



## Original Article

# Establishment of a Method for Amplifying the Full Genome of Human Papillomavirus 16 Using Overlapping Extension Polymerase Chain Reaction



Ziyu Chen<sup>1#</sup> , Huiying Li<sup>2#</sup> , Shaobing Zhan<sup>2</sup> , Xiaoguang Zhang<sup>2\*</sup> , Hong Yu<sup>1</sup> and Shuying Li<sup>1\*</sup>

<sup>1</sup>School of Basic Medicine Sciences (Hebei Key Laboratory for Chronic Diseases), North China University of Science and Technology, Tangshan, Hebei, China; <sup>2</sup>National Key Laboratory of Intelligent Tracking and Forecasting for Infectious Diseases, National Institute for Viral Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, China

Received: September 18, 2025 | Revised: December 12, 2025 | Accepted: December 19, 2025 | Published online: December 30, 2025

## Abstract

**Background and objectives:** Human papillomavirus (HPV) is a double-stranded circular DNA virus with a genome of approximately 7–8 kb. This study aimed to establish an overlapping extension polymerase chain reaction method for the amplification of the entire genome of HPV16.

**Methods:** The HPV16 genome was divided into two larger fragments (with lengths of 3.9 kilobases and 5.3 kilobases, respectively), each of which had overlapping regions of more than 500 base pairs. A nested primer (outer primer: Fout/Rout; inner primer: Fin/Rin) was used to amplify each fragment. The key reaction parameters were optimized, including the selection of two highly accurate DNA polymerases; and a series of diluted samples (initial concentration of 2,000 copies/microliter, diluted to 2, 20, 200, and 2,000 copies/microliter) were used for amplification tests to evaluate the sensitivity of this method.

**Results:** This study demonstrated high sensitivity for HPV16 detection, with effective amplification of samples as low as 2 copies/μL. For low-concentration samples (<200 copies/μL), the Thermo Fisher enzyme showed 50% and 75% effective amplification success rates at 2 copies/μL and 20 copies/μL, respectively, while the Vazyme enzyme achieved 0% success at both concentrations. Both enzymes enabled stable amplification of high-concentration samples (≥200 copies/μL). The amplified products matched the theoretical size, and Illumina sequencing confirmed Q30 ≥ 96% and >98% identity with the HPV16 reference sequence (K02718.1).

**Conclusions:** This study provides a highly sensitive and specific method for the full-genome sequence analysis of HPV16, which is applicable to HPV16 full-genome sequencing, variation analysis, and other research.

**Keywords:** Overlapping extension PCR; HPV16; Full genome; Nested PCR; High-fidelity DNA polymerase; Low viral load; Sensitivity assessment; Illumina sequencing.

\*Correspondence to: Xiaoguang Zhang, National Key Laboratory of Intelligent Tracking and Forecasting for Infectious Diseases, National Institute for Viral Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing 102206, China. ORCID: <https://orcid.org/0009-0001-2208-2081>. Tel: +86-010-83546064, E-mail: [zhangxg@ivdc.chinaacdc.cn](mailto:zhangxg@ivdc.chinaacdc.cn); Shuying Li, School of Basic Medicine Sciences (Hebei Key Laboratory for Chronic Diseases), North China University of Science and Technology, Tangshan, Hebei 063210, China. ORCID: <https://orcid.org/0000-0001-9781-2336>. Tel: +86-0315-8805515, E-mail: [lsy5001@sina.com](mailto:lsy5001@sina.com)

<sup>#</sup>These authors contributed equally to this work.

**How to cite this article:** Chen Z, Li H, Zhan S, Zhang X, Yu H, Li S. Establishment of a Method for Amplifying the Full Genome of Human Papillomavirus 16 Using Overlapping Extension Polymerase Chain Reaction. *Oncol Adv* 2025;2025;3(4):e00025. doi: 10.14218/OnA.2025.00025.

