

Research Article

Prevalence of Human Papillomavirus Variants and Genetic Diversity in the L1 Gene and Long Control Region of HPV16, HPV31, and HPV58 Found in North-East Brazil

Ana Pavla Almeida Diniz Gurgel,¹ Bárbara Simas Chagas,¹
Carolina Medeiros do Amaral,¹ Kamylla Conceição Gomes Nascimento,¹
Lígia Rosa Sales Leal,¹ Jacinto da Costa Silva Neto,²
Maria Tereza Cartaxo Muniz,³ and Antonio Carlos de Freitas¹

¹Laboratory of Molecular Studies and Experimental Therapy (LEMTE), Department of Genetics, Federal University of Pernambuco, Avenida Professor Moraes Rego S/N, 50670-901 Recife, PE, Brazil

²Department of Histology and Embryology, Federal University of Pernambuco, Avenida Professor Moraes Rego S/N, 50670-901 Recife, PE, Brazil

³Molecular Biology Laboratory, Pediatric Oncohematology Center, University of Pernambuco, Avenida Agamenon Magalhaes S/N, 50100-130 Recife, PE, Brazil

Correspondence should be addressed to Antonio Carlos de Freitas; acf_ufpe@yahoo.com.br

Received 14 November 2014; Accepted 8 January 2015

Academic Editor: Paul K. S. Chan

Copyright © 2015 Ana Pavla Almeida Diniz Gurgel et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study showed the prevalence of human papillomavirus (HPV) variants as well as nucleotide changes within L1 gene and LCR of the HPV16, HPV31, and HPV58 found in cervical lesions of women from North-East Brazil.

1. Introduction

Cervical cancer is the second most significant cause of cancer in women worldwide, with more than 529,000 new cases diagnosed and 275,000 deaths in 2011 [1]. Among these, 85% of the total number of cervical cancer cases occur in developing countries [1]. In Brazil, cervical cancer is the third most common cancer among women [2].

It is well-established that persistent infections caused by human papillomavirus (HPV) is a key aetiological factor in the development of cervical lesions and cervical cancer [3]. To date, 184 HPV types have been identified (<http://www.hpvcenter.se/html/refclones.html>) and 62 belong to the *Alphapapillomavirus* genus. Among these, epidemiological data showed that HPV16 and HPV18 are responsible for 70% of cases of invasive cervical cancer worldwide [4]. Moreover, other *Alphapapillomavirus* genotypes, such as HPV31, HPV33,

HPV35, HPV45, HPV52, and HPV58, are involved in 18% of cases of squamous cell carcinoma cancer worldwide [4].

Previous studies revealed that different variants of HPV16 coevolved with the three main human phylogenetic branches: African, Caucasian, and Asian [5, 6]. Hence, variants of HPV16 were grouped into five distinct categories spread across different geographical regions: Europe (E), Asia (As), Asian-America (AA), Africa 1 (AF-1), and Africa 2 (Af-2) [5, 6]. However, recent studies have redefined variant as a nucleotide sequence that differs by approximately 1% between two or more variants of the same HPV type [7–9]. In addition, sublineage also was redefined as a nucleotide sequence that differs from 0.5 to 0.9% within a full genome of the same HPV type [7–9]. Hence, according to this analysis the HPV16 has four variant lineages (A, B, C, and D) and nine sublineages; HPV31 showed three viral lineages (A, B, and C) and seven

TABLE 1: Primers for HPV16, HPV31, and HPV58 L1 gene and LCR analysis.

HPV type	Region amplified	Sequence primer 5'-3'	Amplicon size
HPV16	L1	5'-ACGGTACCCAGGTGACTTTTATTTACATCC-3' 5'-TAAGTCGACCAGCTTACGTTTTTTTGC-3'	1500 pb
	LCR	5'-TTCTGCAGACCTAGATCAGTTTC-3' 5'-GTGCATAACTGTGGTAACTTTCTGG-3'	1057 pb
HPV31	L1	5'-GATTTCGAAAAATGCTCTCTGTGGC-3' 5'-TAAGTCGACCTTTTTAGTTTTTTTTACG-3'	1550 pb
	LCR	5'-TACCTCCAAAGGAAAAGGAAGACCC-3' 5'-TTGGCACAAATCATGCAATGTTTCG-3'	1152 pb
HPV58	L1	5'-AACCTGGTCCAGACATTGCATC-3' 5'-CCACCAAACGCAAAAAGGTTA-3'	1574 pb
	LCR	5'-CATGTTCTATGTCCTTGTCAG-3' 5'-TTGCCAGGTGTGGACTAA-3'	1000 pb

sublineages; and HPV58 has four variant lineages (A, B, C, and D) and seven sublineages [7].

Several studies have demonstrated that HPV16, HPV31, and HPV58 variants are associated with oncogenicity, persistence, and the progression of infection [6, 10–26]. Although there have been important studies about genomic diversity, there are few records of the variants of HPV16, HPV31, and HPV58 that are widespread in Brazil [21, 22, 27–30] and more specifically in North-East Brazil [31–33]. Previous studies have shown that HPV16, HPV31, HPV33, and HPV58 are the most common HPV types found in cervical samples from North-East Brazil [33–36]. For this reason, there is a need for studies concerning the genomic characterization of circulating HPV16, HPV31, and HPV58 variants due to their biological differences, which could explain at least part of the differences in infectivity and pathogenicity of some HPV variants. For instance, nucleotide changes in the L1 gene may affect its immune response to HPV16, HPV31, and HPV58. In addition, the polymorphic sites in long control region (LCR) may affect the transcriptional activity of E6 promoters [37].

Thus, the aim of this study was to detect nucleotide changes within L1 and LCR of HPV16, HPV31, and HPV58 in cervical samples obtained from North-East Brazil. In silico prediction of B-cell and T-cell epitopes in the L1 gene of HPV16, HPV31, and HPV58 was performed. Moreover, binding sites of transcriptional factors prediction were also performed in LCR of HPV16, HPV31, and HPV58. Finally, a phylogenetic analysis was conducted to determine which variants of HPV16, HPV31, and HPV58 are found in North-East Brazil.

2. Materials and Methods

2.1. Study Population and Ethics Statement. A total of 206 samples were randomly collected from women during their medical consultation at two locations: the Gynecological Unit of the Integrated Medicine Center, in Sergipe State, and Salgadinho Medical Care Center, in Maceió, Alagoas

State, North-East Brazil, between November 2010 and July 2011. The study included women with low-grade intraepithelial lesions (LSIL) and high-grade intraepithelial lesions (HSIL). This study was approved by the Ethics Committee of the University of Alagoas (UFAL 004650/2010-55) and the Ethics Committee of the Federal University of Sergipe (CEP/CCS/UFPE N° 491/11).

2.2. Nucleic Acid Isolation and Detection of HPV Types 16, 31, and 58. The cervical cells were collected by using cytobrush and placed in polyethylene tubes containing phosphate-buffered saline and transferred to the Molecular Studies and Experimental Therapy Laboratory (LEMTE) and stored at -80°C until analysis. Nucleic acids were extracted by means of the DNeasy Blood and Tissue Kit 135 (Qiagen), in accordance with the manufacturer's instructions. A polymerase chain reaction (PCR) was performed with the *MDM2* gene to avoid the false negative and to assess the quality of the extracted DNA. Positive HPV16, HPV31, and HPV58 DNA were detected by using PCR with degenerate primers MY09/11 followed by direct sequencing. The positive HPV DNA was purified with the Invisorb Fragment Cleanup (Invitex) Kit and sequenced (in duplicate) by using BigDyeTM Terminator Cycle Sequencing Read Reaction Kit (Applied Biosystems) and ABI PRISM (Applied Biosystems) to obtain both the forward and reverse sequences.

