



Article Genetic Variability in the E6, E7, and L1 Genes of Human Papillomavirus Types 16 and 18 among Women in Saudi Arabia

Madain Alsanea ¹, Asma Alsaleh ², Dalia Obeid ³, Faten Alhadeq ¹, Basma Alahideb ¹ and Fatimah Alhamlan ^{1,4,5,*}

- ¹ Department of Infection and Immunity, King Faisal Specialist Hospital and Research Center, Riyadh 11564, Saudi Arabia
- ² Department of Botany and Microbiology, College of Science, King Saud University, Riyadh 11564, Saudi Arabia
- ³ Public Health Laboratories, Public Health Authority, Riyadh 11564, Saudi Arabia
- ⁴ Department of Pathology and Laboratory Medicine, King Faisal Specialist Hospital and Research Centre, Riyadh 11564, Saudi Arabia
- ⁵ College of Medicine, Alfaisal University, Riyadh 11564, Saudi Arabia
- * Correspondence: falhamlan@kfshrc.edu.sa; Tel.: +966-11-442-4365; Fax: +966-11-442-4519

Abstract: Cervical cancer is the eighth most frequent cancer in Saudi Arabia, and most cases are associated with human papillomavirus (HPV) types 16 and 18. HPV-induced carcinogenesis may be associated with the intra-type variant, genetic mutation, or the continuous expression of viral oncogenes E6 and E7. Infection efficiency and virus antigenicity may be affected by changes in the L1 gene. Thus, this retrospective cohort study analyzed E6, E7, and L1 gene mutations in cervical specimens collected from Saudi women positive for HPV16 or HPV18 infection. HPV16 and HPV18 lineages in these specimens were predominantly from Europe. The L83V mutation in the E6 gene of HPV16 showed sufficient oncogenic potential for progression to cervical cancer. By contrast, the L28F mutation in the E7 gene of HPV16 was associated with a low risk of cervical cancer. Other specific HPV16 and HPV18 mutations were associated with an increased risk of cancer, cancer progression, viral load, and age. Four novel mutations, K53T, K53N, R365P, and K443N, were identified in the L1 genes among women in Saudi Arabia may inform the design and development of effective molecular diagnostic tests and vaccination strategies for the Saudi population.

Keywords: HPV16; HPV18; E6 gene; E7 gene; L1 gene; cervical cancer; cancer progression; mutations; lineages; variants

1. Introduction

Human papillomaviruses (HPVs) are a group of small (8 kbp), non-enveloped, doublestranded DNA viruses [1,2]. Their genome can be divided into three major regions: (1) the early region (E), which contains six common open reading frames (E1, E2, E4, E5, E6, and E7) that encode nonstructural proteins; (2) the late region (L), which is divided into two open reading frames that encode one major (L1) and one minor (L2) capsid protein; and (3) the long control region, also known as the noncoding region, which does not have a protein-coding function but contains the origin of replication as well as multiple binding sites for transcription factors [1,3–6].

This group of viruses is responsible for numerous epithelial cell infections, varying from warts to cancer [7,8]. Low-risk HPV types, such as types HPV6 and HPV11, are mainly observed in genital warts, whereas high-risk HPV types, such as HPV16 and HPV18, are found in precancerous lesions and cervical cancer [7]. In 2020, cervical cancer was the fourth most common cancer among women worldwide [9]. In Saudi Arabia, cervical cancer



Citation: Alsanea, M.; Alsaleh, A.; Obeid, D.; Alhadeq, F.; Alahideb, B.; Alhamlan, F. Genetic Variability in the E6, E7, and L1 Genes of Human Papillomavirus Types 16 and 18 among Women in Saudi Arabia. *Viruses* **2023**, *15*, 109. https:// doi.org/10.3390/v15010109

Academic Editors: Susana Tereno Valente, Barbara Baker and Sharon Lewin

Received: 25 October 2022 Revised: 6 December 2022 Accepted: 14 December 2022 Published: 30 December 2022



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). is the eighth most frequent cancer among women between the ages of 15 and 44 years, with an estimated 358 new cases and 179 deaths occurring annually [10].

HPV16 and HPV18 account for 70% or more of all cervical cancer cases reported worldwide [11–14]. In Saudi Arabia, HPV16 and HPV18 were the most common types detected in 63.4% and 11.1% of cervical cancer samples, respectively, in 2011 [15]. In 2020, a 10-year retrospective analysis conducted in Saudi Arabia found that HPV16 was detected in 56.3% of cervical cancer samples, and HPV18 was detected in 7.3% of cervical cancer samples [16]. However, data assessing the HPV burden in the general population are not available [10].

Some HPV infections progress to high-grade cervical lesions and cancer. This means that there must be additional factors related to either the virus or the host contributing to virus persistence and progression [17,18]. Viral factors, such as virus integration, high-risk (HR)-HPV mutations, and viral load, and the presence of host-related factors, including age, marital status, number of sexual partners, number of pregnancies, and smoking, may increase the risk of virus persistence and influence the progression to cancer [19–23].

Some genetic mutations alter the biological function of the protein encoded by the mutant gene, which in turn may influence the persistence of infection, the progression of lesions to cancer, and morbidity from cervical cancer [24–26]. The rate of spontaneous mutagenesis is exceedingly low in normal human cells, but it dramatically increases with the expression of high-risk HPV E6/E7 proteins [5]. When oncogenes E6 and E7 act cooperatively, they interact with and inhibit the activities of tumor suppressors to enable malignant cell transformation [8,27]. The L1 gene, which plays an important role in the initial interaction of the capsid with host cells and mediates the integration into the nucleus [6,24,28], can be used to design HPV detection primers for clinical use [29,30]. The identification, screening, and analyses of various genomic regions of HR-HPV genotypes, such as HPV16 and HPV18, are important for developing HPV diagnostics, vaccines, and other therapeutic approaches to managing HPV-induced carcinogenesis [24,26,31,32].

Because the accurate detection and typing of HPVs are essential in screening and disease management protocols, the present study aimed to investigate the genetic variability in the E6, E7, and L1 genes of HPV16 and HPV18 in HPV-positive specimens isolated from Saudi female patients. Our study also aimed to assess HPV16 and HPV18 genetic mutations and the risk of cervical cancer in these specimens. By understanding the genetic variability in our population, we can design better screening programs and improved assays.

2. Materials and Methods

2.1. Ethics Statement

This retrospective cohort study was approved by the Research Advisory Council (Ethics Committee) at King Faisal Specialist Hospital and Research Centre (approval number: RAC 1005-033) with waived consent because the cervical biopsies were archived, and the collected patient data were coded and thus deidentified.

2.2. Study Population

In total, 1460 archived cervical specimens or biopsy specimens obtained from women attending Family Medicine and OB/GYN clinics at King Faisal Specialist Hospital and Research Centre in Riyadh were included in this study. All specimens were collected between 2017 and 2019. Clinical and sociodemographic data, as well as the viral load data of the specimens, were obtained from previous studies [23,33,34]. Participant inclusion criteria for the present study were (1) being a Saudi woman (2) who was married, divorced, or widowed, (3) having definite histological or cytological results for collected cervical specimens, and (4) having an HPV genotype test result positive for HPV16 or HPV18 infection. Abnormal cytological stages were identified using the Bethesda classification system [35]. A post-inclusion criterion was that the total sample size was 60.

2.3. Genomic DNA Extraction and Quality Control

For specimens obtained through Pap testing, DNA was extracted by using a Gentra Puregene Cell Kit according to the manufacturer's instructions (Qiagen, Hilden, Germany). For formalin-fixed paraffin-embedded cervical specimens, DNA was extracted using a QIAamp DNA FFPE Tissue Kit by following the manufacturer's instructions (Qiagen, Valencia, CA, USA). The β -globin housekeeping gene was tested to ensure the quality of the genomic DNA. The housekeeping gene primers used are given in Table 1.

Table 1. Housekeeping gene primers.

Primer	Sequence	Length, Bases	Product Size, Base Pairs	Melting Temperature
GH20	CAACTTCATCCACGTTCACC	20	249	F0 0 <i>C</i>
PC04	GAAGAGCCAAGGACAGGTAC	20	248	58 °C

The PCR cycling parameters are given in Table 2. The amplified products were visualized using 1.5% agarose gels stained with Green View[™] Plus that had been electrophoresed for 90 min at 120 volts per meter. A 100-base-pair (bp) DNA ladder was used to identify the product size. The Gel Doc[™] EZ imaging system from Bio-Rad was used to capture images.

Table 2. PCR c	ycling parameters.
----------------	--------------------

Step	Temperature, °C	Time	Cycle
Initial Denaturation	95	5 min	1
Denaturation	95	30 s	
Annealing	58 *	30 s	45
Extension	72	30 s	
	72	5 min	1
Final Extension	4	X	0

* For HPV16 primers, the touchdown PCR ranged from 64 $^{\circ}$ C to 53 $^{\circ}$ C, whereas for HPV18, the range was from 61 $^{\circ}$ C to 56 $^{\circ}$ C.