## Introduction

Human papillomavirus (HPV) is a double-stranded circular DNA virus with a genome of approximately 7–8 kb.<sup>1,2</sup> To date, researchers have identified over 200 genotypes, among which HPV16 is one of the major high-risk subtypes and is closely associated with the occurrence of various malignant tumors, such as cervical and anal cancer.<sup>3,4</sup> Currently, the commonly used HPV detection methods in clinical practice mainly target partial viral fragments for typing, such as the Cobas 4800 detection system.<sup>5,6</sup> Although these methods enable rapid and accurate identification of the types of HPV, their limitation lies in the inability to provide complete full-genome information. Therefore, current research still relies on other molecular techniques to achieve full-genome amplification

**Table 1.** HPV16 whole-genome sequence amplification primers

Primer	Sequence (5' to 3')	Location	Length	Tm value (°C)	GC content (%)
16.1Fout	ATCATCAAGAACACGTAGAGAAACCC	526–551		59	42
16.1Rout	CGACCCTGTTCCAATTCTAACCC	4,403–4,426	3,902	61	54
16.1Fin	TGTGACTCTACGCTTCGGTTG	742–762		60	52
16.1Rin	GGTAGCCGATGCACGTT	4,266–4,282	3,542	57	59
16.2Fout	CAACATTACTGGCGTGCTTT	3,874–3,893		56	45
16.2Rout	CCATACCCGCTGTCTTCG	1,250–1,267	5,299	57	61
16.2 Fin	CCTATTAATACGTCGCTGCT	3,925–3,945		57	48
16.2 Rin	CAGTAAACAACGCATGTGCT	1,062–1,081	5,060	56	45

In the table, “out” represents primers outside the PCR reaction, and “in” represents primers inside the PCR reaction. GC, guanine–cytosine; HPV, human papillomavirus; PCR, polymerase chain reaction.

and sequencing of HPV16.

However, current methods for amplifying the full HPV16 genome, such as long-fragment polymerase chain reaction (PCR) and rolling circle amplification (RCA), still have several limitations. Long-fragment PCR shows low amplification efficiency when handling low-copy-number templates and often produces non-specific amplification products. In addition, long-fragment PCR requires high-quality template DNA. Degraded or damaged DNA templates can further reduce amplification efficiency or even cause complete amplification failure.<sup>7</sup> RCA can achieve efficient DNA amplification; however, its primer binding sites are relatively limited, and it lacks a sequence-specific selection mechanism, which often leads to non-specific products.<sup>8</sup> Moreover, RCA requires long reaction times and is costly, which limits its use in high-throughput clinical applications.<sup>9</sup>

To address these challenges, we established an overlapping extension PCR method for efficient and specific amplification of the full HPV16 genome. The genome was divided into two large fragments (3.9 kb and 5.3 kb) and amplified using a nested PCR strategy. We systematically evaluated the performance of two high-fidelity DNA polymerases under optimized conditions. The high sensitivity of this method enables its application in samples with low viral loads, providing a reliable approach for full-genome sequencing, variant analysis, molecular evolution studies, and clinical diagnostics of HPV16.

Materials and methods

Nucleic acids extraction and plasmid standard curve equation

HPV16 DNA extraction: All HPV16-positive nucleic acid samples were extracted from cervical exfoliated cells of patients at Peking University First Hospital. The samples were confirmed to be HPV16-positive by clinical HPV genotyping detection. For sample selection, the inclusion criteria were: 1) samples with complete clinical detection records; 2) nucleic acid purity meeting experimental requirements (A260/280 ratio between 1.7 and 1.9, determined by ultraviolet spectrophotometry); and 3) no obvious DNA degradation (verified by agarose gel electrophoresis). The exclusion criteria were: 1) samples with ambiguous HPV genotyping results; 2) nucleic acid samples contaminated during extraction; and 3) samples with insufficient volume for repeated experiments. Plasmid construction: The HPV16 E6 standard fragment was synthesized and ligated into the plasmid vector (pUC). The recombinant plasmid was then transformed into Escherichia coli DH5α

and cultured. Positive clones were selected and confirmed by PCR and sequencing. Plasmid DNA was extracted from the confirmed clones, and its concentration was measured using a nucleic acid analyzer. The DNA copy number was determined to serve as the standard stock solution. Serial dilutions were performed, and real-time quantitative PCR was used to amplify the samples. Data were collected and used to generate a standard curve equation.