2.3. Analysis of L1 Gene and LCR of HPV16, HPV31, and HPV58 by PCR and Sequencing. HPV16 (L1 $n = 14$, LCR $n = 22$), HPV31 (L1 $n = 4$, LCR $n = 8$), and HPV58 (L1 $n = 8$, LCR $n = 5$) found in cervical samples were further characterized by amplification of partial sequence of L1 and LCR by means of the specific primer pairs described in Table 1. The reactions were performed with a final volume of 25 μL containing 50 ng of DNA, 20 pmol of each primer, and 1X PCR Master Mix (Promega). The PCR cycling conditions were as follows: initial denaturation at 95°C for 5 minutes, 35 cycles of denaturation at 95°C for 30 seconds, annealing

at 56°C for 1 minute, elongation at 72°C for 2 minutes, and a final extension at 72°C for 10 minutes. The PCR products were run on the agarose gel (1%). Following this, the amplicons were purified with the Invisorb Fragment Cleanup (Invitek) Kit and nucleotide sequences were obtained by means of fluorescent BigDye™ Terminator Cycle Sequencing using v 3.1 Ready Reaction ABI PRISM (Applied Biosystems) to obtain both the forward and reverse sequences. PCR and sequencing were performed in duplicate.

2.4. Data Analysis. The obtained sequences were assembled by means of the Staden package [38]. They were then evaluated to determine the nucleotide divergence relative to the nucleotide sequences of HPV16 (K02718), HPV31 (J04353), and HPV58 (D90400). Sequence comparisons were carried out using the Basic Local Alignment Search Tool (BLAST) and multiple alignments were performed by using the CLUSTALW (Mega 5.2, Beta version) program [39].

The Neighbor-Joining algorithm and the Kimura 2-Parameter model trees, with 1000 bootstrapped replicates, were built by using the MEGA package, version 5.2 [39]. Phylogenetic analyses were performed with LCR sequences of HPV16, HPV31, and HPV58. The partial sequence of L1 and LCR genes of the HPV16, HPV31, and HPV58 was deposited in the NCBI GenBank database, under the following accession numbers: HPV16 L1 gene: KJ467225-467238; HPV16 LCR: KJ452220-452242; HPV31 L1 gene: KJ452216-452219; HPV31 LCR: KJ435060-435067; HPV58 L1 gene: KJ467239-477246; HPV58 LCR: KJ567247-467252. The references for the viral sequences used to construct the phylogenetic branches were collected from the GenBank sequence database and are listed in Table 2.

2.5. B-Cell and T-Cell Epitope Prediction. The putative impact of the HPV variants was estimated in silico by predicting the B-cell and T-cell epitopes. In this study, it was assumed that changes in the amino acid sequences of L1 proteins within the B-cell epitope regions could affect the binding affinities of the neutralizing antibodies and in the case of the T-cell did not initiate an epitope-specific immune response. Thus, the B-cell epitope of prototype sequences was predicted by means of the BcePred server, which is available from URL: <http://www.imtech.res.in/raghava/bcepred/>. The prediction was carried out with the aid of physicochemical parameters, such as hydrophilicity, flexibility/mobility, accessibility, polarity, exposed surface, turns and antigenic propensity [40].

The T-cell epitope predictions were performed by using ProPred and ProPred I servers. The ProPred I server (available from URL: <http://www.imtech.res.in/raghava/propred1/>) was used to predict MHC Class-I binding regions [41], while the ProPred server, (available from URL: <http://www.imtech.res.in/raghava/propred/>) was used to predict MHC Class-II binding peptide [42].

2.6. Transcription Factor in the Binding Sites Prediction. The PROMO server (available from URL: http://algen.lsi.upc.es/cgi-bin/promo_v3/promo/promoinit.cgi?dirDB=TF_8.3) was

TABLE 2: Reference sequences used in phylogenetic analysis.

Reference sequences (GeneBank accession number)	HPV type/lineage	References
K02718	HPV16/A	(Burk et al., 2013) [7]
AF536179	HPV16/A	
HQ644236	HPV16/A	
AF534061	HPV16/A	
AF536180	HPV16/B	
HQ644298	HPV16/B	
AF472509	HPV16/C	
HQ644257	HPV16/D	
AY686579	HPV16/D	
AF402678	HPV16/D	
J04353	HPV31/A	(Burk et al., 2013) [7]
HQ537675	HPV31/A	
HQ537676	HPV31/B	
HQ537680	HPV31/B	
HQ537682	HPV31/C	
HQ537684	HPV31/C	
HQ537685	HPV31/C	
D90400	HPV58/A	(Burk et al., 2013) [7]
HQ537752	HPV58/A	
HQ537758	HPV58/A	
HQ537762	HPV58/B	
HQ537764	HPV58/B	
HQ537774	HPV58/C	
HQ537768	HPV58/D	
HQ537770	HPV58/D	

used to search within the LCR of the HPV16, HPV31, and HPV58 for potential binding sites for cellular and viral transcriptional factors. The transcriptional factors analyzed were the following: AP-1, E2, NF-1, Oct-1, YY1, C/EBP, and Sp1. Transcriptional factors were predicted within a dissimilarity margin of less than or equal to 15% [43, 44].

3. Results

3.1. Characteristics of the Population. A total of 206 cervical smear tests were carried out to detect HPV DNA and the results showed that 121 (59%) were positive for HPV. Among these, 94/121 (77.7%) were infected with one HPV type and 27/121 (22.3%) were infected with more than one HPV type (Figure 1). As a single infection, 40/94 (42.6%) samples were positive for HPV16, 3/94 (3.1%) samples were HPV18, 9/94 (9.6%) samples were positive for HPV31, 5/94 (5.3%) samples were HPV33, and 24/94 (25.5%) samples were positive for HPV58. A total of 13/94 (13.9%) of the samples were infected by other HPV types (Figure 1). With regard to coinfection, 19/27 (70.4%) samples were positive for HPV16/31, 1/27 (3.7%) sample was positive for HPV16/33, 3/27 (11.1%) samples were HPV31/33, and 4/27 samples (14.8%) were infected with HPV31/58 genotypes (Figure 1). Positive samples of HPV16, HPV31, and HPV58 were separated from the total, for molecular characterization.

TABLE 3: HPV16 L1 sequence variations at the genome and protein level and the putative biological function affected.

Genome position	Variation	AA changed	Protein position	Number of isolates	Biological function
5862	C/T	H102Y	β -C	5/14	MHC Class-I/MHC Class-II
5909	T/C	—		5/14	
5981	G/A	—		1/14	
6023	A/T	—		1/14	
6163	C/A	T202N	EF-loop	5/14	B-cell/MHC Class-I
6176	A/C	N207T	EF-loop	2/14	B-cell/MHC Class-I/MHC Class-II
6240	C/G	H228D	EF-loop	14/14	B-cell/MHC Class-I
6245	T/C	—		5/14	
6300	A/C	—		1/14	
6315	A/G	—		4/14	
6436	A/G	T292A	FG-loop	5/14	MHC Class-I/MHC Class-II
6484	T/C	—		1/14	
6561	C/T	S308P		5/14	
6570	T/G	—		1/14	
6697	A/C	T379P	HI-loop	5/14	B-cell/MHC Class-I
6723	G/A	—		5/14	
6805	A/T	T415S	h2	1/14	MHC Class-I/MHC Class-II/VLP formation
6856	C/T	—		5/14	
6867	C/T	—		5/14	
6903	Ins ATC	Ins T and S 448	h4	14/14	MHC Class-I/
6955	Del GAT	Del D 465	h4- β -J		MHC Class-II/VLP formation
6972	C/T	—		5/14	
6996	G/A	—		5/14	

(—) Synonymous variations with regard to the prototype reference K02718.1. Del: deletion, Ins: insertion, H: histidine, Y: tyrosine, T: threonine, N: asparagine, D: aspartic acid, A: alanine, S: serine, P: proline, and Q: glutamine.