2.4. HPV Detection and Genotyping

Well-established sets of nested PCR primers were used to detect the L1 region in the HPV genome, including the MY09/MY11 primer set, which targets a 450 bp conserved sequence, and the GP5+/GP6+ primer set, which targets a 150 bp sequence within the 450 bp product. For positive controls, SiHa and HeLa cell lines were used for HPV16 and HPV18, respectively. HPV genotyping was performed by using two assays: a reverse line blotting hybridization assay that used 23 oligonucleotide probes and a reverse dot blot (GenoFlow) HPV array test (Diagcor Bioscience Incorporation Limited; Hong Kong), with the ability to genotype 33 HPV types. All typing assays have been described previously [36]. Viral loads for HPV16 and HPV18 were assessed using in-house relative assays (TaqMan), which have also been described previously [37]. Both assays were performed using the cycling conditions given in Table 2 with the 7500 Fast Real-Time PCR system and software (Applied Biosystems, Foster City, CA, USA).

2.5. Amplification of Target Genes

In-house target-specific primers for HPV16 E6, E7, and L1 genes and for HPV18 E6, E7, and L1 genes were designed using Primer3 version 4.1.0 (https://primer3.ut.ee/ (accessed on 3 February 2021). The primers were designed referring to the HPV16 (NC_001526) and HPV18 (NC_001357.1) prototypes published in GenBank (National Center for Biotechnology Information; NCBI) [30]. All primers are shown in Tables 3 and 4.

For annealing temperature optimization, gradient PCR was performed. The 23 μ L reaction solutions contained 12.5 μ L of GoTaq[®] Green Master Mix (Promega, Madison, WI USA,), 9.5 μ L of UltraPure DNase/RNase–free water, 0.5 μ L of forward primer, and 0.5 μ L of reverse primer. The primer concentration was 95 pmol (5 μ L of primer added

to 95 μ L of 1X TE buffer) and included 2 μ L of the DNA template. SiHa and HeLa cell lines were used for HPV16 and HPV18, respectively. For the negative control, UltraPure DNase/RNase–free water was used in this optimization. For the HPV16 primers, the temperature for touchdown PCR ranged from 64 °C to 53 °C, and the HPV18 range was from 61 °C to 56 °C. Cycling parameters for both HPV16 and HPV18 primers are given in Table 2. The bands resulting from the amplified products were compared visually after gel electrophoresis.

Primer		Length, Bases	Product Size, bp	Tm, °C	Gene Location (Sequence Length)
HPV16-E6 (F) HPV16-E6 (R)	ACCGGTTAGTATAAAAGCAGACA AGCGTAGAGTCACACTTGCA	23 20	700	57.59 59.33	56–755 (699)
	Sequence	Length, Bases	Product Size, bp	Tm, °C	Gene Location
HPV16-E7 (F) HPV16-E7 (R)	TGTCAAAAGCCACTRTGTCC GTCATYTGATAYRGCATCCCCTGT	20 24	553	59.73 59.83	419–972 (553)
	Sequence	Length, Bases	Product Size, bp	Tm, °C	Gene Location
HPV16-L1 fragment 1 (F)	AACTGACCAARCTCCTTCATT	21	661	57.48	5504-6164
HPV16-L1 fragment 1 (R)	GATCCTTTGCCCCAGTGTTC	20		58.82	(000)
_	Sequence	Length, Bases	Product Size, bp	Tm, °C	Gene Location
HPV16-L1 fragment 2 (F)	TGGATGACACMGAAAATGCTAGT	23	678	58.09	6023–6692 (669)
HPV16-L1 fragment 2 (R)	CTGAAGTWGATATGGCAGCACA	22		57.94	(00))
	Sequence	Length, Bases	Product Size, bp	Tm, °C	Gene Location
HPV16-L1 fragment 3 (F)	GAACCATATGGCGACAGCTT	20	865	60.10	6364–7228 (864)
HPV16-L1 fragment 3 (R)	GCACATACAAGCACATACAAACA	23		58.70	(001)

Table 3. In-house target-specific primers for HPV16 E6 and E7 and the multiple primer sets for L1.

Abbreviations: F, represents forward; R, reverse; bp, base pair; Tm, melting temperature.

Table 4. In-house target-specific primers for HPV18 E6 and E7 and the multiple primer sets for L1.

Primer		Length, Bases	Product Size, bp	Tm, °C	Gene Location (Sequence Length)
HPV18-E6 (F)	GGGACCGAAAACGGTGTAT	19	611	59.66	55–665
HPV18-E6 (R)	GAAGGTCAACCGGAATTTCA	20		59.91	(610)
	Sequence	Length, Bases	Product Size, bp	Tm, °C	Gene Location
HPV18-E7 (F)	CTCCAACGACGCAGAGAAAC	20	495	59.22	552-1064
HPV18-E7 (R)	ACCCTGTGTCTGTTGCATTT	20		57.93	(512)
	Sequence	Length, Bases	Product Size, bp	Tm, °C	Gene Location
HPV18-L1 fragment 1 (F)	CGTYCTACTACCTCCTYTGCA	21	700	57.41	5333-6032
HPV18-L1 fragment 1 (R)	AGAAACATTAGACGTRGCGG	20		58.28	(699)
	Sequence	Length, Bases	Product Size, bp	Tm, °C	Gene Location
HPV18-L1 fragment 2 (F)	CAACGTTTAGTGTGGGCCTG	20	698	59.41	5898–6595 (697)
HPV18-L1 fragment 2 (R)	TGATTATGCCAGCARAYACCA	21		58.48	(077)

	Primer	Length, Bases	Product Size, bp	Tm, °C	Gene Location (Sequence Length)
	Sequence	Length, Bases	Product Size, bp	Tm, °C	Gene Location
HPV18-L1 fragment 3 (F)	CCAAGTGGCTCTATTGTTACCTC	23	696	58.81	6532–7136 (604)
HPV18-L1 fragment 3 (R)	ACACARGACATACAAACACAACA	23		58.73	(004)

Table 4. Cont.

Abbreviations: F, represents forward; R, reverse; bp, base pair; Tm, melting temperature.

2.6. Sequencing and Sequence Analysis

The PCR products were sequenced at the Sequencing Core Facility, Department of Genetics, King Faisal Specialist Hospital and Research Centre, using an ABI3730XL DNA Analyzer (Applied Biosystems, Foster City, CA, USA). Sanger shotgun sequencing was performed for the L1 gene. HPV16 and HPV18 L1 genes were fragmented into three pieces. The HPV16 fragments were 660 bp, 669 bp, and 864 bp, and the HPV18 fragments were 700 bp, 698 bp, and 696 bp. DNASTAR Lasergene 15.0 package SeqMan Pro, version 15, was used for sequence analyses and the assembly of the L1 gene contigs. Sequences were then aligned against the prototype references NC_001526 for HPV16 and NC_001357.1 for HPV18 to detect variants by using ClustalW and MegAlign Pro, version 15.

To determine the amino acid variants, all sequences were translated into amino acids using EditSeq, version 15, and realigned by MUSCLE to the following reference protein sequences: for HPV16, E6 (accession No. NP_041325.1), E7 (accession No. NP_041326.1), and L1 (accession No. NP_041332.2); for HPV18, E6 (accession No. NP_040310.1), E7 (accession No. NP_040311.1), and L1 (accession No. NC_001357.1). Both nucleotide and amino acid variant positions were revised and confirmed by using SeqBuilder Pro, version 15, and by comparing them with published variants.

2.7. Phylogenetic Tree Construction

Phylogenetic trees were constructed using IQ-TREE (http://iqtree.cibiv.univie.ac.at (accessed on 3 March 2021) [38] based on the E6 and L1 genes. Maximum likelihood values were calculated with 1000 replicates to test the robustness of the major phylogenetic group. FigTree 1.4 was used to view and edit the phylogenetic trees [39]. For the construction of the HPV16 and HPV18 phylogenetic trees, sublineage references were obtained from the NCBI GenBank database and the PapillomaVirus Episteme database, as shown in Table 5 [39,40].

Table 5. HPV16 and HPV18 sublineage references for phylogenetic tree construction.

HPV16 Sublineage	Accession Number	Database	HPV18 Sublineage	Accession Number	Database
A1	HQ644280.1 HQ644282.1 HQ644283.1 HQ644268.1 NC_001526.4	NCBI	A1	AY262282.1	NCBI and PaVE
A2	KY549170.1 AF536179.1	NCBI NCBI and PaVE	A2	EF202146.1	NCBI and PaVE
A3	AB889494.1 HQ644236.1	NCBI NCBI and PaVE	A3	EF202147.1	NCBI and PaVE
A4	HQ644248.1 AF534061.1	NCBI NCBI and PaVE	A4 A5	EF202150.1 GQ180787.1	NCBI NCBI and PaVE
B1	HQ644290.1 HQ644293.1 AF536180.1	NCBI NCBI NCBI and PaVE	B1	EF202155.1	NCBI and PaVE

HPV16 Sublineage	Accession Number	Database	HPV18 Sublineage	Accession Number	Database
Po	KU053915.1	NCBI	B2	KC470225.1	NCBI and PaVE
B2	HQ644298.1	NCBI and PaVE	B3	EF202152.1	NCBI and PaVE
	HQ644249.1	NCBI		KC470229.1	
С	KY549231.1	NCBI	С	KC 470000 1	NCBI and PaVE
	AF472509.1	NCBI and PaVE		KC470229.1	
D1	HQ644257.1	NCBI and PaVE	-	-	-
D2	HQ644279.1	NCBI and PaVE	-	-	-
D2	AF402678.1	NCBI and PaVE	-	-	-
D3	HQ644255.1	NCBI	-	-	-

Table 5. Cont.