Primer design

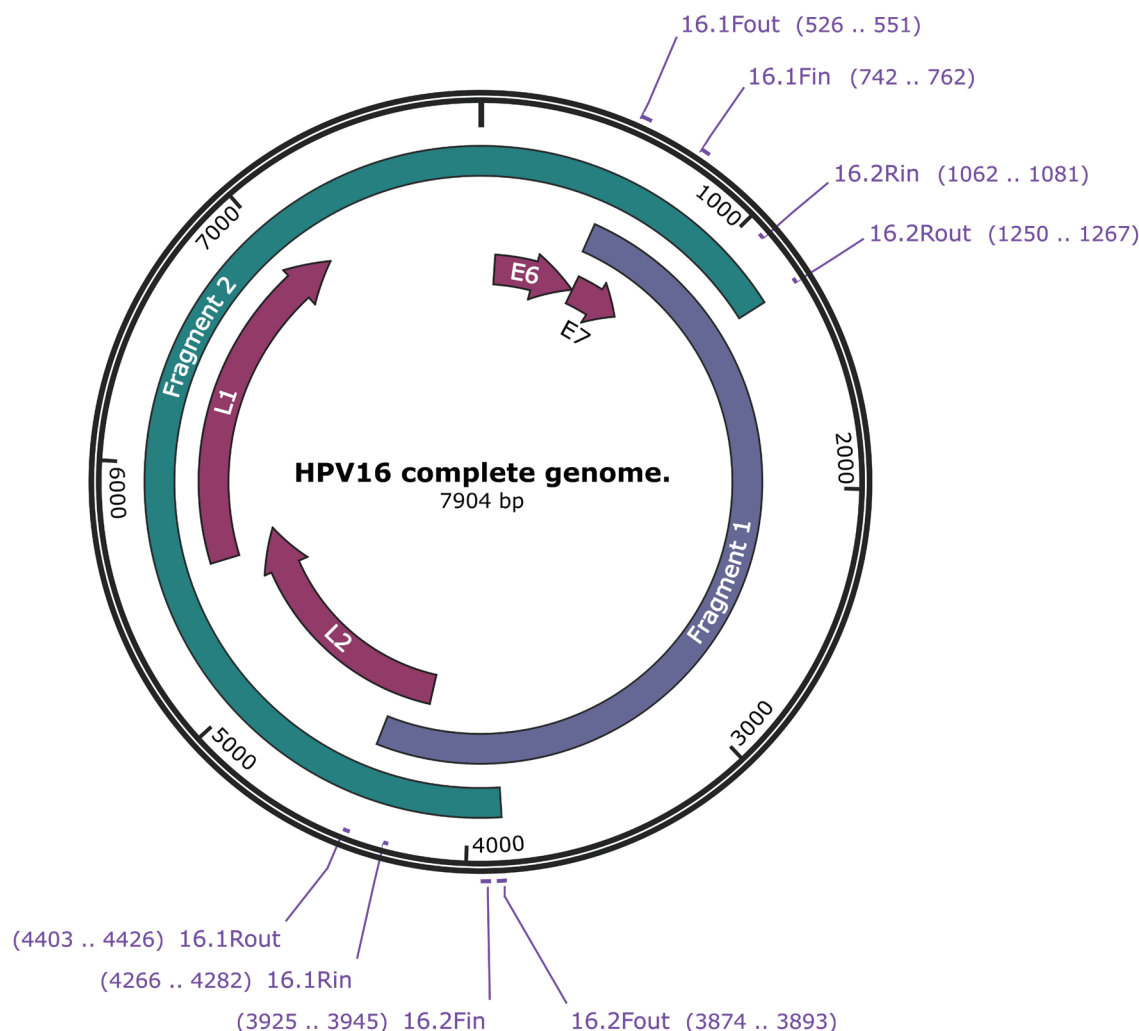
Based on the HPV16 whole-genome reference sequence (K02718.1) from the NCBI Papillomavirus Episteme database, nested PCR primers (Table 1) were designed using Primer 5 software. Fout/Rout served as the outer primers for nested PCR, while Fin/Rin were the inner primers. The overlapping regions of the amplified products were designed to be greater than 500 bp to significantly enhance the efficiency of fragment splicing by providing sufficient homologous sequences for overlapping extension PCR. This design also ensured adequate overlapping read lengths during next-generation sequencing (NGS), thereby improving the accuracy of sequence assembly and the reliability of variant detection. All primers were synthesized by TianyiHuiyuan Biotech.