3.2. *HPV16*. 1300 pb nucleotide sequences of HPV16 L1 gene ($n = 14$) were compared with the reference sequence (K02718.1). DNA sequence analysis revealed twenty-three single nucleotide changes in the L1 gene, in which 10/23 (43.5%) showed nonsynonymous variations. The C6240G variation, which leads to H228D in the EF loop of the L1 protein, was observed in 100% of the samples and is embedded in B-cell and T-cell epitopes. Similarly, the insertion of ATC as well as the deletion of GAT, which leads to 447 threonine/448 serine and 445 aspartate amino acid changes respectively, was observed in all of the samples. These amino acid changes are located in the H4 and B-J regions and are embedded in T-cell epitopes. The nonsynonymous variations C5862T (H102Y), C6163A (T202N), A6176C (N207T), A6436G (T292A), and A6697C (T379P) are located in the EF, FG, and HI external loops and are embedded in B-cell and/or T-cell epitopes. Compared with the prototype of the HPV sequences, there was no evidence of premature stop codons in the HPV16 L1 gene variants. The detected variations are summarized in Table 3.

With regard to the HPV16 LCR sequences, thirty nucleotide changes were observed, in which 24/30 (80%) are embedded in the binding sites of transcriptional factors. The most common variations, insC7434, C7436G and delA7869, were found in 100% of the samples and are embedded in

the YY1, NF-1, and E2 binding sites of the transcriptional factors, respectively. Moreover, the G7521A variation was found in 10/23 (43.5%) of the total number of samples, followed by A7489C (39.1%), G7493A (39.1%), C7693A (39.1%), C7768T (39.1%), and C7790T (39.1%). Among these variations, C7693A and C7790T are embedded in the E2 and YY1 binding sites, respectively. Some of these nucleotide changes are “diagnostic SNPs” that are conducted to detect the lineages and sublineages of HPV16 [15]. The detected variations are summarized in Table 4.

The phylogenetic analyses showed that 63.63% (14/22) samples belong to the A variant and 36.36% (8/22) samples belong to the D variant (Figure 2). No variants in the B or C variant were found in this study (Figure 2).

3.3. *HPV31*. The HPV31 L1 gene was analyzed through an alignment of 809 pb nucleotide sequences. The DNA sequence study revealed nine nucleotide changes in the L1 gene, two of which (2/9) were nonsynonymous variations and 7/9 were synonymous variations. When compared with the L1 protein of HPV16, the variations of A6025G (T267A) are located in the FG loop and embedded in the T-cell epitopes. Moreover, the C6379A nucleotide (T274N) is embedded in the FG loop as well as within the B-cell and T-cell epitopes. Compared with the prototype HPV sequences (J04353.1),

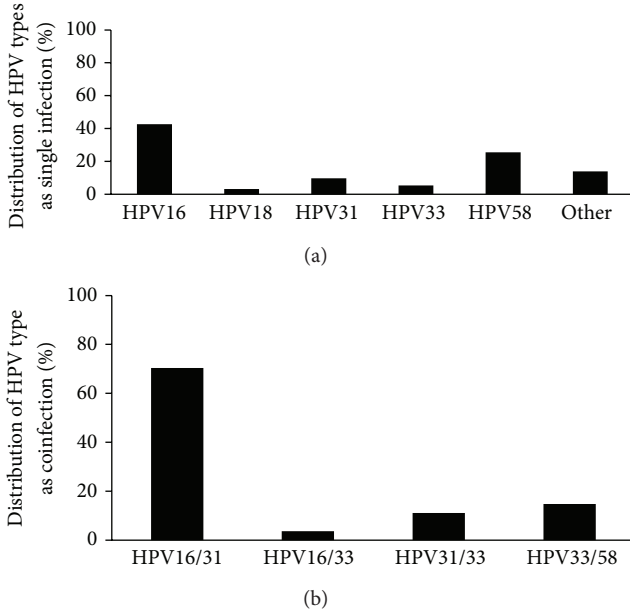


FIGURE 1: HPV type prevalence in North-East Brazil. (a) Distribution of HPV type as a single infection found in cervical lesions of women from North-East Brazil. (b) Distribution of HPV type as a coinfection found in cervical lesions of women from North-East Brazil.

there was no evidence of premature stop codons in the HPV31 L1 gene of variants. The detected variations are summarized in Table 5.

With regard to the HPV31 LCR nucleotide sequence, fragments of 883 pb were analyzed. Among these, thirty nucleotide changes were observed, 14/30 (47%) of which are embedded in the binding sites of the transcriptional factors. The most common variations were a deletion of TGTTCCT-GCT at positions 7341-7450 (8/8, 100%) and located within the transcriptional binding sites of NF-1. C7480T and T7871G were found in 20% of the samples and are located within the binding sites of the E2 transcriptional factor. The detected variations are summarized in Table 6.

The phylogenetic analysis showed 62.5% of variants are clustered into A branches ($n = 5$) and 37.5% are clustered into C branches ($n = 3$) (Figure 3). However, there were no observed variants clustered into B branches.

3.4. HPV58. The HPV58 L1 gene was analyzed through an alignment of the 1264 pb nucleotide sequence. Altogether, thirty-five single nucleotide polymorphisms were found, seven of which (7/35) were nonsynonymous variations. The most common are nucleotide changes were A6540G (I335M), C6828A (N422D), A6014C (L150F), G5994A (V144I), A6799G (I412V), G6823A (D420N), and C6689A (T375N); these are either located in the external loop (DE/HI loop) or alpha helix (H2 and H3) regions. Moreover, these polymorphisms are embedded in B-cell and/or T-cell epitopes. Compared with the prototype HPV sequence (D90400.1), insertion and deletion events were not identified and there was no evidence of premature stop codons or nucleotide

TABLE 4: HPV16 LCR sequence variations with regard to the reference sequence K02718.1. The table also shows the putative transcriptional factor binding sites in the nucleotide altered.

Genome position	Nucleotide variation	Number of isolates	Transcriptional factor
7175	A/C	1/23	C/EBPbeta
7194	G/T	10/23	
7229	A/C	1/23	C/EBPbeta
7234	A/C	10/23	
7258	G/A	1/23	YY1
7281	G/T	1/23	
7316	A/C	1/23	YY1, AP-1
7340	A/T	7/23	C/EBPbeta
7360	A/G	1/23	
7361	G/C	1/23	
7395	C/T	7/23	C/EBPbeta
7434	Ins C	23/23	NF-1, YY1
7436	C/G	23/23	YY1
7441	T/G	1/23	YY1
7485	A/C	9/23	NF1
7490	G/A	9/23	
7521	G/A	10/23	YY1
7670	C/T	9/23	
7690	C/A	9/23	E2
7730	A/C	8/23	Oct-1
7730	A/T	1/23	Oct-1
7747	T/G	7/23	C/EBPbeta
7768	C/T	9/23	NF-1
7790	C/T	1/23	YY1
7790	C/T	9/23	YY1
7838	G/T	2/23	C/EBPbeta
7869	Del A	23/23	E2
7874	G/A	1/23	NF-1
7881	C/A	1/23	C/EBPbeta
7891	C/G	7/23	C/EBPbeta

deletions in the L1 HPV58 sequences analyzed. The detected variations are described in Table 7.