Abbreviations: PaVE represents PapillomaVirus Episteme; NCBI, National Center for Biotechnology Information.

2.8. Statistical Analysis

Data were analyzed using Statistical Analytical Software (SAS), version 9.4 (SAS Institute Inc.; Cary, NC, USA), and GraphPad, version 8 (San Diego, CA, USA). *T*-tests or Wilcoxon rank-sum tests were used to assess differences in numerical values between two groups. For analyses of multiple factors, analysis of variance (ANOVA) or Kruskal–Wallis tests were used, as appropriate. Chi-square tests for associations were conducted to evaluate the distribution of categorical variables. Odds ratios (ORs) were used to evaluate cancer risk by variant. Descriptive analyses were conducted to assess the distribution of cases by group with time and were visualized using bar graphs and scatter plots. A 2-sided *p*-value of alpha < 0.05 was considered statistically significant.

3. Results

3.1. Study Population

Tests assessing a housekeeping gene were performed on 67 archived specimens that had passed our initial inclusion criteria. We found 60 (89.6%) specimens that were positive for the housekeeping gene, and we excluded 7 (10.4%) specimens that were negative for the gene. Of 60 specimens included in the ensuing analyses, 54 specimens were positive for the presence of HPV16, and 6 specimens were HPV18-positive (Figure 1).

In total, 70.3% of the HPV16-positive specimens and 100% of the HPV18-positive specimens were obtained from married women. The mean (SD) age of the women providing the HPV16-positive specimens was 53.9 (12.9) years, and the mean (SD) age of the women providing the HPV18-positive specimens was 48 (8.9) years (Table 6). Approximately 81% of the HPV16-positive specimens were collected from women who were diagnosed as having cervical cancer, whereas all of the HPV18-positive specimens were collected from system were collected from women who were diagnosed as having cervical cancer, the histology and cytology results of all cervical tissues collected for this study are given in Table 6.

Table 6. Summary of the demographic and clinical characteristics of all our study participants.

	Human Papillomavirus Type 16 Total Sample Size = 54		Human Papillomavirus Type 18 Total Sample Size = 6	
	n (%)	$\chi^2/(p$ -Value)	n (%)	$\chi^2/(p ext{-Value})$
Marital status				
Divorced	3 (5.6)	0(// 00001)*	0	N 7.4
Married	38 (70.3)	$36/(p < 0.0001)^{+}$	6 (100)	NA
Widowed	13 (24.1)		0	
Age mean (SD)	53.	53.9 (12.9)		8 (8.9)
Minimum	30		37	
Maximum	82		60	

	Human Papillomavirus Type 16 Total Sample Size = 54		Human Papillor Total Samp	navirus Type 18 Je Size = 6
	n (%)	$\chi^2/(p$ -Value)	n (%)	$\chi^2/(p$ -Value)
Previously diagnosed with cervical cancer				
Yes	44 (81.5)	57/(p < 0.0001) *	6 (100)	NA
No	8 (14.8)	•	0	
Unknown	2 (3.7)		0	
Histology results	、			
CIN 1	1 (2.9)		0	
CIN 2	3 (8.6)		0	
CIN 3	15 (42.9)	27.1/(p < 0.0001) *	0	NA
Invasive carcinoma	14 (40)	N	1 (20)	
Cell carcinoma	0		3 (60)	
NIEL	2 (5.7)		1 (20)	
Cytology test				
HSIL	11 (25.6)	24.7/(0	NT A
Invasive carcinoma	29 (67.4)	$24.7/(p < 0.0001)^{+}$	3 (100)	NA
NIEL	3 (6.9)		0	
SCC grades, for $N = 29$				
SCC Grade I	16 (55.2)	8 7 /(m < 0.012E) *	1 (25)	NT A
SCC Grade II	10 (34.5)	$8.7/(p < 0.0123)^{+}$	1 (25)	NA
SCC Grade III	3 (10.3)		2 (50)	
AC SCC, $N = 29$		N T A		NT A
Adenocarcinoma	2 (6.9)	NA	0	NA
Viral load Mean (SD), viral copies/μL	5074 (12890)		22850 (49144)	

Table 6. Cont.

Abbreviations: CIN 1, mild dysplasia; CIN 2, moderate dysplasia; CIN 3 severe dysplasia; NIEL, negative for intraepithelial lesion; HSIL, high-grade squamous intraepithelial lesion; SCC, squamous cell carcinoma; AC, Adenocarcinoma; NA, not applicable. * Indicates statistical significance (p < 0.05). NA, not applicable for statistical testing due to the poor sample size.



Figure 1. Schematic of the final study sample population after application of inclusion and exclusion criteria. Numerals in boxes represent the number of specimens. HKG indicates the housekeeping gene assay.

3.2. Sequence Analyses

For specimens that were HPV16-positive (N = 54), 72 nucleic acid variants were detected: 19 were located in the E6 gene, 11 were in the E7 gene, and 42 were in the L1

gene. Among these 72 variants, 30 resulted in amino acid mutations, with 12 of these amino acid mutations detected in the E6 gene, 2 in the E7 gene, and 16 in the L1 gene. A summary of the frequencies and percentages of the detected variants is given in the Supplementary Materials Table S1.

In 6 sequences that were obtained from HPV18-positive specimens, 20 nucleic acid mutations were detected, 4 in the E6 gene, 4 in the E7 gene, and 12 in the L1 gene. Among these 20 nucleic acid mutations, 8 resulted in amino acid mutations, 2 in the E7 gene and 6 in the L1 gene. A summary of the frequencies and percentages of the detected variants is given in the Supplementary Materials Table S1.

The HPV16 L1 gene had four nonsynonymous novel mutations: K53T, K53N, R365P, and K443N. The chromatograms are shown in Figure 2.



Figure 2. Novel mutations detected in the study population. (**A**) In specimen 219, A5796C led to changing lysine in position 53 to threonine (K53T). (**B**) In specimen 184, A5797C led to changing lysine in position 53 to asparagine (K53N). (**C**) In specimen 163, G6732C led to changing arginine in position 365 to proline (R365P). (**D**) In sample 136, A6967C led to changing lysine in position 443 to asparagine (K443N).

3.3. Association of HPV16 Variants with Cervical Cancer Risk

We evaluated the risk of cervical cancer associated with the various HPV16 variants, but not the HPV18 variants because all of the HPV18-positive specimens were collected from women who had already received a diagnosis of cervical cancer. Using Fisher's exact test, we found a significant excess risk of cervical cancer associated with three HPV16 variants, one of which was observed in each of the studied genomic regions. In the E6 gene, the highest increased risk of cervical cancer was for the L83V amino acid mutation (OR: 10.1; 95% confidence interval (CI): 1.1–89.4). In the E7 gene, the L28F mutation was associated with a significantly lower risk of cervical cancer (OR: 0.14; 95% CI: 0.07–0.27). In the L1 gene, the highest significant risk was for the variant with the T266A amino acid change (OR: 4.5; 95% CI: 0.92–21.9). Thus, the two main amino acid mutations that showed an increased risk of cervical cancer associated with HPV16 infection were L84V and T266A, whereas the L28F mutation was associated with a lower risk. All tables reporting the ORs and statistical tests are shown in the Supplementary Materials Tables S2–S4.

3.4. Association of Variants with Cervical Dysplasia Evaluated by Histological and Cytological Grades

We used chi-squared tests to assess the associations of the HPV16 and HPV18 variants with the histological and cytological grades of the cervical specimens. The summary of the analysis is shown in the Supplementary Materials Tables S5–S8. For the HPV16 E6 variants, we found no association between any variant and the histological grade. A significant association with the cytological grade was detected only for the R10I mutation

($\chi^2 = 6.7$, p = 0.048). For the HPV16 E7 gene variants, no association was found between the cytological grade and gene mutation. However, a significant association was detected between the histological grade and the L28F ($\chi^2 = 35$, p < 0.0001) and T795C ($\chi^2 = 10.9$, p = 0.027) mutations. For both E6 and E7 gene mutations, there was no significant association between the gene variant and squamous cell carcinoma (SCC) grade. For HPV16 L1 gene variants, a significant association with the cytological grade was detected only for the V194I mutation ($\chi^2 = 13.7$, p = 0.001). However, significant associations were detected between the histological grade and the T266A ($\chi^2 = 11.1$, p = 0.025), G6994A ($\chi^2 = 10.9$, p = 0.028) and L474F ($\chi^2 = 10.9$, p = 0.028) mutations. Three mutations were significantly associated with the SCC grade: T353P ($\chi^2 = 7.8$, p = 0.02), G6721A ($\chi^2 = 7.8$, p = 0.02), and C6854T ($\chi^2 = 7.8$, p = 0.02).

Due to the sample size, detected mutations in the E6, E7, and L1 regions of the HPV18 variants were not evaluated by the histological, cytological, or SCC grade.