Two-segment primer set validation

To verify the feasibility, amplifications were performed using two high-fidelity DNA polymerases: 2× Platinum™ SuperFi™ II PCR Master Mix (Thermo Fisher) and 2× Phanta® Max Master Mix (Vazyme). The HPV16 full-length genome was amplified using a two-step nested PCR method. The primer positions are shown in Figure 1, and the primer information is detailed in Table 1.

For the first-round PCR using the Thermo Fisher DNA polymerase, the total reaction volume was 20 μL, which included 10 μL of 2× Platinum™ SuperFi™ II PCR Master Mix, 1 μL of forward primer (10 μmol), 1 μL of reverse primer (10 μmol), 5 μL of template DNA, and 3 μL of ddH<sub>2</sub>O. The reaction conditions were as follows: initial denaturation at 98°C for 30 s; followed by 30 cycles of denaturation at 98°C for 10 s, annealing at 58°C for 10 s, and extension at 72°C for 3 m; and a final extension at 72°C for 5 m, followed by holding at 4°C. For the second-round PCR, the total reaction volume was 50 μL, which included 25 μL of 2× Platinum™ SuperFi™ II PCR Master Mix, 2.5 μL of forward primer (10 μmol), 2.5 μL of reverse primer (10 μmol), 5 μL of template DNA, and 15 μL of ddH<sub>2</sub>O. The reaction conditions were the same as those for the first round.

For the first-round PCR using the Vazyme DNA polymerase,



**Fig. 1. Schematic representation of the location of the two-segment primers in the HPV16 genome.** HPV, human papillomavirus.

the total reaction volume was 25  $\mu$ L, which included 12.5  $\mu$ L of 2 $\times$  Phanta<sup>®</sup> Max Master Mix, 1  $\mu$ L of forward primer (10  $\mu$ mol), 1  $\mu$ L of reverse primer (10  $\mu$ mol), 5  $\mu$ L of template DNA, and 5.5  $\mu$ L of ddH<sub>2</sub>O. The reaction conditions were as follows: initial denaturation at 95°C for 3 m; followed by 30 cycles of denaturation at 95°C for 15 s, annealing at 58°C for 15 s, and extension at 72°C for 3 m; and a final extension at 72°C for 5 m, followed by holding at 4°C. For the second-round PCR, the total reaction volume was 50  $\mu$ L, which included 25  $\mu$ L of 2 $\times$  Phanta<sup>®</sup> Max Master Mix, 2  $\mu$ L of forward primer (10  $\mu$ mol), 2  $\mu$ L of reverse primer (10  $\mu$ mol), 5  $\mu$ L of template DNA, and 16  $\mu$ L of ddH<sub>2</sub>O. The reaction conditions were the same as those for the first round.

In the first-round amplification, two HPV16-positive nucleic acid samples were used as templates. The A260/280 ratio of the samples was measured by ultraviolet spectrophotometry and was found to be between 1.7 and 1.9. The copy numbers of the samples were determined by real-time quantitative PCR to be 2,000 copies/ $\mu$ L and 200 copies/ $\mu$ L, respectively. Nested PCR amplification was performed using two-segment primers. The second-round amplification used the products from the first round as templates. After amplification, 5  $\mu$ L of the amplified product was subjected

to 1% agarose gel electrophoresis. The electrophoresis conditions were 120 V for 25 m. After electrophoresis, the results were observed using a gel electrophoresis imaging system.