With regard to the HPV58 LCR sequences, thirty-five nucleotide changes were observed, in which 12/35 are embedded within the binding sites of transcriptional factors. The most common variations C7745A and A7794G were found in 50% of the samples and are embedded within the NF-1 and E2 binding sites of transcriptional factors, respectively. The detected variations are summarized in Table 8.

In addition, the phylogenetic analyses showed 50% of isolates belong to the A variant, followed by B (16.6%), C (16.6%), and D (16.6%) variants (Figure 4).

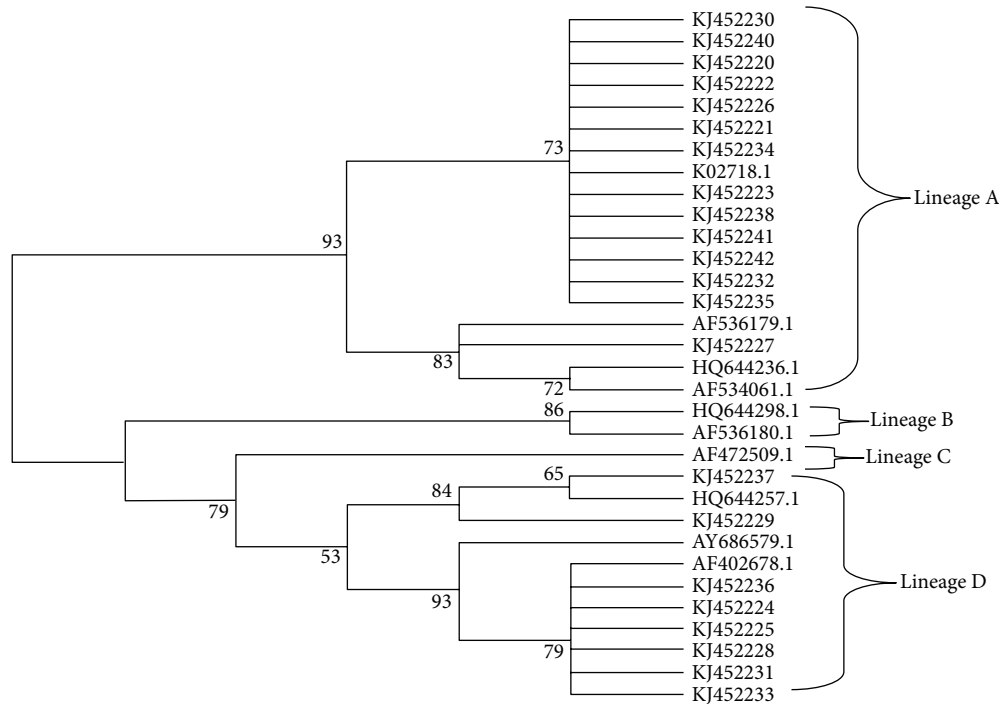


FIGURE 2: Neighbor joining phylogenetic tree of HPV16 variants based on 759 pb of LCR. Four clusters were identified as lineages A, B, C, and D. Reference sequences are listed in Table 2. Viral lineages analyzed in this study (KJ452220-452242) are clustered into A and D branches. Only bootstrap values above 50% are represented in the branches.

TABLE 5: HPV31 L1 sequence variations at the genome and protein level and the putative biological function affected.

Genome Position	Variation	AA changed	Protein position (relative to HPV16 L1 gene)	Number of isolates	Biological function
5926	A/G	—		1/4	
5927	T/C	—		1/4	
6025	A/G	T267A	FG loop	1/4	MHC Class-I/HC Class-II
6091	C/T	—		1/4	
6335	G/A	—		1/4	
6357	A/G	—		1/4	
6374	C/T	—		1/4	
6379	C/A	T274N	FG loop	1/4	B-cell/MHC Class-I
6593	T/G	—		1/4	

(—) Synonymous variations with regard to the reference sequence J04353.1. T: threonine; N: asparagine; A: alanine.

4. Discussion

Several studies have demonstrated that variants of HPV16, HPV31, and HPV58 may affect the oncogenicity, persistence, and progression of viral infection [6, 10–22, 24–26]. In this study, we evaluated the genetic diversity within L1 and LCR of HPV16, HPV31, and HPV58 in cervical samples from North-East Brazil. With regard to the HPV16, 23 nucleotide changes in L1 gene and 30 nucleotide changes in LCR were found. In addition, 9 nucleotide changes were found in L1 gene of HPV31 and 30 nucleotide changes also were found in LCR of HPV31. Moreover, 35 nucleotide changes in the L1 gene and LCR of HPV58 were found. Some of these nucleotide changes

are putatively found in T-cell or B-cell epitope and in binding sites of transcriptional factor. Furthermore, two nucleotide changes in LCR of HPV31 and one deletion of seven base pair in LCR of HPV58 were described for the first time in this study. As far as we are aware, this is the first study of the genetic diversity of HPV16, HPV31, and HPV58 L1 and LCR in cervical samples from North-East Brazil.

Nucleotide changes within the HPV16 L1 gene can play an important role in the structure of the capsid protein, immune recognition, and viral neutralization [45]. Hence, viral polymorphisms in the L1 gene can affect the self-assembly of L1 protein in virus-like particles (VLPs) [46]. As a result, Kirnbauer et al. demonstrated that nucleotides change

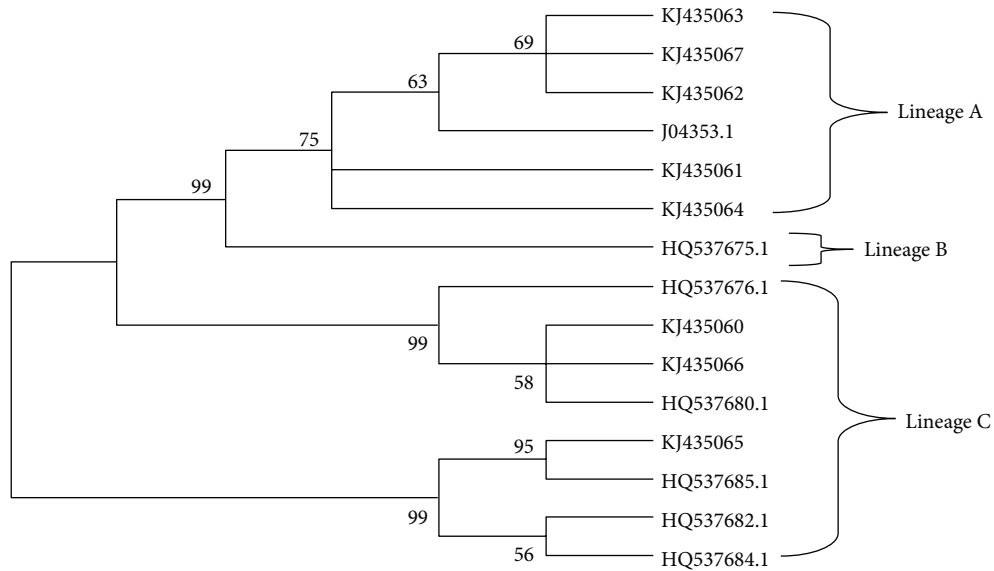


FIGURE 3: Neighbor joining phylogenetic tree of HPV31 variants based on 783 pb of LCR. Reference sequences are listed in Table 2. Three clusters were identified as lineages A, B, and C. Viral lineages analyzed in this study (KJ435060-KJ435067) are clustered into A and C branches. Only bootstrap values above 50% are represented in the branches.

