, p. e12697, set. 2010.

LASKOWSKI, R. A. et al. PROCHECK: a program to check the stereochemical quality of protein structures. **Journal Of Applied Crystallography**, [S.l.], v. 26, n. 2, p.283-291, abr. 1993.

LI, Yang et al. Progressive miRNA expression profiles in cervical carcinogenesis and identification of HPV-related target genes for miR-29. **The Journal Of Pathology**, [S.l.], v. 224, n. 4, p.484-495, 18 abr. 2011.

LIAO, Shujie et al. A novel “priming-boosting” strategy for immune interventions in cervical cancer. **Molecular Immunology**, [S.l.], v. 64, n. 2, p.295-305, abr. 2015.

LIU, D. W. et al. Cytotoxic T-Lymphocyte Responses to Human Papillomavirus Type 16 E5 and E7 Proteins and HLA-A*0201-Restricted T-Cell Peptides in Cervical Cancer Patients. **Journal Of Virology**, [S.l.], v. 81, n. 6, p.2869-2879, jan. 2007.

LIU, D. W. et al. Induction of CD8 T cells by vaccination with recombinant adenovirus expressing human papillomavirus type 16 E5 gene reduces tumor growth. **J. Virol.**, [S.l.], v. 74, n. 19, p. 9083-9089, out. 2000.

LIU, Zhida et al. A novel dendritic cell targeting HPV16 E7 synthetic vaccine in combination with PD-L1 blockade elicits therapeutic antitumor immunity in mice. **Oncoimmunology**, [S.l.], v. 5, n. 6, e1147641, mar. 2016.

LONDESBOROUGH, Philip et al. Human papillomavirus genotype as a predictor of persistence and development of high-grade lesions in women with minor cervical abnormalities. **International Journal of Cancer**, [S.l.], v. 69, n. 5, p.364-368, out. 1996.

LONGWORTH, M. S.; LAIMINS, L. A. Pathogenesis of Human Papillomaviruses in Differentiating Epithelia. **Microbiology And Molecular Biology Reviews**, [S.l.], v. 68, n. 2, p.362-372, jun. 2004.

MARIANI, Luciano; PAGLIUSI, Sonia. Vaccination and screening programs: harmonizing prevention strategies for HPV-related diseases. **Journal Of Experimental & Clinical Cancer Research**, [S.l.], v. 27, n. 1, p.84-89, 2008.

MAUFORT, J. P. et al. Human papillomavirus 16 E5 oncogene contributes to two stages of skin carcinogenesis. **Cancer research**, [S.l.], v. 67, n. 13, p. 6106-6112, 2007.

MAYRAND, M. H. et al. Detection of human papillomavirus type 16 DNA in consecutive genital samples does not always represent persistent infection as determined by molecular variant analysis. **Journal of clinical microbiology**, [S.l.], v. 38, n. 9, p. 3388-3393, 2000.

MCBRIDE, A. A. The papillomavirus E2 proteins. **Virology**, [S.l.], v. 445, n. 1, p. 57-79, 2013.

MCGUFFIN, L. J.; BRYSON, K.; JONES, D. T. The PSIPRED protein structure prediction server. **Bioinformatics Applications Note**, [S.l.], v. 16, n. 4, p. 404-405, abr. 2000.

MICHELS, K. B.; HAUSEN, H. Z. HPV vaccine for all. **The Lancet**, [S.l.], v. 374, n. 9686, p. 268-270, 2009.

MOISE, L. et al. In silico-accelerated identification of conserved and immunogenic variola/vaccinia T-cell epitopes. **Vaccine**, v. 27, n. 46, p. 6471-6479, out. 2009.

MONFRÉ, Elaine Rodrigues Mello. **Avaliação dos Níveis de Citocinas e HLA-G Solúvel em Linhagens Celulares Tumoriais de Colo Uterino Tratadas com Alcalóides de *Pterogyne nitens***. 2011. 176 f. Dissertação (Mestrado) - Curso de

Ciências Farmacêuticas, Faculdades de Ciências Farmacêuticas, Universidade Estadual Paulista "Júlio de Mesquita Filho", Araraquara, 2011.

MOYLE, Peter Michael; TOTH, Istvan. Modern Subunit Vaccines: Development, Components, and Research Opportunities. **Chemmedchem**, [S.l.], v. 8, n. 3, p.360-376, jan. 2013.

MUÑOZ, N. et al. HPV in the etiology of human cancer. **Vaccine**, [S.l.], v. 24, p. S1-S10, 2006.