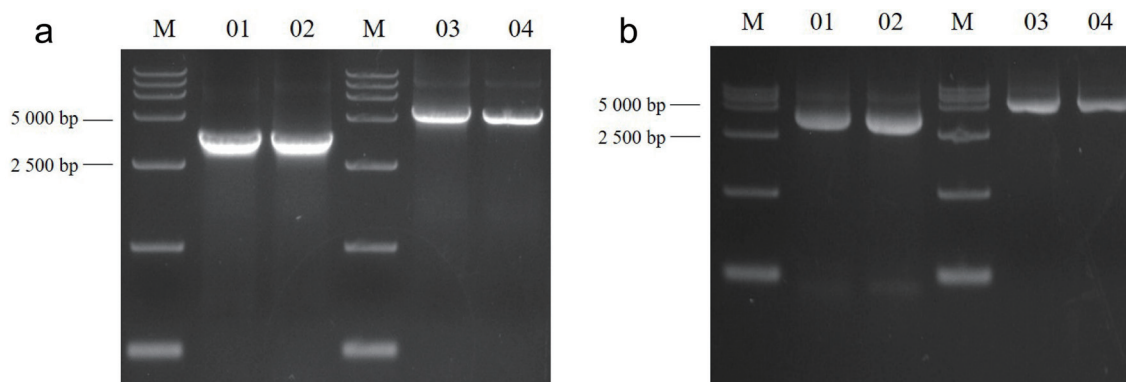
#### **Annealing temperature optimization for PCR amplification**

After validating the nested primer feasibility, HPV16 full-genome nested PCR amplification was performed using Thermo Fisher and Vazyme DNA polymerases at four annealing temperatures (57.2°C, 58.6°C, 59.5°C, and 60°C). The first-round PCR products served as templates for the second-round amplification with inner primers, with a no-template control (ddH<sub>2</sub>O replacing template DNA). The amplified products were analyzed by 1% agarose gel electrophoresis.

For quantitative analysis, ImageJ software was used to quantify and grade band intensity to screen for the optimal annealing temperature, and the effect of different temperatures on non-specific amplification was also evaluated (see Section 2.6 for detailed methods).

#### **Sensitivity assessment for PCR amplification**

The pUC-HPV16 E6 standard curve equation ( $Y = -3.232X + 35.648$ , where Y represents the Ct value and X represents the sam-



**Fig. 2. Agarose gel electrophoresis of two-segment products of HPV16 full genome amplified by Thermo Fisher (a) and Vazyme (b) DNA polymerases.** Lane 01: Fragment 1 (3.9 kb), template 2,000 copies/ $\mu$ L. Lane 02: Fragment 1 (3.9 kb), template 200 copies/ $\mu$ L. Lane 03: Fragment 2 (5.3 kb), template 2,000 copies/ $\mu$ L. Lane 04: Fragment 2 (5.3 kb), template 200 copies/ $\mu$ L. M: 15 kb DNA marker. HPV, human papillomavirus.

ple concentration in logarithmic form) was used to calculate the concentration of the sample template. An HPV16-positive nucleic acid sample with a Ct value of 25 was selected, and its initial copy number was calculated to be 2,000 copies/ $\mu$ L based on the standard curve equation.

The sample was then serially diluted using sterile distilled water ( $\text{ddH}_2\text{O}$ ) as the diluent at dilution factors of  $10^0$ ,  $10^{-1}$ ,  $10^{-2}$ , and  $10^{-3}$ , yielding concentrations of 2,000, 200, 20, and 2 copies/ $\mu$ L, respectively. Quadruplicate amplification tests were conducted using both Thermo Fisher and Vazyme DNA polymerases on the serially diluted samples. The optimal annealing temperatures for Fragment 1 and Fragment 2 obtained from the methods in Section 2.4 were applied for amplification. Using “simultaneous successful amplification of two target fragments” as the criterion for effective amplification, the amplified products were detected by 1% agarose gel electrophoresis. The effective amplification success rate and its 95% confidence interval were calculated, and differences in amplification between the two enzymes in high- and low-concentration groups were compared (see Section 2.6 for detailed statistical methods).