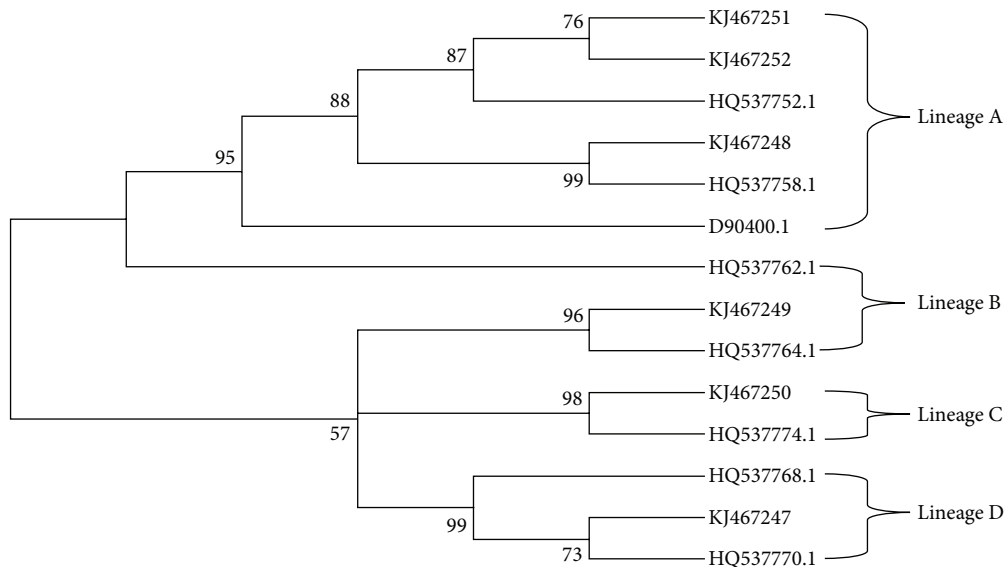


FIGURE 4: Neighbor joining phylogenetic tree of HPV58 variants based on 696 pb of LCR. Four clusters were identified as lineages A, B, C, and D. Reference sequences are listed in Table 2. Viral lineages analyzed in this study (KJ467247-KJ467252) are clustered into A, B, C, and D branches. Only bootstrap values above 50% are represented in the branches.

C6240G, and this leads to a change in the amino acid at position H202D, which is self-assembled within the VLPs with more efficiency in a heterologous system than with a prototype sequence [47]. In addition, it was found that variations in the 83–97 residues of the L1 gene have an impact on the yield of the L1 protein [48]. The nonsynonymous variations found in the L1 gene of HPV16, HPV31, and HPV58 of this study were reported in previous studies [49–58]. Some

of these polymorphisms are located within hypervariable immuno-dominant regions (BC, DE, EF, FG, and HI loops) of L1 protein, which can be recognized as conformational epitopes of HPV [59, 60]. For instance, the A6436G polymorphisms (T292A) found in HPV16 and A6025G (T267A) and C6379A (T274N) found in HPV58 L1 genes are located in the FG loop of the L1 protein. In addition, the A6697C polymorphism (T379P) of HPV16 and C6689A (T375N) of

TABLE 6: HPV31 LCR sequence variation with regard to the reference sequence J04353.1. The table also shows the transcriptional factor binding sites in the nucleotide altered.

Genome position	Nucleotide variation	Number of isolates	Transcriptional factor
7157	A/G	1/8	YY1
7190	T/A*	1/8	C/EBPbeta
7305	Del CTATTTTATA	2/8	
7311	T/C	1/8	
7334	T/C	3/8	
7338	A/G	7/8	
7341	Del TGTTTCCTGCT	8/8	NF-1
7360	A/G	3/8	AP-1, NF-1
7375	T/C	1/8	
7378	G/C	3/8	
7389	A/C*	1/8	
7390	G/A	5/8	
7400	C/A	3/8	YY1, NF-1
7455	G/A	3/8	YY1
7456	A/C	2/8	YY1
7463	G/A	3/8	
7476	A/C	1/8	
7480	C/T	2/8	E2
7512	C/T	2/8	C/EBPbeta
7521	C/A	1/8	C/EBPbeta
7531	G/A	3/8	
7539	T/C	1/8	E2
7581	T/C	3/8	NF-1
7648	C/A	1/8	AP-1
7665	T/G	1/8	
7713	A/G	1/8	
7716	C/T	2/8	
7755	A/C	2/8	
7760	C/A	2/8	
7871	T/G	2/8	E2

Del: deletion. * Novel HPV31 mutation.

HPV58 are located within the HI loop. Both the FG and HI loop constitute the immunodominant epitope region [61]. Furthermore, the polymorphisms found in helix 4 (including threonine and serine at 448 and 465 of L1 protein of HPV16) are implicated in the VLP formation [60].

Nucleotide variation within LCR may influence the binding affinity of the cellular and viral transcriptional factor. For instance, nucleotide changes may result in a loss or insertion of transcriptional factors that regulate the transcription of the of HR-HPV genes [62]. Hence, nucleotide changes in LCR

of specific variants of HPV16, HPV31, and HPV58 may be involved in the alteration in the E6 and/or E7 oncogenes expression which could explain the potential carcinogenesis of some variants [62]. Some of the variations reported in this work are embedded in the putative binding sites for E2, C/EBPbeta, YY1, AP-1, NF-1, and Oct-1 transcriptional factors. These viral and cellular transcriptional factors are involved with early viral genes and differentiation of the epithelium, respectively. Hence, the nucleotide changes found in LCR of HPV16, HPV31, and HPV58 could be an impact directly or indirectly in the expression of E6 and E7 oncogenes.

In addition, we performed a phylogenetic analysis of HPV16 by using fragments of LCR. The results showed 63.63% of isolates belong to the A variant and 33.36% belong to the D variant for HPV16. These results are similar to the previous study performed in Central-West Brazil [63], which showed high prevalence of A and D variants and low frequency of B and C variants. In contrast, a recent study performed in South-Eastern Brazil showed A and C variants as the most prevalent, followed by D and B variants [64]. A previous study in 27 countries and using 953 cervical samples showed the A variant as the most prevalent, followed by C, B, and D variants [8]. These differences in the prevalence of HPV16 variants in different regions of Brazil and worldwide may be explained by geographic origin and ethnicity of the infected patients.

The LCR contains more phylogenetic information than other regions of the HPV16 genome and can distinguish both the lineages and sublineages [8]. Due to the lineage fixation and a putative nonrecombination process, studies have proposed diagnostic polymorphisms to classify both HPV16 lineages and sublineages [8, 9]. Cornet et al. proposed that variant lineages could be detected by using 32 SNP combinations in the LCR of HPV16 [8]. In the light of this, some of these diagnostic SNPs were found in the present study. For instance, the T7747G found in seven isolates of this study are diagnostic SNPs for the AA1 sublineage. Furthermore, G7891G found in seven isolates are diagnostic SNPs for the AA2 sublineage. Both AA1 and AA2 sublineages belong to lineage D [8].

With regard to the HPV31, the phylogenetic trees showed the presence of A and C variants in the North-East Brazil. These results are similar to the results obtained by Chagas et al., which reported high prevalence of A and C variants and very low prevalence of B variant in North-East Brazil [31, 32]. In addition, a recent study performed in Northern China also showed high prevalence of A and C variants [65]. In this study, we did not find any variants that belong to variant B, which was probably due to the small number of isolates analysed or the low frequency of this isolate in the North-East Brazil.