MYERS, G. et al. Human Papillomaviruses: a compilation and analysis of nucleic acid and amino acid sequences. **Publication Los Alamos, N.M., U.S.A. : Theoretical Biology and Biophysics Group T-10**. Los Alamos National Laboratory, 1994.

NATALE, C. et al. Computer-assisted analysis of molecular mimicry between human papillomavirus 16 E7 oncoprotein and human protein sequences. **Immunology and Cell Biology**, [S.l.], v. 78, n. 6, p. 580-585, 2000.

NATH, R. et al. Analyses of variant human papillomavirus type-16 E5 proteins for their ability to induce mitogenesis of murine fibroblasts. **Cancer Cell International**, [S.l.], v. 6, n. 1, p. 19, 2006.

NIELSEN, M. et al. MHC class II epitope predictive algorithms. **Immunology**, v. 130, n. 3, p. 319-328, jul. 2010.

NINDL, I. et al. The E7 protein of human papillomavirus (HPV) type 16 expressed by recombinant vaccinia virus can be used for detection of antibodies in sera from cervical cancer patients. **Journal of Virological Methods**, [S.l.], v. 62, n. 1, p. 81-85, 1996.

PALADINI, Laura et al. Targeting microRNAs as key modulators of tumor immune response. **Journal Of Experimental & Clinical Cancer Research**, [S.l.], v. 35, n. 1, p.103-121, jun. 2016.

PARREM, P. W. et al. Antibodies in human infectious disease. **Immunol. Res.**, [S.l.], v. 21, n. 2-3, p. 265–278, 2000.

PATEL, P. et al. Incidence of types of cancer among HIV-infected persons compared with the general population in the United States, 1992-2003. **Ann. Intern. Med.**, [S.l.], v. 148, n. 10, p. 728-736, maio 2008.

PELLEQUER, Jean-luc; WESTHOF, Eric; VAN REGENMORTEL, Marc H.v.. Correlation between the location of antigenic sites and the prediction of turns in proteins. **Immunology Letters**, [S.l.], v. 36, n. 1, p.83-99, abr. 1993.

PERALTA-ZARAGOZA, Oscar; BERMUDEZ-MORALES; PEREZ-PLASENCIA. Targeted treatments for cervical cancer: a review. **Oncotargets And Therapy**, [S.l.], p.315-328, nov. 2012.

QUEIROZ, Alda Maria Alves; CANO, Maria Aparecida Tedeschi; ZAIA, José Eduardo. O papiloma vírus humano (HPV) em mulheres atendidas pelo SUS na cidade de Patos de Minas – MG. **Revista Brasileiras de Análises Clínicas**, [S.l.], v. 39, n. 2, p. 151-157, abr.-jun., 2007.

RAJEEVAN, Mangalathu S. et al. Epidemiologic and viral factors associated with cervical neoplasia in HPV-16-positive women. **International journal of cancer**, [S.l.], v. 115, n. 1, p. 114-120, maio. 2005.

RANNALA, Bruce; YANG, Ziheng. Phylogenetic Inference Using Whole Genomes. **Annual Review Of Genomics And Human Genetics**, [S.l.], v. 9, n. 1, p.217-231, set. 2008.

REBRIKOV, Denis V. et al. Complete genome sequence of a novel extrachromosomal virus-like element identified in planarian *Girardia tigrina*. **BMC genomics**, [S.l.], v. 3, n. 1, p. 15-23, 2002.

ROCHA, E. **Folhas de Bioinformática e Análise de Sequências**. [S.l.]: Instituto Pastier, 2000.

ROCHA, Nuno; NEEFJES, Jacques. MHC class II molecules on the move for successful antigen presentation. **The Embo Journal**, [S.l.], v. 27, n. 1, p.1-5, nov. 2008.

RONQUIST, Fredrik et al. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. **Systematic Biology**, [S.l.], v. 61, n. 3, p.539-542, fev. 2012.

ROSALLES, Ricardo. Immune therapy for human papillomaviruses-related cancers. **World Journal Of Clinical Oncology**, [S.l.], v. 5, n. 5, p.1002-1019, dez. 2014.

SAHAB, Z. et al. Quantitative Measurement of Human Papillomavirus Type 16 E5 Oncoprotein Levels in Epithelial Cell Lines by Mass Spectrometry. **Journal Of Virology**, [S.l.], v. 86, n. 17, p.9465-9473, jun. 2012.