### Statistical analysis

For annealing temperature optimization, the ImageJ software was used for quantitative analysis of agarose gel electrophoresis bands. First, a fixed-area rectangular frame was set to select target bands, and the integrated density of the bands was measured after background subtraction. Then, band intensities were classified into three levels (“+”, low; “++”, medium; “+++”, high) based on integrated density values. The temperature corresponding to the “+++” intensity was selected as the optimal annealing temperature to ensure the highest amplification efficiency.

For sensitivity assessment, using “simultaneous successful amplification of two HPV16 target fragments (3.9 kb and 5.3 kb)” as the criterion for effective amplification, the success rate of quadruplicate experiments for each sample concentration (2, 20, 200, and 2,000 copies/ $\mu$ L) was calculated (number of effective amplifications/total number of replicates  $\times$  100%). The Wilson score method was used to estimate the 95% confidence interval of the success rate. Samples were divided into a high-concentration group ( $\geq 200$  copies/ $\mu$ L) and a low-concentration group ( $< 200$  copies/ $\mu$ L). Fisher’s exact test was used to compare differences in amplification success rates between the two enzymes, with a  $P$ -value  $< 0.05$  indicating a statistically significant difference.

## Results

### Validation of amplification primers

Amplifications were performed using Thermo Fisher DNA polymerase and Vazyme DNA polymerase, with two different copy-number positive HPV16 nucleic acid samples as templates. The results are shown in Figure 2. When the template copy numbers were 2,000 copies/ $\mu$ L and 200 copies/ $\mu$ L, both DNA polymerases successfully amplified the two fragments, with the product band sizes consistent with the theoretical values. Moreover, PCR products were validated by Illumina sequencing, achieving a Q30  $\geq 96\%$  and  $>98\%$  identity with the HPV16 reference sequence (K02718.1), confirming accurate and complete amplification (Fig. 3).