3.5. Viral Load by HPV16 or HPV18 Variant

We used Wilcoxon rank-sum tests to evaluate whether specific variants were associated with the viral load. For the HPV16 E6 gene, five mutations (G145T, Q14H, T286A, A289G, and H78Y) were associated with higher viral loads (Table 7). For the HPV16 E7 gene, two mutations, T789C and T795G, were significantly associated with higher viral loads. For the HPV16 L1 gene, 12 mutations (G5698A, H76Y, T5911C, T266A, T353P, G6721A, C6854T, C6865T, C6970T, G6994A, and L474F) were significantly associated with higher viral loads.

Viral Load, Mean (SD), Viral Copies/µL) Wilcoxon Rank-Sum Nucleic Acid/Amino Gene Test *p*-Value **Acid Mutation** Present (Mutation) Absent (Wild Type) 145 G -> T 7770.8 (16,798.7) 2026.4 (4854.8) (0.017)* 5259.7 (10,039.4) 145 G-> T/Q14H 4992.7 (14103.9) (0.035)* 7510 (16,528.1) (0.0162) * 286 T -> A 2085.4 (4960.6) E6 289 A -> G 7510 (16,528.1) 2085.4 (4960.6) (0.0162) * 335 C -> T/H78Y 7510 (16,528.1) 2085.4 (4960.6) $(0.0162)^{3}$ 789 T -> C 7509.7 (16,528) 2085.8 (4960.4) (0.0174) * E7 795 T -> G 7509.7 (16,528) 2085.5 (4960.4) (0.0174) * 5698 G -> A 7242.1 (16,281) 2184.3 (5060.8) (0.0448) * 5864 C -> T/H76Y 7242.1 (16,281) 2184.3 (5060.8) (0.0448) * 5911 T -> C 7242.1 (16,281) 2184.3 (5060.8) (0.0448) * (0.0449) * 6434 A -> G/T266A 5754.3 (14,149.7) 2423.1 (5574.9) 6559 C -> T (0.0224) * 7389.1 (17,003.8) 2663.4 (5761.8) 6695 A -> C/T353P 14,596.3 (22,832) 1986.3 (4715.8) (0.0004) * L1 1986.3 (4715.8) 6721 G -> A 14,596.3 (22,832) (0.0004) * 6854 C -> T 14,596.3 (22,832) 1986.3 (4715.8) (0.0004) * 6865 C -> T 14,596.3 (22,832) 1986.3 (4715.8) (0.0004) * 6970 C -> T 15,682.7 (23,619) 2003.7 (4625) (0.0007)* 6994 G -> A 17,206 (24,321.1) 1963.8 (4598) (0.0003) * 7060 G -> T/L474F 17,206 (24,321.1) 1963.8 (4598) (0.0003) *

Table 7. HPV16 gene variants significantly associated with higher viral loads.

* Indicates statistical significance (p < 0.05).

For HPV18, due to the small sample size, we could not perform the statistical test.

3.6. Phylogenetic Trees

For HPV16, 43 of 54 (79.6%) specimens were successfully classified, 38 specimens based on the E6 gene and 5 based on the complete L1 gene. The results of the lineage analysis showed that 46.5% of the specimens belonged to lineage A, 9.3% belonged to lineage B, 11.6% belonged to lineage C, and 32.6% belonged to lineage D. For lineage A, 41.9% of the specimens were classified as A1, which is a European lineage, and 4.7% were classified as A4, which is an Asian lineage. For lineage B, 4.7% of the specimens were classified as both B1 and B2, both of which are African lineages. For lineage C, 11.6%

belonged to an African lineage. For lineage D, 20.9% were classified as D1, which is a North American lineage, and 11.6% were classified as D3, an Asian American sublineage. The phylogenetic trees are illustrated in Figure 3. No specimens belonging to A2, A3, A5, or D2 sublineages were detected.



Figure 3. HPV16 phylogenetic trees. (**A**) Maximum likelihood analysis of the E6 nucleotide sequence. Specimens that did not cluster to any reference sequence were excluded. (**B**) Excluded specimens with complete L1 genes were reanalyzed by using maximum likelihood of the L1 nucleotide sequence. Numbers on the nodes indicate bootstrap values.

For HPV18, 5 of 6 (83.3%) specimens were successfully classified. Only lineage A was detected, with 60% belonging to the A4 sublineage and 40% belonging to the A5 sublineage, both of which are European lineages. The phylogenetic tree is shown in Figure 4.



Figure 4. HPV18 phylogenetic tree. Maximum likelihood analysis of the E6 nucleotide sequence. Numbers on the nodes indicate bootstrap values. HSIL represents high-grade squamous intraepithe-lial lesion; SCC, squamous cell carcinoma.

3.7. Association of Lineage with Patient Demographic Characteristics

There was no significant association between patient demographic variables and HPV16 lineages or HPV18 lineages.

4. Discussion

The role of HR-HPV intra-type variants and genetic mutations in the development of malignant neoplasms is still under investigation, and the results of many studies assessing the clinical and prognostic significance of HPV detection in cervical cancer screening are contentious [20,26]. E6 and E7 gene mutations may influence malignant transformation processes, including cell cycle arrest and the absence of apoptosis caused by E6 and hyperproliferation caused by E7 [12,20,26,41,42]. L1 gene mutations may change the host's cellular immune response and may play a role in vaccine effectiveness because the L1 region is typically used to generate vaccines [11,12,20]. Therefore, identifying and understanding HR-HPV genetic mutations are essential for studying molecular pathogenesis and for developing molecular diagnostics, vaccines, and other therapeutic approaches aimed at

controlling or eliminating virus-induced diseases [20,26]. The present study was, to our knowledge, the first to investigate mutations in the E6, E7, and L1 genes among Saudi women who tested positive for HPV16 or HPV18 infection.

We found 12 nonsynonymous mutations in the HPV16 E6 gene: R10I, Q14D, Q14H, H24N, D25Y, D25E, I27R, V31A, A61G, D64E, H78Y, and L83V. Previous studies have demonstrated that R10I and Q14D mutations may affect p53 affinity and significantly alter its degradation rate [19,43,44]. D25E also plays a major role in p53 degradation [45,46]. Our results indicated that R10I was the only mutation that was significantly associated with the cytological grade ($\chi^2 = 6.7$, p = 0.048). A recent study evaluating E6 structural changes using an in silico approach found that the Q14H mutation is located at the binding interface of the E6-p53 core and thus may play an important role in p53 binding. Those authors also suggested that an I27R mutation may lead to changes in molecular interactions [47]. Although H24N, D25Y, V31A, A61G, and D64E mutations have been detected in previous studies, their structural and functional effects require further investigation [43,48–51].

The mutations most frequently detected in our study were H78Y and L83V, with a rate of 57.4% each. There are fewer studies assessing the role of H78Y compared with those for L83V. It has been suggested that the region where amino acid 83 is located could be responsible for the greater oncogenic potential of HR-HPVs [52]. The authors of that study conducted functional analyses of the mutations found in 10 premalignant and malignant tissues. The biochemical activities of these mutations were evaluated using both in vivo and in vitro assays. Their comparison of the L83V mutation with the E6 gene prototype indicated that L83V displayed moderate to high activities in most functional assays, more efficient binding to the E6-binding protein, and more efficient degradation of the apoptosis regulator Bax compared with the other variants [19]. These results could explain our finding of high odds for detecting the L83V amino acid mutation in cervical cancer specimens (OR: 10.1; 95% CI: 1.1–89.4).

The genetic basis of HPV carcinogenicity is geographically different among variants and mutations [53–55]. It is well known that both L83V and D25E mutations in the E6 gene are associated with an elevated risk of cervical carcinoma [43]. However, in the present study, the excess risk associated with cervical cancer and HPV mutation was significant only for the L83V mutation. Our results are similar to those of a hospital-based study that was conducted in Korea and reported a significant association of L83V with cervical cancer [44]. By contrast, previous studies that were performed in China did not find an association between L83V and cervical cancer [40,56]. These inconsistent results assessing the E6-L83V role in the development of cervical cancer may be explained in part by (1) the geographic origins or races of the patients, (2) the small number of cases investigated, and (3) the differences in experimental design [26,38,42]. More studies investigating the functional implications of the L83V mutation are required [57].

The L28F and N29S mutations were the only nonsynonymous E7 mutations detected in the present study, and the N29S mutation was the most prevalent mutation detected. Consistent with our findings, various studies evaluating E7 mutations have reported a higher prevalence of N29S compared with other mutations [44,54,56]. Previous studies have also reported that the E7 protein linked to cervical cancer is more conserved than the E6 protein [43]. In our study, the number of E6 mutations was higher than the number of E7 mutations by 113%. Moreover, the E7 mutation L28F was associated with a significantly lower risk of cervical cancer. It is known that L28F affects DNA synthesis, pRb binding, and nuclear localization signals. Considering these known functional implications, the association between L28F and the decreased risk of both cervical dysplasia and invasive cancer should be further investigated [44].

Mutations that occur in surface loops BC, DE, EF, FG, and HI or in sites between them could affect virus infection efficiency or alter virus antigenicity [11,20,58]. A Dutch study of genetic variation in the L1 proteins of HPV16 and HPV18 among young adolescents found that 78% of L1 amino acid mutations were located in the encoding region of the alpha helix, beta sheets, surface loops, or connecting loops [58–60]. Of the 16 amino acid

mutations detected in the present study (K53N, K53T, H76Y, T176N, N181S, N181T, V194I, T266A, S282P, S351A, T353P, R365P, T389S, S396P, K443N, and L474F), 75% were located in the surface loops and in the encoding regions of the alpha helix and beta sheet. Structural analyses of the L1 protein in previous studies have shown that T176C, N181S, N181T, and V194I are located in the EF surface loop, T266A and S282P are in the FG surface loop, and S351A and T353P are in the HI surface loop. T389S, S396P, and L474F are in the alpha-helix encoding region, whereas H76Y is located in the region that encodes the beta sheet [57–60].