SANJOSE, Silvia de et al. Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. **The Lancet Oncology**, [S.l.], v. 11, n. 11, p.1048-1056, nov. 2010.

SAPP, Martin; BIENKOWSKA-HABA, Malgorzata. Viral entry mechanisms: human papillomavirus and a long journey from extracellular matrix to the nucleus. **Febs Journal**, [S.l.], v. 276, n. 24, p.7206-7216, out. 2009.

SCHIFFMAN, M. et al. A Study of the Impact of Adding HPV Types to Cervical Cancer Screening and Triage Tests. **Jnci Journal Of The National Cancer Institute**, [S.l.], v. 97, n. 2, p.147-150, jan. 2005.

SCHIFFMAN, M.; WENTZENSEN, N. Human Papillomavirus Infection and the Multistage Carcinogenesis of Cervical Cancer. **Cancer Epidemiology Biomarkers & Prevention**, [S.l.], v. 22, n. 4, p.553-560, abr. 2013.

SCHIFFMAN, Mark et al. The carcinogenicity of human papillomavirus types reflects viral evolution. **Virology**, [S.l.], v. 337, n. 1, p.76-84, jun. 2005.

SCHIFFMAN, Mark; WENTZENSEN, Nicolas. From Human Papillomavirus to Cervical Cancer. **Obstetrics & Gynecology**, [S.l.], v. 116, n. 1, p.177-185, jul. 2010.

SCHNELL, Jason R.; CHOU, James J. Structure and mechanism of the M2 proton channel of influenza A virus. **Nature**, [S.l.], v. 451, n. 7178, p.591-595, jan. 2008.

SICHERO, L. et al. High grade cervical lesions are caused preferentially by non-European variants of HPVs 16 and 18. **International journal of cancer**, v. 120, n. 8, p. 1763-1768, abr. 2007.

SICHERO, L.; VILLA, L. Epidemiological and functional implications of molecular variants of human papillomavirus. **Brazilian journal of medical and biological research**, v. 39, n. 6, p. 707-717, jun. 2006.

SMITH, B. et al. Sequence imputation of HPV16 genomes for genetic association studies. **Plos one**, v. 6, n. 6, p. e21375, jun. 2011.

SONG, Yong Sang et al. Major Sequence Variants in E7 Gene of Human Papillomavirus Type 16 from Cervical Cancerous and Noncancerous Lesions of Korean Women. **Gynecologic Oncology**, [S.l.], v. 66, n. 2, p.275-281, ago. 1997.

SOUZA, Patrícia Savio de Araujo; SICHERO, Laura; MACIAG, Paulo Cesar. HPV variants and HLA polymorphisms: the role of variability on the risk of cervical cancer. **Future Oncology**, [S.l.], v. 5, n. 3, p.359-370, abr. 2009.

STADEN, Rodger. The staden sequence analysis package. **Molecular Biotechnology**, [S.l.], v. 5, n. 3, p.233-241, jun. 1996.

STEENBERGEN, Renske D. M. et al. HPV-mediated transformation of the anogenital tract. **Journal Of Clinical Virology**, [S.l.], v. 32, supl. 1, p.25-33, mar. 2005.

STRAIGHT, S. W. et al. The E5 oncoprotein of human papillomavirus type 16 transforms fibroblasts and effects the downregulation of the epidermal growth factor receptor in keratinocytes. **Journal of virology**, v. 67, n. 8, p. 4521-4532, ago. 1993.

STRAIGHT, Samuel W. et al. The E5 oncoprotein of human papillomavirus type 16 transforms fibroblasts and effects the downregulation of the epidermal growth factor receptor in keratinocytes. **Journal of virology**, [S.l.], v. 67, n. 8, p. 4521-4532, ago. 1993.

SUBRAMANIAN, Nirmala; CHINNAPPAN, Sudandiradoss. Prediction of promiscuous epitopes in the e6 protein of three high risk human papilloma viruses: a

computational approach. **Asian Pac. J. Cancer Prev**, [S.l.], v. 14, n. 7, p. 4167-4175, 2013.

SUMMERGILL, Kurt F. et al. P53 polymorphism, human papillomavirus infection in the oral cavity, and oral cancer. **Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, And Endodontology**, [S.l.], v. 90, n. 3, p.334-339, set. 2000.

SUPRYNOWICZ, Frank A. et al. The Human Papillomavirus Type 16 E5 Oncoprotein Inhibits Epidermal Growth Factor Trafficking Independently of Endosome Acidification. **Journal Of Virology**, [S.l.], v. 84, n. 20, p.10619-10629, ago. 2010.