With regard to HPV58, variants that belong to the A, B, C, and D variants were found in North-East Brazil. In this study, A variant was the most prevalent (50%), followed by B (16.6%), C (16.6%), and D (16.6%) variants. In contrast, variant distribution worldwide and in the American continent showed the A variant as the most prevalent, followed by C, D, and E variants [66]. Additional studies should be performed

TABLE 7: HPV58 L1 sequence variations at the genome and protein level and the putative biological function affected.

Genome position	Variation	AA changed	Number of isolates	Protein position relative to HPV16 L1 gene	Biological function
5593	C/T	—	2/8		
5747	T/C	—	1/8		
5789	T/C	—	1/8		
5801	A/G	—	1/8		
5819	A/G	—	1/8		
5861	C/G	—	1/8		
5939	G/A	—	1/8		
5994	G/A	V144I	3/8	DE loop	B-cell/MHC Class-I
6014	A/C	L150F	5/8	DE loop	B-cell/MHC Class-II
6038	C/T	—	1/8		
6039	A/G	—	1/8		
6046	G/A	—	1/8		
6051	C/A	—	1/8		
6172	A/G	—	1/8		
6206	G/A	—	1/8		
6222	A/G	—	1/8		
6405	G/A	—	4/8		
6417	A/G	—	4/8		
6435	T/C	—	4/8		
6440	A/C	—	1/8		
6441	A/C	—	1/8		
6459	G/A	—	1/8		
6459	G/T	—	1/8		
6460	G/A	—	1/8		
6497	T/G	—	1/8		
6497	C/T	—	1/8		
6540	A/G	I325M	6/8	β -G1	MHC Class-I/MHC Class-II
6642	G/A	—	4/8		
6689	C/A	T375N	2/8	HI loop	B-cell/MHC Class-I/MHC Class-II
6693	G/A	—	1/8		
6698	G/A	—	1/8		
6712	G/A	—	1/8		
6799	A/G	I412V	3/8	H2	MHC Class-I/VLP formation
6823	G/A	D420N	2/8	H3	B-cell/MHC Class-I/MHC Class-II/VLP formation
6828	C/A	N422D	4/8	H3	B-cell/MHC Class-I/MHC Class-II/VLP formation
6829	A/G	—	3/8		

(—) Synonymous variations with regard to the prototype reference D90400.1. T: threonine; N: asparagine; D: aspartic acid; V: valine; I: isoleucine; L: leucine; F: phenylalanine; M: methionine.

TABLE 8: HPV58 LCR sequence variation with regard to the reference sequence D90400.1. The table also shows the transcriptional factor binding sites in the nucleotide altered.

Genome position	Nucleotide variation	Number of isolates	Transcriptional factor
7150	G/T	1/6	
7169	Del TATACAT*	1/6	
7187	T/C	1/6	
7188	T/A	1/6	
7189	A/C	1/6	
7189	Del A	1/6	
7188	Del TATGT	1/6	
7197	G/C	1/6	
7202	C/T	1/6	
7210	T/A	3/6	
7233	G/A	1/6	
7244	C/A	1/6	YY1
7268	C/G	1/6	YY1
7269	C/T	2/6	YY1
7283	Ins TGTCAGTTTCCT	1/6	
7299	C/G	3/6	
7319	A/G	1/6	
7347	G/A	1/6	Oct-1, C/EBPbeta
7360	T/C	1/6	NF-1
7384	T/G	1/6	C/EBPbeta
7410	G/A	1/6	
7436	G/A	1/6	Oct-1, C/EBPbeta
7444	T/G	1/6	
7446	T/G	1/6	
7450	A/G	1/6	
7498	T/G	1/6	
7537	T/C	2/6	
7701	G/A	1/6	
7729	A/C	1/6	
7739	A/G	2/6	
7745	C/A	3/6	NF-1
7760	G/C	1/6	AP-1
7760	A/G	1/6	AP-1
7793	T/C	1/6	E2
7794	A/G	3/6	E2

Del: deletion,Ins: insertion. * Novel HPV58 deletion.

to clarify whether these differences in prevalence of HPV58 variants are due to small number of isolates analysed or differences in prevalence of HPV58 variants in North-East Brazil.

In summary, this study reported the prevalence of HPV16, HPV31, and HPV58 variants and sequence variations in the L1 gene and LCR of HPV16, HPV31, and HPV58 isolates from North-East Brazil. Some of the polymorphisms found in the L1 gene are embedded within B-cell or T-cell epitopes. Moreover, some of the variations found in LCR are located

within binding sites of transcriptional factors. Further studies should be carried out to throw light on both the pathological differences and the prevalence of these variants in different geographical regions.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Authors' Contribution

Ana Pavla Almeida Diniz Gurgel carried out all the experiments and wrote the paper; Maria Tereza Cartaxo Muniz and Jacinto da Costa Silva Neto were involved in data collection; Bárbara Simas Chagas, Carolina Medeiros do Amaral, Kamylla Conceição Gomes Nascimento, and Lígia Rosa Sales Leal participated in some of the experiments; Antonio Carlos de Freitas is the supervisor of this study and reviewed the final draft of this paper.

Acknowledgments

This research was supported by the Brazilian Federal Agency for the Support of Postgraduate Education (CAPES), National Council for Scientific and Technological Development (CNPq), and Foundation for the Support of Science and Technology in the State of Pernambuco (FACEPE).

References

- [1] A. Jemal, F. Bray, M. M. Center, J. Ferlay, E. Ward, and D. Forman, "Global cancer statistics," *CA Cancer Journal for Clinicians*, vol. 61, no. 2, pp. 69–90, 2011.
- [2] A. C. de Freitas, "Vaccine strategies against human papillomavirus: a discussion focused on developing countries," *Journal of Clinical & Cellular Immunology*, vol. 1, no. S4, 2012.
- [3] H. Zur Hausen, "Papillomavirus infections—a major cause of human cancers," *Biochimica et Biophysica Acta*, vol. 1288, no. 2, pp. F55–F78, 1996.
- [4] S. de Sanjose, W. G. Quint, L. Alemany et al., "Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study," *The Lancet Oncology*, vol. 11, no. 11, pp. 1048–1056, 2010.
- [5] L. Ho, S.-Y. Chan, R. D. Burk et al., "The genetic drift of human papillomavirus type 16 is a means of reconstructing prehistoric viral spread and the movement of ancient human populations," *Journal of Virology*, vol. 67, no. 11, pp. 6413–6423, 1993.
- [6] T. Yamada, M. M. Manos, J. Peto et al., "Human papillomavirus type 16 sequence variation in cervical cancers: a worldwide perspective," *Journal of Virology*, vol. 71, no. 3, pp. 2463–2472, 1997.
- [7] R. D. Burk, A. Harari, and Z. Chen, "Human papillomavirus genome variants," *Virology*, vol. 445, no. 1-2, pp. 232–243, 2013.
- [8] I. Cornet, T. Gheit, S. Franceschi et al., "Human papillomavirus type 16 genetic variants: phylogeny and classification based on E6 and LCR," *Journal of Virology*, vol. 86, no. 12, pp. 6855–6861, 2012.