The most frequently detected mutations in our specimens were T176N and H76Y. In a previous study, pseudovirion (PsV) constructs of 39 L1 variants were generated to analyze the effectiveness of these variants [60]. Owing to the absence of any detectible proteins or PsV particles when investigated under the electron microscope, T176N presumably reduced the expression level or stability of L1 or affected PsV assembly. A recent study reported that T176N was a part of linear B-cell epitope regions, whereas H76Y and T266A were part of known T-cell epitopes. In our study, T266A was associated with the highest risk of cervical cancer (OR: 4.5; 95% CI: 0.92–21.9). A previous study using the Site-Directed Mutator tool to predict the effects of variants on protein stability found that T266A may result in the loss of two hydrogen bonds with phenylalanine in position 360 (F360) and lysine in the next position (K361) of the adjacent chain [61].

Previous studies have reported that HPV16 variants and mutations associated with the occurrence of cervical cancer may differ by geographic region or ethnicity [39,44,47,55,62]. One study in India found that the N181T/I, T353P, and T389S mutations are associated with the histological grade [61]. We found that the presence of V194I was significantly associated with the histological grade, whereas the presence of T266A, L474F, and G6994A was significantly associated with the cytological grade. Thus, further investigations assessing the oncogenicity of variants/mutations in larger populations are needed that take into consideration geography and ethnicity effects [44].

We identified four amino acid mutations of the HPV16 L1 gene: K53T, K53N, R365P, and K443N. To the best of our knowledge, none of these four mutations have been reported. A recent study in India reported a mutation in position 443 of this gene for the first time in Indian isolates. However, in that study, in HPV-18, lysine in position 443 was replaced by glutamine, whereas in our study, the replacement was by asparagine, which has not been previously reported. Those authors hypothesized that lysine in that position may stabilize the structure of L1 [61].

No amino acid mutations were detected in the E6 gene. By contrast, there were two nonsynonymous mutations in the E7 gene, L8V and M61L, with L8V being more prevalent. Previous studies have reported L8V and M61L mutations in the E7 gene; however, the biological role of these mutations will require additional investigation [43,63]. In general, genetic studies assessing HPV18 are insufficient, making comparisons difficult [64].

Viral antigenicity and infection efficiency may be affected by amino acid mutations in the L1 gene [11]. In the present study, there were six amino acid mutations: R25Q, L64M, P91R, T149N, P344R, and P399R. All of these mutations have been previously reported, although insufficient studies are available exploring their roles and effects [11,24,42,65]. The amino acid mutations most frequently detected in the present study were R25Q, P91R, and T149N. The authors of a previous study suggested that R25Q and P91R may affect immune responses to the HPV18 capsid protein [24]. Amino acid mutations at position 149 have been detected in many HPV genotypes, and these mutations may decrease the probability of other mutation effects on L1 gene function [65].

We also found that 19 mutations of HPV16 (5 in E6, 2 in E7, and 12 in L1) were significantly associated with higher viral loads. By contrast, none of the HPV18 mutations were significantly associated with higher viral loads, but this finding may be due to the small number of specimens assessed in the present study. A previous study investigated the same specimens we used in our study but categorized cytological grades as normal and abnormal. Their logistic regression model showed a significant association between a higher risk of cervical abnormalities and high viral loads for both HPV16 and HPV18 [33]. We used three categories of cytological grades: negative for intraepithelial lesions, high-grade squamous intraepithelial lesions, and invasive carcinoma. Consistent with the results of that previous study, we found that higher viral loads were mainly observed in invasive carcinoma, with lower viral loads detected mainly in cases negative for intraepithelial lesions and cases with high-grade squamous intraepithelial lesions. However, a cross-sectional study conducted in China reported that the HPV16 viral load increased significantly with the increased invasiveness of the cervical lesion grade and was predominant in cervical cancer, whereas the HPV18 viral load showed no significant increase, although it was slightly higher in cervical cancer cases [37]. Another cross-sectional study performed in India found no association between the HPV18 viral load and cervical lesion grade but did detect such an association for the HPV16 viral load [66]. Controlling the factors that may affect the results, such as the method of virus detection and the tissue classification system, are required to be similar to enable an accurate comparison [37]. Further studies investigating the associations between mutations, viral loads, and cervical lesion grades are needed in Saudi Arabia, as well as globally, because the studies in this region are limited.

In Southeast Asia, Asian sublineages are the most prevalent variants, whereas European sublineages predominate in all other geographic regions except Africa [54]. According to the phylogenetic trees generated from our study, for HPV16 and HPV18, most of the sequences were clustered within the European sublineages.

A study by Sait et al. conducted in Saudi Arabia based on sequence analyses of the LI gene reported that 62.5% of HPV16 specimens belonged to lineage A, 50% of which were from European sublineages, and no specimens belonged to sublineages A3, A5, B2, D1, or D3 [31]. For HPV18, 100% of the specimens belonged to lineage A, with 71.4% of them from European sublineages, and no specimens belonged to sublineage A2 or lineages B, C, or D. Their results are consistent with ours, although we constructed phylogenetic trees based on the E6 and L1 genes separately. The most frequently detected lineage in our HPV16 and HPV18 specimens was lineage A, which represented 46% and 100% of the population, respectively. For HPV16, 41.9% of the specimens belonged to European sublineages, and no specimens belonged to the A3, A5, or D2 sublineages. For HPV18, all specimens belonged to European sublineages, and no specimens belonged to the A1 or A2 sublineages or to the B, C, or D lineages.

The distribution of the mutations in the HPV16 and HPV18 lineages appeared to be geographically related. The L83V mutation detected in the European variant of the HPV16 E6 gene was the most prevalent in Central and South America (52%) and Europe (44%). However, the D25E mutation in the Asian variant was the most prevalent in Southeast Asia (26%), in contrast to Europe (2%), Central and South America (0%), and Africa (0%) [55]. In our study, the prevalence of L83V was 57.4%, and the prevalence of D25E was 4.3%.

In addition to this potential geographic effect, the distribution of the mutations appeared to be associated with the race of the women living in the same geographical region. Of 1025 women who were included in a race-related variant distribution study [67–72], 70% were White, 24.5% were African American, 3.5% were Asian/Pacific Islander, and 1.8% were American Indian/Alaskan living in the United States. The study found that 162 African American women were positive for HPV16, of which 43 isolates (26.5%) belonged to the African variant, whereas for 584 White women who were infected with HPV16, 25 isolates (4.3%) belonged to this variant. The statistical analysis of the distribution of HPV16 variants among women who were African American indicated a significant difference from that observed among White women (p < 0.001) and among American Indian/Alaskan women or among Asian/Pacific Islander women (p < 0.001). For HPV18 specimens, the African variant was predominant in African American women. The difference in the overall distribution of HPV18 variants between African American and White women was statistically significant (p < 0.001) [72].

Our study was conducted in Riyadh, whereas the study by Sait et al. [31] was conducted in Jeddah, where the ethnic diversity is known to be greater than that in Riyadh. This diversity may be another reason, in addition to the targeted genes and different protocols, for the reported differences in some sublineage distributions. However, the conclusions from their study and ours—that the most prevalent HPV16 and HPV18 lineages in Saudi Arabia were European—are consistent.

A major challenge for preserving genetic resources is the high-quality long-term storage of DNA samples [68]. The samples enrolled in this study were collected between 2017 and 2019 and had been used in previously published studies. Multiple repetitive freeze–thawing cycles may affect DNA quality. When we used housekeeping gene assays to assess the DNA quality, we indeed found seven specimens that were negative for the gene and were thus excluded from the present study. In addition, 4 specimen vials were empty, and 12 specimens had insufficient DNA for use in sequencing all of the genes of interest.

The gold-standard technology for verifying PCR results and analyzing nucleic acid sequences is Sanger sequencing [68,69]. The automated technology can support the analysis of nucleic acid sequences up to 800–1000 bp [68]. In the present study, we used an automated DNA analyzer to sequence the targeted genes. All of the nucleic acid sequences ranged from 512 bp to 966 bp. However, the low quality of the first 15–40 bp of the sequences was a major limitation of this approach. These unreadable areas are known to be primer binding sites and they often contain indels. Thus, areas must be trimmed, with only high-quality target nucleic acid sequences used for sequence assembly [68]. Therefore, in-house primers were designed to flank the regions of interest of \geq 15 bp. Another major limitation in our study is our use of genomic DNA from formalin-fixed paraffin-embedded specimens, which partially degraded; also, a high number of PCR cycles were performed, which could cause artifacts. However, we have sequenced each sample two times to overcome this limitation.