### Annealing temperature optimization

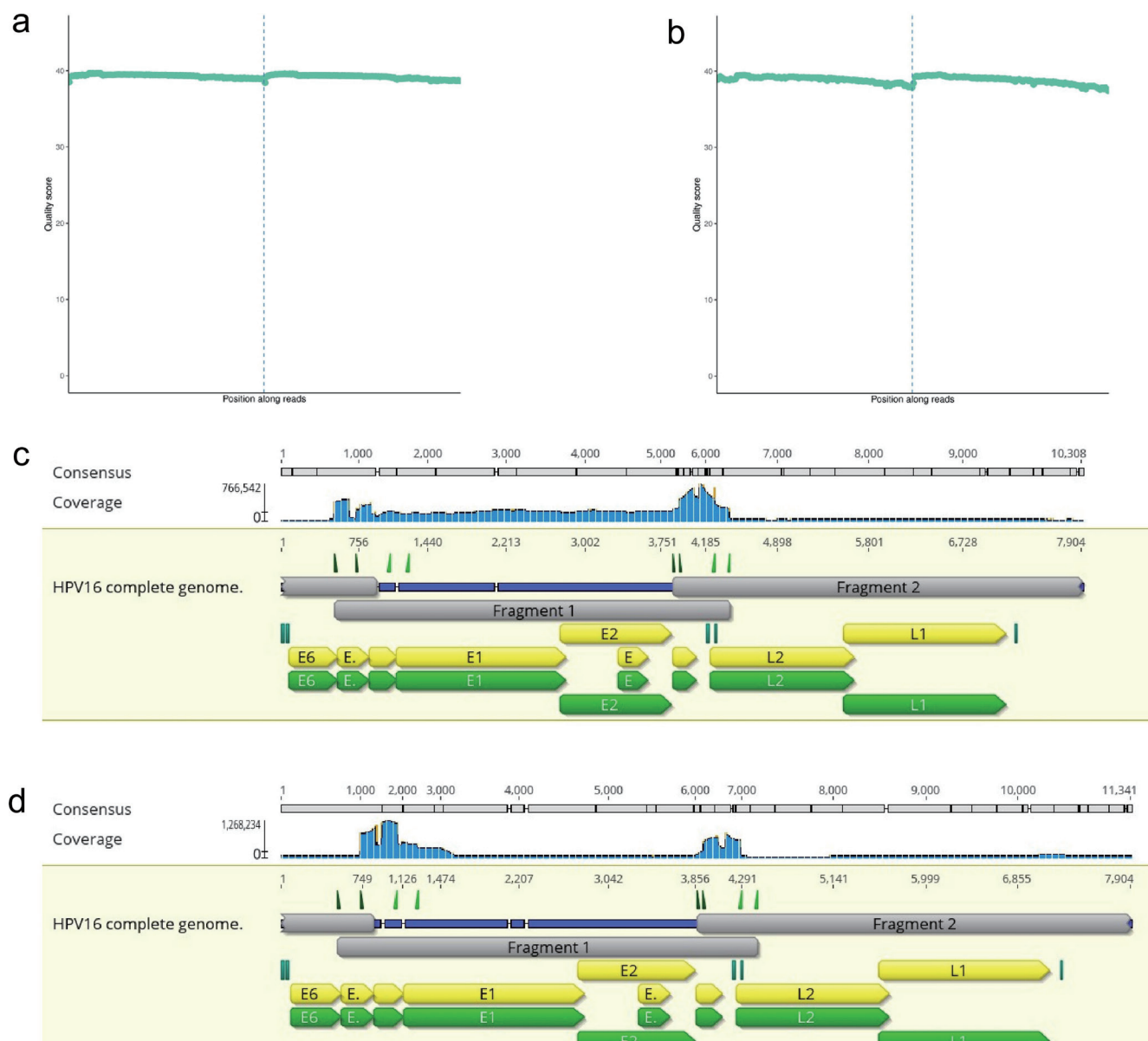
Amplifications were performed using Thermo Fisher DNA polymerase and Vazyme DNA polymerase at different annealing temperatures, and the results are shown in Figure 4. No-template control showed no specific bands, confirming no contamination or non-specific amplification. The amplification results were quantitatively analyzed by ImageJ (Table 2). Based on the summarized data, the annealing temperature that yielded the highest product intensity was selected for amplification. Specifically, for Thermo Fisher DNA polymerase, the annealing temperature for both Fragment 1 and Fragment 2 was determined to be  $60^\circ\text{C}$ . For Vazyme DNA polymerase, the annealing temperature for both Fragment 1 and Fragment 2 was determined to be  $58.6^\circ\text{C}$ .

### Sensitivity assessment

The results showed that Thermo Fisher DNA polymerase could clearly amplify specific bands when the template concentration was as low as 2 copies/ $\mu$ L. Vazyme DNA polymerase could clearly amplify specific bands when the template concentration was as low as 200 copies/ $\mu$ L, while the sample at 20 copies/ $\mu$ L failed to produce a detectable band (Fig. 5).

The statistical analysis of four replicate experiments (Table 3) showed that both enzymes achieved success rates over 75% in samples with concentrations  $\geq 200$  copies/ $\mu$ L, and their 95% confidence intervals overlapped. Fisher’s exact test revealed no significant difference between the two enzymes in this concentration range ( $P = 1.000$ ). In contrast, at concentrations  $< 200$  copies/ $\mu$ L, Thermo Fisher enzyme had a success rate of 62.5% (5/8), while Vazyme enzyme had a success rate of 0% (0/8), indicating a statis-





**Fig. 3. Illumina sequencing validation results of HPV16 amplified products.** a/b: Sequencing quality control base quality plot. c/d: HPV16 sequence assembly and alignment plots after sequencing. HPV, human papillomavirus.