TAMURA, Koichiro et al. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. **Molecular Biology And Evolution**, [S.l.], v. 30, n. 12, p.2725-2729, out. 2013.

TAN, S. et al. Anticancer drugs aimed at E6 and E7 activity in HPV-positive cervical cancer. **Current cancer drug targets**, v. 12, n. 2, p. 170-184, fev. 2012.

TINDLE, Robert W. Immune evasion in human papillomavirus-associated cervical cancer. **Nature Reviews Cancer**, [S.l.], v. 2, n. 1, p.59-64, jan. 2002.

VAN DOORSLAER, Koenraad. Evolution of the Papillomaviridae. **Virology**, [S.l.], v. 445, n. 1-2, p.11-20, out. 2013.

VENUTI, Aldo et al. Papillomavirus E5: the smallest oncoprotein with many functions. **Molecular Cancer**, [S.l.], v. 10, n. 1, p.140-157, 2011.

VIART, B. et al. EPI-peptide designer: a tool for designing peptide ligand libraries based on epitope–paratope interactions. **Bioinformatics**, [S.l.], v. 32, n. 10, p.1462-1470, jan. 2016.

VICI, Patrizia et al. Immunologic treatments for precancerous lesions and uterine cervical cancer. **Journal Of Experimental & Clinical Cancer Research**, [S.l.], v. 33, n. 1, p.29-43, 2014.

VILLA, L. L. Biología Molecular: Conceitos e Princípios Básicos. In: MARTINS, N.V.; PEREYRA, E.A. G. (Ed.) **Conhecendo o HPV Patologia do trato-genital inferior Colposcopia e CAF**. São Paulo: Frontis Editorial, 2000. p. 137-43.

VILLIERS, Ethel-Michele de et al. Classification of papillomaviruses. **Virology**, [S.l.], v. 324, n. 1, p. 17-27, 2004.

VITA, Randi et al. The Immune Epitope Database 2.0. **Nucleic Acids Research**, [S.l.], v. 38, suppl. 1, p. D854-D862, nov. 2010.

WARD, M. J. et al. Estimating the Rate of Intersubtype Recombination in Early HIV-1 Group M Strains. **Journal Of Virology**, [S.l.], v. 87, n. 4, p.1967-1973, dez. 2013.

WEE, Lawrence J. K. et al. SVM-based prediction of linear B-cell epitopes using Bayes Feature Extraction. **Bmc Genomics**, [S.l.], v. 11, n. 4, p.21-30, jan. 2010.

WENTZENSEN, Nicolas et al. Multiple human papillomavirus genotype infections in cervical cancer progression in the study to understand cervical cancer early endpoints and determinants. **International Journal Of Cancer**, [S.l.], v. 125, n. 9, p.2151-2158, nov. 2009.

WETHERILL, Laura F.; ROSS, Rebecca; MACDONALD, Andrew. HPV E5: An Enigmatic Oncoprotein. **Small DNA Tumour Viruses**, v. 55, 2012.

WHO. World Health Organization. Human papillomavirus (HPV) and cervical cancer. 2016. Disponível em: <http://www.who.int/immunization/topics/hpv/en/>. Acesso em: 10 de agosto de 2018.

WILSON, Patrick C.; ANDREWS, Sarah F. Tools to therapeutically harness the human antibody response. **Nature Reviews Immunology**, [S.l.], v. 12, n. 10, p.709-719, out. 2012.

WOODMAN, Ciaran B. J.; COLLINS, Stuart I.; YOUNG, Lawrence S.. The natural history of cervical HPV infection: unresolved issues. **Nature Reviews Cancer**, [S.l.], v. 7, n. 1, p.11-22, jan. 2007.

WORTH, C. L.; PREISSNER, R.; BLUNDELL, T. L. SDM - a server for predicting effects of mutations on protein stability and malfunction. **Nucleic Acids Res.** v.39, P. W215-222, jul. 2011.

XU, Dong; ZHANG, Yang. Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. **Proteins: Structure, Function, and Bioinformatics**, [S.l.], v. 80, n. 7, p.1715-1735, jul. 2012.

YANG, Z. PAML 4: Phylogenetic Analysis by Maximum Likelihood. **Molecular Biology And Evolution**, [S.l.], v. 24, n. 8, p.1586-1591, abr. 2007.