Shotgun sequencing is a well-established approach [69]. In this technique, DNA is typically randomly fragmented into smaller pieces, sequenced, and assembled as a contig [69,70]. The contig is an assemblage of overlapping sequence reads [71]. However, in our study, the L1 genes of both HPV16 and HPV18 were not fragmented randomly. We used target-specific primer sets to fragment the L1 genes by binding to specific regions in three fragments. The sequence reads of the fragments were overlapped to eliminate the possibility of having unreadable areas and to guarantee the whole gene was sequenced. However, we were unable to obtain all three fragments of the L1 gene from all specimens. Unlike the E6 and E7 genes, the L1 gene can be lost during HPV genome integration during the progression from a low-grade lesion to cancer [36,72]. A previous study compared the performance of MY09/11 primers designed to target the conserved region of the L1 gene with type-specific primers, mainly directed at E6 and E7 genes, for HPV detection. The study found that the MY09/11 primers failed to detect 409 (83.3%) of the 491 HR-HPV-positive specimens that were detected with the type-specific primers. The 409 false-negative specimens were followed up, and it was found that 104 (25.4%) were classified as CIN2+ (moderate dysplasia or higher) [36]. In our study, 82.9% of the HPV16 specimens and 60% of the HPV18 specimens were classified as CIN3+ (severe dysplasia or higher), which may explain our inability to obtain the complete L1 gene from all specimens.

In conclusion, this retrospective cohort study contributes to the HPV epidemiology knowledge for a Saudi population. We provide data on the HPV16 and HPV18 lineages and on mutations in the E6, E7, and L1 genes. Our findings confirmed the predominance of the European lineage in HPV16 and HPV18 specimens in Saudi Arabia. We observed that the number of E6 mutations was 113% higher than E7 mutations in HPV16. We also provided evidence that the L83V mutation in the E6 gene of HPV16 in this Saudi population carried sufficient oncogenic potential to be responsible for the progression to cervical cancer. By contrast, the L28F mutation in the E7 gene of HPV16 was associated with a significantly low risk of cervical cancer. In addition, specific mutations in HPV16 and HPV18 were associated with the risk of cancer, cancer progression, viral load, and age, suggesting that these mutations represented increased oncogenic activity. We also identified four novel mutations, K53T, K53N, R365P, and K443N, in the L1 gene of HPV16. Although this study offers data essential for informing the design and development of effective molecular diagnostic tests and vaccination strategies for the Saudi population,

further studies assessing the associations between mutations, viral loads, and cervical lesion grades are needed in Saudi Arabia as well as globally.

Supplementary Materials: The following supporting information can be downloaded at: https: //www.mdpi.com/article/10.3390/v15010109/s1. Table S1: Mutations detected in the E6, E7, and L1 genes of HPV16 in the study population. Table S2. Cervical Cancer association with HPV16 E6 mutations observed in the Saudi population. Table S3. Cervical Cancer association with HPV16 E7 mutations observed in the Saudi population. Table S4. Cervical Cancer association with HPV16 L1 mutations observed in the Saudi population. Table S5. HPV 16 E6 mutations distributed by histology, cytology and SCC grades. Table S6. HPV 16 E7 mutations distributed by histology, cytology and SCC grades. Table S7. HPV 16 L1 mutations distributed by histology, cytology and SCC grades.

Author Contributions: Conceptualization, F.A. (Fatimah Alhamlan) and D.O.; methodology, M.A.; validation, M.A.; formal analysis, D.O.; investigation, F.A. (Faten Alhatheq) and B.A.; resources, F.A. (Faten Alhatheq); writing—original draft preparation, M.A.; writing—review and editing, A.A.; supervision, F.A. (Fatimah Alhamlan) and A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by King Faisal Specialist Hospital and Research Centre, funding number: RAC 1005-033.

Institutional Review Board Statement: This retrospective cohort study was approved by the Research Advisory Council (Ethics Committee) at King Faisal Specialist Hospital and Research Centre (approval number: RAC 1005-033).

Informed Consent Statement: Patient consent was waived due to the type of patient samples. The cervical biopsies were archived, and the collected patient data were coded and thus deidentified.

Data Availability Statement: The data and codes used in this study are available upon request.

Acknowledgments: The authors thank the King Faisal Specialist Hospital and Research Centre for supporting the laboratory work and reagents and thank the sequencing core facility at the Research Centre for providing Sanger sequencing.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Zheng, Z.-M.; Baker, C.C. Papillomavirus genome structure, expression, and post-transcriptional regulation. *Front. Biosci.* 2006, 11, 2286–2302. [CrossRef]
- 2. Hausen, H. Zur Human papillomavirus & cervical cancer. Indian J. Med. Res. 2009, 130, 209. [PubMed]
- 3. Man, S. Human cellular immune responses against human papillomaviruses in cervical neoplasia. *Expert Rev. Mol. Med.* **1998**, *1*, 1–19. [CrossRef] [PubMed]
- 4. Harari, A.; Chen, Z.; Burk, R.D. Human papillomavirus genomics: Past, present and future. *Curr. Probl. Dermatol.* **2014**, *45*, 1–18. [CrossRef] [PubMed]
- Münger, K.; Baldwin, A.; Edwards, K.M.; Mu, K.; Hayakawa, H.; Nguyen, C.L.; Owens, M.; Grace, M. Mechanisms of Human Oncogenesis MINIREVIEW Mechanisms of Human Papillomavirus-Induced Oncogenesis. *Society* 2004, 78, 11451–11460. [CrossRef]
- 6. Burd, E.M.; Dean, C.L. Human Papilloma Virus. Microbiol. Spectr. 2014, 2, 1–16. [CrossRef]
- 7. Alsbeih, G.; Ahmed, R.; Al-Harbi, N.; Venturina, L.A.; Tulbah, A.; Balaraj, K. Prevalence and genotypes' distribution of human papillomavirus in invasive cervical cancer in Saudi Arabia. *Gynecol. Oncol.* **2011**, *121*, 522–526. [CrossRef]
- Estêvão, D.; Costa, N.R.; Gil da Costa, R.M.; Medeiros, R. Hallmarks of HPV carcinogenesis: The role of E6, E7 and E5 oncoproteins in cellular malignancy. *Biochim. Biophys. Acta-Gene Regul. Mech.* 2019, 1862, 153–162. [CrossRef]
- 9. World Health Organization. Cervical Cancer. Available online: https://www.who.int/news-room/fact-sheets/detail/cervical-cancer (accessed on 2 February 2021).
- Bruni, L.; Albero, G.; Serrano, B.; Mena, M.; Collado, J.J.; Gómez, D.; Muñoz, J.; Bosch, F.X.; de Sanjosé, S. Human Papillomavirus and Related Diseases in Saudi Arabia. 2021. Available online: https://hpvcentre.net/statistics/reports/SAU.pdf?t=1664946162590 (accessed on 2 February 2021).
- 11. Frati, E.; Bianchi, S.; Colzani, D.; Zappa, A.; Orlando, G.; Tanzi, E. Genetic variability in the major capsid L1 protein of human papillomavirus type 16 (HPV-16) and 18 (HPV-18). *Infect. Genet. Evol.* **2011**, *11*, 2119–2124. [CrossRef]