- [9] B. Smith, Z. Chen, L. Reimers et al., "Sequence imputation of HPV16 genomes for genetic association studies," *PLoS ONE*, vol. 6, no. 6, Article ID e21375, 2011.
- [10] J. Berumen, R. M. Ordoñez, E. Lazcano et al., "Asian-American variants of human papillomavirus 16 and risk for cervical cancer: a case-control study," *Journal of the National Cancer Institute*, vol. 93, no. 17, pp. 1325–1330, 2001.
- [11] R. D. Burk, M. Terai, P. E. Gravitt et al., "Distribution of human papillomavirus types 16 and 18 variants in squamous cell carcinomas and adenocarcinomas of the cervix," *Cancer Research*, vol. 63, no. 21, pp. 7215–7220, 2003.
- [12] E. Burroni, S. Bisanzi, C. Sani, D. Puliti, and F. Carozzi, "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*, vol. 85, no. 1, pp. 83–90, 2013.
- [13] Y.-J. Chang, H.-C. Chen, B.-H. Lee et al., "Unique variants of human papillomavirus genotypes 52 and 58 and risk of cervical neoplasia," *International Journal of Cancer*, vol. 129, no. 4, pp. 965–973, 2011.
- [14] J. Chansaenroj, A. Theamboonlers, P. Junyangdikul, S. Swangvaree, A. Karalak, and Y. Poovorawan, "Whole genome analysis of human papillomavirus type 16 multiple infection in cervical cancer patients," *Asian Pacific Journal of Cancer Prevention*, vol. 13, no. 2, pp. 599–606, 2012.
- [15] I. Cornet, T. Gheit, G. M. Clifford et al., "Human papillomavirus type 16 E6 variants in France and risk of viral persistence," *Infectious Agents and Cancer*, vol. 8, no. 1, article 4, 2013.
- [16] T. Gheit, I. Cornet, G. M. Clifford et al., "Risks for persistence and progression by human papillomavirus type 16 variant lineages among a population-based sample of Danish women," *Cancer Epidemiology Biomarkers and Prevention*, vol. 20, no. 7, pp. 1315–1321, 2011.
- [17] C.-M. Ho, S.-S. Yang, T.-Y. Chien, S.-H. Huang, C.-J. Jeng, and S.-F. Chang, "Detection and quantitation of human papillomavirus type 16, 18 and 52 DNA in the peripheral blood of cervical cancer patients," *Gynecologic Oncology*, vol. 99, no. 3, pp. 615–621, 2005.
- [18] P. Londsborough, L. Ho, G. Terry, J. Cuzick, C. Wheeler, and A. Singer, "Human papillomavirus genotype as a predictor of persistence and development of high-grade lesions in women with minor cervical abnormalities," *International Journal of Cancer*, vol. 69, no. 5, pp. 364–368, 1996.
- [19] K. D. Quint, M. N. C. de Koning, L.-J. van Doorn, W. G. V. Quint, and E. C. Pirog, "HPV genotyping and HPV16 variant analysis in glandular and squamous neoplastic lesions of the uterine cervix," *Gynecologic Oncology*, vol. 117, no. 2, pp. 297–301, 2010.
- [20] M. Schiffman, A. C. Rodriguez, Z. Chen et al., "A population-based prospective study of carcinogenic human papillomavirus variant lineages, viral persistence, and cervical neoplasia," *Cancer Research*, vol. 70, no. 8, pp. 3159–3169, 2010.
- [21] L. Sichero, S. Ferreira, H. Trottier et al., "High grade cervical lesions are caused preferentially by non-European variants of HPVs 16 and 18," *International Journal of Cancer*, vol. 120, no. 8, pp. 1763–1768, 2007.
- [22] L. L. Villa, L. Sichero, P. Rahal et al., "Molecular variants of human papillomavirus types 16 and 18 preferentially associated with cervical neoplasia," *Journal of General Virology*, vol. 81, no. 12, pp. 2959–2968, 2000.
- [23] L. F. Xi, M. Schiffman, L. A. Koutsky et al., "Association of human papillomavirus type 31 variants with risk of cervical intraepithelial neoplasia grades 2-3," *International Journal of Cancer*, vol. 131, no. 10, pp. 2300–2307, 2012.
- [24] L. F. Xi, L. A. Koutsky, A. Hildesheim et al., "Risk for high-grade cervical intraepithelial neoplasia associated with variants of human papillomavirus types 16 and 18," *Cancer Epidemiology Biomarkers and Prevention*, vol. 16, no. 1, pp. 4–10, 2007.
- [25] I. Zehbe, H. Lichtig, A. Westerback, P. F. Lambert, M. Tommasino, and L. Sherman, "Rare human papillomavirus 16 E6 variants reveal significant oncogenic potential," *Molecular Cancer*, vol. 10, article 77, 2011.
- [26] R. E. Zuna, E. Tuller, N. Wentzensen et al., "HPV16 variant lineage, clinical stage, and survival in women with invasive cervical cancer," *Infectious Agents and Cancer*, vol. 6, no. 1, article 19, 2011.
- [27] T. R. Alencar, D. M. Cerqueira, M. R. D. Cruz, P. S. Wyant, E. D. Ramalho, and C. R. F. Martins, "New HPV-16 European and non-European variants in central Brazil," *Virus Genes*, vol. 35, no. 1, pp. 1–4, 2007.
- [28] M. R. Cruz, D. M. Cerqueira, W. B. Cruz et al., "Prevalence of human papillomavirus type 16 variants in the Federal District, Central Brazil," *Memorias do Instituto Oswaldo Cruz*, vol. 99, no. 3, pp. 281–282, 2004.
- [29] D. Marreco Cerqueira, G. N. De Lima Camar, M. R. Da Cruz et al., "Variants of human papillomavirus types 53, 58 and 66 identified in Central Brazil," *Virus Genes*, vol. 26, no. 1, pp. 83–87, 2003.
- [30] T. Raiol, P. S. Wyant, R. M. S. de Amorim et al., "Genetic variability and phylogeny of the high-risk HPV-31, -33, -35, -52, and -58 in central Brazil," *Journal of Medical Virology*, vol. 81, no. 4, pp. 685–692, 2009.
- [31] B. S. Chagas, M. V. D. A. Batista, S. Crovella et al., "Novel E6 and E7 oncogenes variants of human papillomavirus type 31 in Brazilian women with abnormal cervical cytology," *Infection, Genetics and Evolution*, vol. 16, pp. 13–18, 2013.
- [32] B. S. Chagas, M. V. A. Batista, V. Guimarães, V. Q. Balbino, S. Crovella, and A. C. Freitas, "New variants of E6 and E7 oncogenes of human papillomavirus type 31 identified in Northeastern Brazil," *Gynecologic Oncology*, vol. 123, no. 2, pp. 284–288, 2011.
- [33] A. P. A. D. Gurgel, B. S. Chagas, C. M. M. do Amaral 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*, vol. 2013, Article ID 546354, 7 pages, 2013.
- [34] M. F. P. T. B. da Silva, B. S. Chagas, V. Guimarães et al., "HPV31 and HPV33 incidence in cervical samples from women in Recife, Brazil," *Genetics and Molecular Research*, vol. 8, no. 4, pp. 1437–1443, 2009.
- [35] J. V. Fernandes, R. V. Meissner, M. G. Carvalho et al., "Prevalence of human papillomavirus in archival samples obtained from patients with cervical pre-malignant and malignant lesions from Northeast Brazil," *BMC Research Notes*, vol. 3, no. 1, article 96, 2010.
- [36] J. V. Fernandes, R. D. V. Meissner, M. G. F. Carvalho et al., "Human papillomavirus prevalence in women with normal cytology and with cervical cancer in Natal, Brazil," *Molecular Medicine Reports*, vol. 4, no. 6, pp. 1321–1326, 2011.
- [37] C. Kämmer, M. Tommasino, S. Syrjänen et al., "Variants of the long control region and the E6 oncogene in European human papillomavirus type 16 isolates: implications for cervical disease," *British Journal of Cancer*, vol. 86, no. 2, pp. 269–273, 2002.