YANG, Z. The power of phylogenetic comparison in revealing protein function. **Proceedings Of The National Academy Of Sciences**, [S.l.], v. 102, n. 9, p.3179-3180, fev. 2005.

YU, Shui-lian et al. Antagonist-mediated down-regulation of toll-like receptors increases the prevalence of human papillomavirus infection in systemic lupus erythematosus. **Arthritis Research & Therapy**, [S.l.], v. 14, n. 2, p.80-82, 2012.

ZEHBE, I. et al. Human papillomavirus 16 E6 polymorphisms in cervical lesions from different European populations and their correlation with human leukocyte antigen class II haplotypes. **International journal of cancer**, v. 94, n. 5, p. 711-716, out. 2001.

ZEHBE, Ingeborg et al. Human papillomavirus 16 E6 polymorphisms in cervical lesions from different European populations and their correlation with human leukocyte antigen class II haplotypes. **International journal of cancer**, v. 94, n. 5, p. 711-716, 2001.

ZHANG, B. et al. The E5 protein of human papillomavirus type 16 perturbs MHC class II antigen maturation in human foreskin keratinocytes treated with interferon- γ . **Virology**, v. 310, n. 1, p. 100-108, Mai. 2003.

ZHANG, Shen-ying et al. Human Toll-like receptor-dependent induction of interferons in protective immunity to viruses. **Immunological Reviews**, [S.l.], v. 220, n. 1, p.225-236, dez. 2007.

ZUR HAUSEN, Harald. Papillomavirus infections — a major cause of human cancers. **Biochimica Et Biophysica Acta (bba) - Reviews On Cancer**, [S.l.], v. 1288, n. 2, p. F55-F78, out. 1996.

ZUR HAUSEN, Harald. Papillomaviruses and cancer: from basic studies to clinical application. **Nature Reviews Cancer**, [S.l.], v. 2, n. 5, p.342-350, maio 2002.

ANEXO A – CÓPIA DA APROVAÇÃO PELO COMITÊ DE ÉTICA INSTITUCIONAL

Plataforma Brasil - Ministério da Saúde

Universidade Federal de Pernambuco Centro de Ciências da Saúde / UFPE-CCS

PROJETO DE PESQUISA

Título: Caracterização do perfil molecular associado à infecção pelo Papilomavirus Humano: identificação de novos marcadores moleculares e alvos terapêuticos.

Área Temática:

Pesquisador: Antonio Carlos de Freitas

Versão: 2

Instituição: Universidade Federal de Pernambuco - UFPE

CAAE:

3606212.7.0000.5208

PARECER CONSUBSTANCIADO DO CEP

Número do Parecer: 67719

Data da Relatoria: 23/07/2012

ANEXO B – CÓPIA DO TCLE



REPÚBLICA FEDERATIVA DO BRASIL
UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE GENÉTICA

TERMO DE CONCENTIMENTO LIVRE E ESCLARECIDO

Convido a Sr. a para participar, como voluntária, da pesquisa "*Caracterização do perfil de expressão molecular associado à infecção pelo papilomavirus humano: identificação de novos marcadores moleculares e alvos terapêuticos*".

Após ser esclarecido (a) sobre as informações a seguir, no caso de aceitar fazer parte do estudo, assine ao final deste documento, que está em duas vias. Uma delas é sua e a outra é do pesquisador responsável. Em caso de recusa você não será penalizado (a) de forma alguma. Em caso de dúvida você pode procurar o Comitê de Ética em Pesquisa Envolvendo Seres Humanos da UFPE no endereço: (Avenida da Engenharia s/n – 1º Andar, Sla 4 - Cidade Universitária, Recife-PE, CEP: 50740-600, Tel.: 2126.8588 – e-mail: cepccs@ufpe.br).

INFORMAÇÕES SOBRE A PESQUISA:

Título do Projeto: CARACTERIZAÇÃO DO PERFIL DE EXPRESSÃO MOLECULAR ASSOCIADO À INFECÇÃO PELO PAPILOMAVIRUS HUMANO: IDENTIFICAÇÃO DE NOVOS MARCADORES MOLECULARES E ALVOS TERAPÊUTICOS.