- Arroyo, S.L.; Basaras, M.; Arrese, E.; Hernáez, S.; Andía, D.; Esteban, V.; Garcia-etxebarria, K.; Jugo, B.M.; Cisterna, R. Human Papillomavirus (HPV) genotype 18 variants in patients with clinical manifestations of HPV related infections in Bilbao, Spain. *Virol. J.* 2012, *9*, 258. [CrossRef]
- 13. Bian, M.L.; Cheng, J.Y.; Li, M.; Xiao, C.; Jun, L.; Ying, C.; Xi, C. Evaluation of the detection of 14 high-risk human papillomaviruses with HPV 16 and HPV 18 genotyping for cervical cancer screening. *Exp. Ther. Med.* **2013**, *6*, 1332–1336. [CrossRef] [PubMed]
- 14. Gradíssimo Oliveira, A.; Delgado, C.; Verdasca, N.; Pista, Â. Prognostic value of human papillomavirus types 16 and 18 DNA physical status in cervical intraepithelial neoplasia. *Clin. Microbiol. Infect.* **2013**, *19*, 447–450. [CrossRef] [PubMed]
- Al-Badawi, I.A.; Al-Suwaine, A.; Al-Aker, M.; Asaad, L.; Alaidan, A.; Tulbah, A.; Bohol, M.F.; Munkarah, A.R. Detection and genotyping of human papilloma virus in cervical cancer specimens from Saudi patients. *Int. J. Gynecol. Cancer* 2011, 21, 907–910. [CrossRef]
- Alhamlan, F.S.; Obeid, D.A.; Khayat, H.H.; Tulbah, A.M.; Al-Badawi, I.A.; Al-Muammar, T.A.; Almutairi, A.N.; Almatrrouk, S.A.; Alfageeh, M.B.; Bakhrebah, M.A.; et al. Prognostic Impact of Human Papillomavirus Infection on Cervical Dysplasia, Cancer, and Patient Survival in Saudi Arabia: A 10-Year Retrospective Analysis. *medRxiv* 2020. [CrossRef] [PubMed]
- Leinonen, M. Prevalence of HPV Infection and Use of HPV Test in Cervical Cancer Screening: Randomised Evaluation within the Organised Cervical Cancer Screening in Finland. University of Helsinki. 2013. Available online: https://integro.ciniba.edu.pl/integro/192906311336/leinonen-maarit/prevalence-of-hpv-infection-and-use-of-hpv-test-in-cervical-cancer-screening-randomised-evaluation?bibFilter=19 (accessed on 2 April 2021).
- 18. Sun, Z.; Lu, Z.; Liu, J.; Wang, G.; Zhou, W.; Yang, L.; Liu, C.; Wang, B. Genetic variations of E6 and long control region of human papillomavirus type 16 from patients with cervical lesion in Liaoning, China. *BMC Cancer* **2013**, *13*, 459. [CrossRef] [PubMed]
- Lichtig, H.; Algrisi, M.; Botzer, L.E.; Abadi, T.; Verbitzky, Y.; Jackman, A.; Tommasino, M.; Zehbe, I.; Sherman, L. HPV16 E6 natural variants exhibit different activities in functional assays relevant to the carcinogenic potential of E6. *Virology* 2006, 350, 216–227. [CrossRef] [PubMed]
- Frati, E.R.; Bianchi, S.; Amendola, A.; Colzani, D.; Petrelli, F.; Zehender, G.; Tanzi, E. Genetic characterization of variants of HPV-16, HPV-18 and HPV-52 circulating in Italy among general and high-risk populations. *Mol. Med. Rep.* 2020, 21, 894–902. [CrossRef] [PubMed]
- 21. Castellsagué, X. Natural history and epidemiology of HPV infection and cervical cancer. Gynecol. Oncol. 2008, 110, 4–7. [CrossRef]
- 22. Ghedira, R.; Mahfoudh, W.; Hadhri, S.; Gabbouj, S.; Bouanene, I.; Khairi, H.; Chaieb, A.; Khelifa, R.; Bouaouina, N.; Remadi, S.; et al. Human papillomavirus genotypes and HPV-16 variants distribution among Tunisian women with normal cytology and squamous intraepithelial lesions. *Infect. Agent. Cancer* **2016**, *11*, 1–10. [CrossRef]
- Alhamlan, F.S.; Khayat, H.H.; Ramisetty-Mikler, S.; Al-Muammar, T.A.; Tulbah, A.M.; Al-Badawi, I.A.; Kurdi, W.I.; Tulbah, M.I.; Alkhenizan, A.A.; Hussain, A.N.; et al. Sociodemographic characteristics and sexual behavior as risk factors for human papillomavirus infection in Saudi Arabia. *Int. J. Infect. Dis.* 2016, 46, 94–99. [CrossRef]
- 24. Shen, M.; Ding, X.; Li, T.; Chen, G.; Zhou, X. Sequence Variation Analysis of HPV-18 Isolates in Southwest China. *PLoS ONE* 2013, *8*, 56614. [CrossRef] [PubMed]
- Villa, L.L.; Sichero, L.; Rahal, P.; Caballero, O.; Ferenczy, A.; Rohan, T.; Franco, E.L. Molecular variants of human papillomavirus types 16 and 18 preferentially associated with cervical neoplasia. *J. Gen. Virol.* 2000, *81*, 2959–2968. [CrossRef]
- Qiu, A.D.; Wu, E.Q.; Yu, X.H.; Jiang, C.L.; Jin, Y.H.; Wu, Y.G.; Chen, Y.; Chen, Y.; Shan, Y.M.; Zhang, G.N.; et al. HPV prevalence, E6 sequence variation and physical state of HPV16 isolates from patients with cervical cancer in Sichuan, China. *Gynecol. Oncol.* 2007, 104, 77–85. [CrossRef]
- Narisawa-Saito, M.; Kiyono, T. Basic mechanisms of high-risk human papillomavirus-induced carcinogenesis: Roles of E6 and E7 proteins. *Cancer Sci.* 2007, *98*, 1505–1511. [CrossRef] [PubMed]
- 28. Buck, C.B.; Day, P.M.; Trusb, B.L. The Papillomavirus Major Capsid Protein L1. Virology 2013, 445, 169–174. [CrossRef] [PubMed]
- Alhamlan, F.S.; Al-Qahtani, A.A.; Al-Ahdal, M.N. Current studies on human papillomavirus in Saudi Arabia. J. Infect. Dev. Ctries. 2015, 9, 571–576. [CrossRef] [PubMed]
- 30. Al-Shabanah, O.A.; Hafez, M.M.; Hassan, Z.K.; Sayed-Ahmed, M.M.; Abozeed, W.N.; Al-Rejaie, S.S.; Alsheikh, A.A. Human papillomavirus genotyping and integration in ovarian cancer Saudi patients. *Virol. J.* **2013**, *10*, 343. [CrossRef] [PubMed]
- Sait, K.; Turki, R.; Abuzenadah, A.M.; Jiffiri, O.H.; Bohmaidah, A.; Sohrab, S.S. Genetic diversity and phylogenetic analysis of HPV 16 & 18 variants isolated from cervical specimens of women in Saudi Arabia. Saudi J. Biol. Sci. 2019, 26, 317–324. [CrossRef]
- 32. Turki, R.; Sait, K.; Anfinan, N.; Sohrab, S.S.; Abuzenadah, A.M. Prevalence of human papillomavirus in women from Saudi Arabia. *Asian Pacific J. Cancer Prev.* **2013**, *14*, 3177–3318. [CrossRef] [PubMed]
- Obeid, D.A.; Almatrrouk, S.A.; Khayat, H.H.; Al-Muammer, T.A.; Tulbah, A.M.; Albadawi, I.A.; Al-Ahdal, M.N.; Alhamlan, F.S. Human papillomavirus type 16 and 18 viral loads as predictors associated with abnormal cervical cytology among women in Saudi Arabia. *Heliyon* 2020, 6, e03473. [CrossRef]
- Alhamlan, F.S.; Khayat, H.H.; Obeid, D.A.; Tulba, A.M.; Baduwais, T.S.; Alfageeh, M.B.; Al-Ahdal, M.N. Clinical comparison of two human papillomavirus detection assays: GenoFlow and reverse line blot. J. Infect. Dev. Ctries. 2020, 14, 97–103. [CrossRef]
- 35. Apgar, B.S.; Zoschnick, L.; Wright, T.C. The 2001 Bethesda System Terminology. Am. Fam. Physician 2003, 68, 1992–1998. [PubMed]
- Depuydt, C.E.; Boulet, G.A.V.; Horvath, C.A.J.; Benoy, I.H.; Vereecken, A.J.; Bogers, J.J. Comparison of MY09/11 consensus PCR and type-specific PCRs in the detection of oncogenic HPV types. J. Cell. Mol. Med. 2007, 11, 881–891. [CrossRef] [PubMed]