- [38] R. Staden, "The Staden sequence analysis package," *Molecular Biotechnology*, vol. 5, no. 3, pp. 233–241, 1996.
- [39] K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar, "MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods," *Molecular Biology and Evolution*, vol. 28, no. 10, pp. 2731–2739, 2011.
- [40] S. Saha and G. P. S. Raghava, "BcePred: prediction of continuous B-cell epitopes in antigenic sequences using physico-chemical properties," in *Proceedings of the 3rd International Conference on Artificial Immune Systems (ICARIS '04)*, G. Nicosia, V. Cutello, P. J. Bentley, and J. Timis, Eds., vol. 3239 of *Lecture Notes in Computer Science*, pp. 197–204, Springer, 2004.
- [41] H. Singh and G. P. S. Raghava, "ProPredI: prediction of promiscuous MHC class-I binding sites," *Bioinformatics*, vol. 19, no. 8, pp. 1009–1014, 2003.
- [42] H. Singh and G. P. S. Raghava, "ProPred: prediction of HLA-DR binding sites," *Bioinformatics*, vol. 17, no. 12, pp. 1236–1237, 2002.
- [43] D. Farré, R. Roset, M. Huerta et al., "Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN," *Nucleic Acids Research*, vol. 31, no. 13, pp. 3651–3653, 2003.
- [44] X. Messeguer, R. Escudero, D. Farré, O. Núñez, J. Martínez, and M. M. Albà, "PROMO: detection of known transcription regulatory elements using species-tailored searches," *Bioinformatics*, vol. 18, no. 2, pp. 333–334, 2002.
- [45] S. Pande, N. Jain, B. K. Prusty et al., "Human papillomavirus type 16 variant analysis of E6, E7, and L1 genes and long control region in biopsy samples from cervical cancer patients in North India," *Journal of Clinical Microbiology*, vol. 46, no. 3, pp. 1060–1066, 2008.
- [46] M. Sapp and P. M. Day, "Structure, attachment and entry of polyoma- and papillomaviruses," *Virology*, vol. 384, no. 2, pp. 400–409, 2009.
- [47] R. Kirnbauer, J. Taub, H. Greenstone et al., "Efficient self-assembly of human papillomavirus type 16 L1 and L1-L2 into virus-like particles," *Journal of Virology*, vol. 67, no. 12, pp. 6929–6936, 1993.
- [48] A. Touze, S. El Mehdaoui, P.-Y. Sizaret, C. Mougin, N. Muñoz, and P. Coursaget, "The L1 major capsid protein of human papillomavirus type 16 variants affects yield of virus-like particles produced in an insect cell expression system," *Journal of Clinical Microbiology*, vol. 36, no. 7, pp. 2046–2051, 1998.
- [49] V. Cento, M. Ciccozzi, L. Ronga, C. F. Perno, and M. Ciotti, "Genetic diversity of human papillomavirus type 16 E6, E7, and L1 genes in Italian women with different grades of cervical lesions," *Journal of Medical Virology*, vol. 81, no. 9, pp. 1627–1634, 2009.
- [50] M. A. de Boer, L. A. W. Peters, M. F. Aziz et al., "Human papillomavirus type 16 E6, E7, and L1 variants in cervical cancer in Indonesia, Suriname, and the Netherlands," *Gynecologic Oncology*, vol. 94, no. 2, pp. 488–494, 2004.
- [51] E. Frati, S. Bianchi, D. Colzani, A. Zappa, G. Orlando, and E. Tanzi, "Genetic variability in the major capsid L1 protein of human papillomavirus type 16 (HPV-16) and 18 (HPV-18)," *Infection, Genetics and Evolution*, vol. 11, no. 8, pp. 2119–2124, 2011.
- [52] J.-H. Liu, Z.-T. Lu, G.-L. Wang et al., "Variations of human papillomavirus type 58 E6, E7, L1 genes and long control region in strains from women with cervical lesions in Liaoning province, China," *Infection, Genetics and Evolution*, vol. 12, no. 7, pp. 1466–1472, 2012.
- [53] C. K. Ntova, C. Kottaridi, A. Chranioti et al., "Genetic variability and phylogeny of high risk HPV type 16, 18, 31, 33 and 45 L1 gene in Greek women," *International Journal of Molecular Sciences*, vol. 13, no. 1, pp. 1–17, 2012.
- [54] M. R. Pillai, R. Hariharan, J. M. Babu et al., "Molecular variants of HPV-16 associated with cervical cancer in Indian population," *International Journal of Cancer*, vol. 125, no. 1, pp. 91–103, 2009.
- [55] Q. Shang, Y. Wang, Y. Fang et al., "Human papillomavirus type 16 variant analysis of E6, E7, and L1 genes and long control region in identification of cervical carcinomas in patients in northeast China," *Journal of Clinical Microbiology*, vol. 49, no. 7, pp. 2656–2663, 2011.
- [56] Z. Sun, G. Ren, X. Cui, W. Zhou, C. Liu, and Q. Ruan, "Genetic diversity of HPV-16 E6, E7, and L1 genes in women with cervical lesions in Liaoning Province, China," *International Journal of Gynecological Cancer*, vol. 21, no. 3, pp. 551–558, 2011.
- [57] C. M. Wheeler, T. Yamada, A. Hildesheim, and S. A. Jenison, "Human papillomavirus type 16 sequence variants: identification by E6 and L1 lineage-specific hybridization," *Journal of Clinical Microbiology*, vol. 35, no. 1, pp. 11–19, 1997.
- [58] Y. Yue, H. Yang, K. Wu et al., "Genetic variability in L1 and L2 genes of HPV-16 and HPV-58 in Southwest China," *PLoS ONE*, vol. 8, no. 1, Article ID e55204, 2013.
- [59] B. Bishop, J. Dasgupta, and X. S. Chen, "Structure-based engineering of papillomavirus major capsid L1: controlling particle assembly," *Virology Journal*, vol. 4, article 3, 2007.
- [60] X. S. Chen, R. L. Garcea, I. Goldberg, G. Casini, and S. C. Harrison, "Structure of small virus-like particles assembled from the L1 protein of human papillomavirus 16," *Molecular Cell*, vol. 5, no. 3, pp. 557–567, 2000.
- [61] N. D. Christensen, N. M. Cladel, C. A. Reed et al., "Hybrid papillomavirus L1 molecules assemble into virus-like particles that reconstitute conformational epitopes and induce neutralizing antibodies to distinct HPV types," *Virology*, vol. 291, no. 2, pp. 324–334, 2001.
- [62] L. Slichero, J. S. Sobrinho, and L. L. Villa, "Identification of novel cellular transcription factors that regulate early promoters of human papillomavirus types 18 and 16," *Journal of Infectious Diseases*, vol. 206, no. 6, pp. 867–874, 2012.
- [63] M. R. Cruz, D. M. Cerqueira, W. B. Cruz et al., "Prevalence of human papillomavirus type 16 variants in the Federal District, Central Brazil," *Memorias do Instituto Oswaldo Cruz*, vol. 99, no. 3, pp. 281–282, 2004.
- [64] L. B. Freitas, Z. Chen, E. F. Muqui et al., "Human papillomavirus 16 non-European variants are preferentially associated with high-grade cervical lesions," *PLoS ONE*, vol. 9, no. 7, Article ID e100746, 2014.
- [65] M. Liu, Z. He, L. Xi et al., "The distribution and common amino acid polymorphisms of human papillomavirus (HPV)-31 variants in 2700 women from Northern China," *PLoS ONE*, vol. 9, no. 6, Article ID e99141, 2014.
- [66] P. K. S. Chan, A. C. S. Luk, J.-S. Park et al., "Identification of human papillomavirus type 58 lineages and the distribution worldwide," *Journal of Infectious Diseases*, vol. 203, no. 11, pp. 1565–1573, 2011.