Pesquisador Responsável: Pro. Dr. Antonio Carlos de Freitas

Endereço/Telefone/e-mail para contato (inclusive ligações a cobrar):

Av.Prof. Moraes Rego, s/n – Laboratório de Estudos Moleculares e Terapia Experimental (LEMTE), Departamento de Genética, Universidade Federal de Pernambuco, CEP 50670-901, Fone: 2126-8520, e-mail: acf_ufpe@yahoo.com.br

Pesquisadores participantes:

- Eliane Campos Coimbra
- Maria da Conceição Gomes Leitão

Telefones para contato: 2126-8520

OBJETIVOS: Avaliar o perfil de moléculas envolvidas no processo da carcinogênese cervical para a identificação de novos biomarcadores para diagnóstico, prognóstico e tratamento do câncer.

PROCEDIMENTOS DO ESTUDO: Se concordar em participar da pesquisa, você responderá a um questionário (perguntas para marcar um "x") sobre alguns hábitos de vida, que podem facilitar a infecção pelo HPV e o aparecimento de lesões no colo do útero. Nesse questionário não haverá o seu nome, e todas as informações que você der serão mantidas em sigilo (segredo), ou seja, só terão acesso a essas informações os pesquisadores envolvidos na pesquisa. Além disso, você permitirá a utilização, de uma amostra da biopsia e do sangue, que já serão coletados de você como parte da rotina de exames para a prevenção do câncer cervical. Nós utilizaremos uma parte das amostras de sangue e biopsia para desenvolver o nosso estudo que tem como objetivo, encontrar uma melhor forma de prevenção e tratamento do câncer cervical.

RISCOS E DESCONFORTOS: A paciente pode se sentir constrangida em responder ao questionário referente aos fatores de risco para desenvolvimento do câncer cervical. Os pesquisadores envolvidos na aplicação do questionário explicarão primeiramente do que se trata o projeto e a importância da colaboração da paciente para o estudo, deixando-a a vontade para decidir se deseja ou não fazer parte do estudo, de forma a minimizar esse desconforto. O risco na coleta de sangue ou biopsia é proveniente do próprio procedimento de rotina para prevenção do câncer cervical, não estando relacionado ao presente projeto de pesquisa. Na coleta de sangue a paciente poderá sentir um desconforto, como ardor durante a picada com a agulha e também pode aparecer um hematoma no local onde foi realizada a coleta (mancha azulada que desaparece em poucos dias. geralmente não há complicações.), mas caso ocorra algum desses sintomas o profissional responsável irá tratar de forma adequada. Em relação à coleta da biopsia a paciente pode sentir um desconforto no local em que será retirada a biopsia, mas caso ocorra o médico irá realizar o tratamento adequado.

BENEFÍCIOS: Será possível retribuir a colaboração voluntária das pacientes, com a detecção e identificação dos tipos de HPV de alto risco mais prevalentes na nossa região, presente nas amostras coletadas, auxiliando no diagnóstico preventivo do câncer cervical.

CUSTO/REEMBOLSO PARA O PARTICIPANTE: Não haverá nenhum gasto com sua participação. As consultas, exames, tratamentos serão totalmente gratuitos, não recebendo nenhuma cobrança com o que será realizado. Você também não receberá nenhum pagamento com a sua participação. Você pode retirar o consentimento a qualquer tempo, sem qualquer prejuízo da continuidade do acompanhamento/ tratamento usual.

CONFIDENCIALIDADE DA PESQUISA: Todos os dados referentes a presente pesquisa serão armazenados no servidor da plataforma de genômica e expressão gênica do Centro

de Ciências Biológicas da Universidade Federal de Pernambuco, sob responsabilidade do Prof. Antonio Carlos de Freitas, com acesso restrito aos pesquisadores, durante o período de vigência do projeto.

Assinatura do Pesquisador Responsável:

CONSENTIMENTO DA PARTICIPAÇÃO DA PESSOA COMO SUJEITO

Eu, _____, RG/CPF/_____, abaixo assinado, _____ concordo em participar do estudo, como sujeito. Fui devidamente informado(a) e esclarecido(a) pelo(a) pesquisador(a) _____ sobre a pesquisa, os procedimentos nela envolvidos, assim como os possíveis riscos e benefícios decorrentes de minha participação. Foi-me garantido que posso retirar meu consentimento a qualquer momento, sem que isto leve a qualquer penalidade ou interrupção de meu acompanhamento/ assistência/tratamento.

Local e data _____

Nome e Assinatura do sujeito ou responsável:

Presenciamos a solicitação de consentimento, esclarecimentos sobre a pesquisa e aceite do sujeito em participar.

02 testemunhas (não ligadas à equipe de pesquisadores):

Nome: _____

Assinatura: _____

Nome: _____

Assinatura: _____