- Wu, Z.; Qin, Y.; Yu, L.; Lin, C.; Wang, H.; Cui, J.; Liu, B.; Liao, Y.; Warren, D.; Zhang, X.; et al. Association between human papillomavirus (HPV) 16, HPV18, and other HR-HPV viral load and the histological classification of cervical lesions: Results from a large-scale cross-sectional study. *J. Med. Virol.* 2017, *89*, 535–541. [CrossRef]
- 38. Trifinopoulos, J.; Nguyen, L.T.; von Haeseler, A.; Minh, B.Q. W-IQ-TREE: A fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* **2016**, *44*, W232–W235. [CrossRef] [PubMed]
- Pérez, S.; Cid, A.; Iñarrea, A.; Pato, M.; Lamas, M.J.; Couso, B.; Gil, M.; Álvarez, M.J.; Rey, S.; López-Miragaya, I.; et al. Prevalence of HPV 16 and HPV 18 lineages in Galicia, Spain. *PLoS ONE* 2014, 9, 104678. [CrossRef] [PubMed]
- 40. Zhou, Z.; Yang, H.; Yang, L.; Yao, Y.; Dai, S.; Shi, L.; Li, C.; Yang, L.; Yan, Z.; Yao, Y. Human papillomavirus type 16 E6 and E7 gene variations associated with cervical cancer in a Han Chinese population. *Infect. Genet. Evol.* **2019**, *73*, 13–20. [CrossRef] [PubMed]
- 41. Salavatiha, Z.; Shoja, Z.; Heydari, N.; Marashi, S.M.; Younesi, S.; Nozarian, Z.; Jalilvand, S. Lineage analysis of human papillomavirus type 18 based on E6 region in cervical samples of Iranian women. *J. Med. Virol.* **2020**, *92*, 3815–3820. [CrossRef]
- 42. Xu, H.H.; Zheng, L.Z.; Lin, A.F.; Dong, S.S.; Chai, Z.Y.; Yan, W.H. Human papillomavirus (HPV) 18 genetic variants and cervical cancer risk in Taizhou area, China. *Gene* 2018, 647, 192–197. [CrossRef]
- Zhao, J.; Zhan, Q.; Guo, J.; Liu, M.; Ruan, Y.; Zhu, T.; Han, L.; Li, F. Phylogeny and polymorphism in the E6 and E7 of human papillomavirus: Alpha-9 (HPV16, 31, 33, 52, 58), alpha-5 (HPV51), alpha-6 (HPV53, 66), alpha-7 (HPV18, 39, 59, 68) and alpha-10 (HPV6, 44) in women from Shanghai. *Infect. Agent. Cancer* 2019, 14, 38. [CrossRef]
- 44. Park, J.S.; Shin, S.; Kim, E.C.; Kim, J.E.; Kim, Y.B.; Oh, S.; Roh, E.Y.; Yoon, J.H. Association of human papillomavirus type 16 and its genetic variants with cervical lesion in Korea. *Apmis* **2016**, 124, 950–957. [CrossRef] [PubMed]
- Boumba, L.M.A.; Assoumou, S.Z.; Hilali, L.; Mambou, J.V.; Moukassa, D.; Ennaji, M.M. Genetic variability in E6 and E7 oncogenes of human papillomavirus Type 16 from Congolese cervical cancer isolates Cancer centers in low- and middle-income countries. *Infect. Agent. Cancer* 2015, 10, 15. [CrossRef] [PubMed]
- 46. Tan, G.; Duan, M.; Li, Y.; Zhang, N.; Zhang, W.; Li, B.; Qu, P. Distribution of HPV 16 E6 gene variants in screening women and its associations with cervical lesions progression. *Virus Res.* **2019**, *273*, 197740. [CrossRef] [PubMed]
- 47. Hang, D.; Gao, L.; Sun, M.; Liu, Y.; Ke, Y. Functional Effects of Sequence Variations in the E6 and E2 Genes of Human Papillomavirus 16 European and Asian Variants. *J. Med. Virol.* **2014**, *86*, 618–626. [CrossRef] [PubMed]
- Rodríguez-Ruiz, H.A.; Garibay-Cerdenares, O.L.; Illades-Aguiar, B.; Montaño, S.; Jiang, X.; Leyva-Vázquez, M.A. In silico prediction of structural changes in human papillomavirus type 16 (HPV16) E6 oncoprotein and its variants. *BMC Mol. Cell Biol.* 2019, 20, 35. [CrossRef] [PubMed]
- 49. Hang, D.; Yin, Y.; Han, J.; Jiang, J.; Ma, H.; Xie, S.; Feng, X.; Zhang, K.; Hu, Z.; Shen, H.; et al. Analysis of human papillomavirus 16 variants and risk for cervical cancer in Chinese population. *Virology* **2016**, *488*, 156–161. [CrossRef]
- Matsumoto, K.; Yasugi, T.; Nakagawa, S.; Okubo, M.; Hirata, R.; Maeda, H.; Yoshikawa, H.; Taketani, Y. Human papillomavirus type 16 E6 variants and HLA class II alleles among Japanese women with cervical cancer. *Int. J. Cancer* 2003, 106, 919–922. [CrossRef]
- 51. Mutagenesis, S.; Spliced, H.P.V.; Protein, E. Site-Directed Mutagenesis of HPV16 Spliced E6 Protein and Localization Patterns. *Thai J. Toxicol.* **2010**, *25*, 154–161.
- 52. Zehbe, I.; Wilander, E.; Delius, H.; Tommasino, M. Human papillomavirus 16 E6 variants are more prevalent in invasive cervical carcinoma than the prototype. *Cancer Res.* **1998**, *58*, 829–833.
- 53. Pillai, M.R.; Sreevidya, S.; Pollock, B.H.; Jayaprakash, P.G.; Herman, B. Human papillomavirus type 16 E6 and E7 gene variations in Indian cervical cancer. *Gynecol. Oncol.* **2002**, *87*, 268–273. [CrossRef]
- 54. Cao, M.; Chenzhang, Y.; Ding, X.; Zhang, Y.; Jing, Y.; Chen, Z. Genetic variability and lineage phylogeny of human papillomavirus type-16 and -53 based on the E6, E7, and L1 genes in Southwest China. *Gene* **2016**, *592*, 49–59. [CrossRef] [PubMed]
- Cai, H.B.; Chen, C.C.; Ding, X.H. Human papillomavirus type 16 E6 gene variations in Chinese population. *Eur. J. Surg. Oncol.* 2010, 36, 160–163. [CrossRef] [PubMed]
- 56. Yang, L.; Yang, H.; Wu, K.; Shi, X.; Ma, S.; Sun, Q. Prevalence of HPV and Variation of HPV 16/HPV 18 E6/E7 Genes in Cervical Cancer in Women in South West China. *J. Med. Virol.* **2014**, *86*, 1926–1936. [CrossRef] [PubMed]
- 57. He, H.; Li, H.; Fan, P.; Zhu, J.; Pan, Z.; Pan, H.; Wu, D.; Ren, X.; Guo, X.; Li, D.; et al. Variants of human papillomaviruses 16 (HPV16) in Uigur women in Xinjiang, China. *Infect. Agent. Cancer* **2016**, *11*, 14. [CrossRef] [PubMed]
- 58. El Aliani, A.; El Abid, H.; Kassal, Y.; Khyatti, M.; Attaleb, M.; Ennaji, M.M.; El Mzibri, M. HPV16 L1 diversity and its potential impact on the vaccination-induced immunity. *Gene* **2020**, 747, 144682. [CrossRef]
- 59. King, A.J.; Sonsma, J.A.; Vriend, H.J.; Van Der Sande, M.A.B.; Feltkamp, M.C.; Boot, H.J.; Koopmans, M.P.G. Genetic Diversity in the Major Capsid L1 Protein of HPV-16 and HPV-18 in the Netherlands. *PLoS ONE* **2016**, *11*, e0152782. [CrossRef]
- Ning, T.; Wolfe, A.; Nie, J.; Huang, W.; Chen, X.S.; Wang, Y. Naturally Occurring Single Amino Acid Substitution in the L1 Major Capsid Protein of Human Papillomavirus Type 16: Alteration of Susceptibility to Antibody-Mediated Neutralization. *J. Infect. Dis.* 2017, 216, 867–876. [CrossRef]
- Mane, A.; Patil, L.; Limaye, S.; Nirmalkar, A.; Kulkarni-Kale, U. Characterization of major capsid protein (L1) variants of Human papillomavirus type 16 by cervical neoplastic status in Indian women: Phylogenetic and functional analysis. *J. Med. Virol.* 2020, 92, 1303–1308. [CrossRef]
- 62. Cornet, I.; Gheit, T.; Iannacone, M.R.; Vignat, J.; Sylla, B.S.; Del Mistro, A.; Franceschi, S.; Tommasino, M.; Clifford, G.M. HPV16 genetic variation and the development of cervical cancer worldwide. *Br. J. Cancer* **2013**, *108*, 240–244. [CrossRef]

- 63. Bokal, E.V.Č.; Kocjan, B.J.; Poljak, M.; Bogovac, Ž.E.; Jančar, N. Genomic variants of human papillomavirus genotypes 16, 18, and 33 in women with cervical cancer in Slovenia. *J. Obstet. Gynaecol. Res.* **2010**, *36*, 1204–1213. [CrossRef]
- 64. Vidal, J.P.C.B. Genetic Diversity of HPV16 and HPV18 in Brazilian Patients With Invasive Cervical Cancer. *Antivir. Ther.* **2016**, *88*, 1279–1287. [CrossRef] [PubMed]
- Lurchachaiwong, W.; Junyangdikul, P.; Termrungruanglert, W.; Payungporn, S.; Sampatanukul, P.; Tresukosol, D.; Niruthisard, S.; Trivijitsilp, P.; Karalak, A.; Swangvaree, S.; et al. Whole-genome sequence analysis of human papillomavirus type 18 from infected thai women. *Intervirology* 2010, *53*, 161–166. [CrossRef] [PubMed]
- 66. Joharinia, N.; Farhadi, A.; Hosseini, S.Y.; Safaei, A.; Sarvari, J. Association of HPV16 and 18 genomic copies with histological grades of cervical lesions. *VirusDisease* **2019**, *30*, 387–393. [CrossRef] [PubMed]
- 67. Al-Awadhi, R.; Chehadeh, W.; Al-Jassar, W.; Al-Harmi, J.; Al-Saleh, E.; Kapila, K. Viral load of human papillomavirus in women with normal and abnormal cervical cytology in Kuwait. *J. Infect. Dev. Ctries.* **2013**, *7*, 130–136. [CrossRef]
- Knebelsberger, T.; Stöger, I. DNA Extraction, Preservation, and Amplification. In DNA Barcodes: Methods and Protocols, Methods in Molecular Biology; Humana Press: Totowa, NJ, USA, 2012; pp. 311–338. [CrossRef]
- 69. França, L.T.C.; Carrilho, E.; Kist, T.B.L. A review of DNA sequencing techniques. Q. Rev. Biophys. 2002, 35, 169–200. [CrossRef]
- Crossley, B.M.; Bai, J.; Glaser, A.; Maes, R.; Porter, E.; Killian, M.L.; Clement, T.; Toohey-Kurth, K. Guidelines for Sanger sequencing and molecular assay monitoring. *J. Vet. Diagnostic Investig.* 2020, *32*, 767–775. [CrossRef] [PubMed]
- Venter, J.C.; Adams, M.D.; Myers, E.W.; Li, P.W.; Mural, R.J.; Sutton, G.G.; Smith, H.O.; Yandell, M.; Evans, C.A.; Holt, R.A. The sequence of the human genome. *Science* 2001, 291, 1304–1351. [CrossRef]
- 72. Tjalma, W.A.A.; Depuydt, C.E. Cervical cancer screening: Which HPV test should be used-L1 or E6/E7? *Eur. J. Obstet. Gynecol. Reprod. Biol.* **2013**, *170*, 45–46. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.