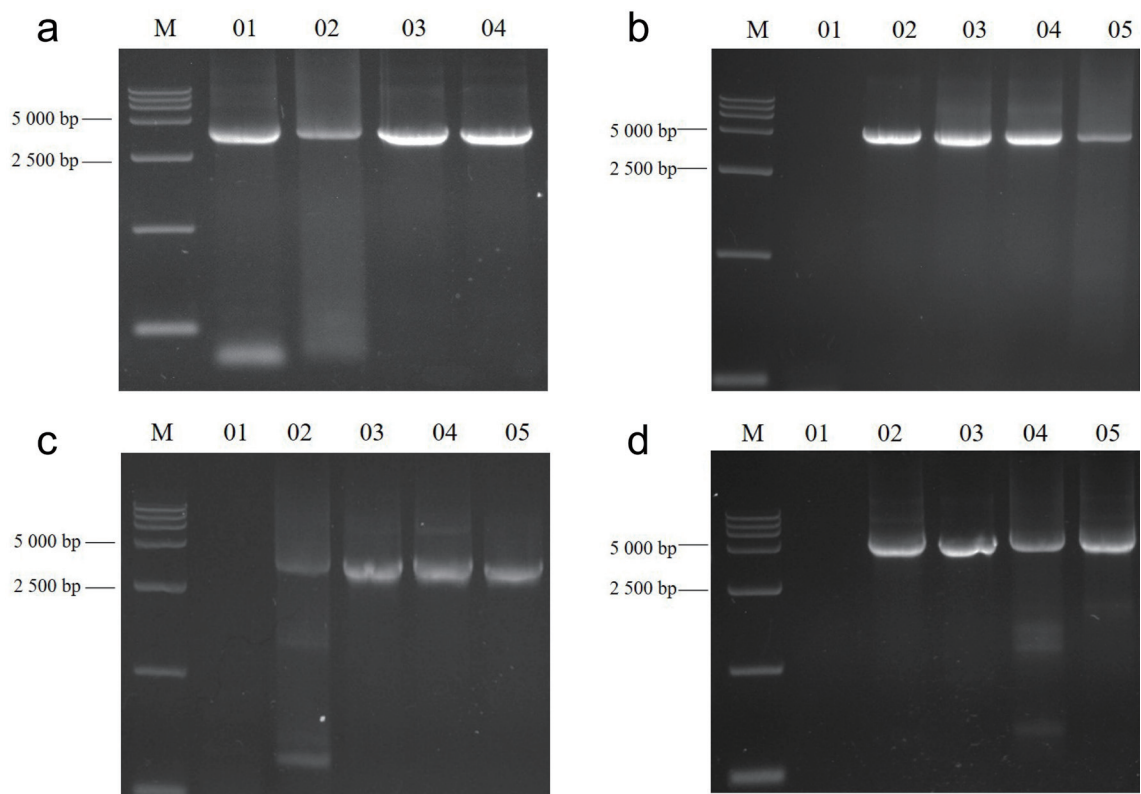
tically significant difference ( $P = 0.037$ ). Two types of failure were observed in replicate experiments: (1) No bands were amplified (the majority of cases); (2) Non-specific bands (non-target fragments) appeared without the expected target bands. These results demonstrated that the method was effective across a concentration range of 2–2,000 copies/ $\mu$ L. While high-concentration samples could be reliably amplified by both enzymes, Thermo Fisher enzyme was particularly effective for low-concentration samples.

## Discussion

HPV is a highly host-specific virus that primarily infects human skin and mucosal epithelial cells. Due to its life cycle's dependence on the specific microenvironment of host cells, HPV is difficult to

effectively replicate *in vitro*,<sup>10,11</sup> which has limited research into its molecular biology. Given the critical role of HPV16 in the development of various malignancies, establishing efficient and reliable whole-genome amplification methods is essential. This will advance our understanding of its oncogenic mechanisms, guide vaccine development, and support precise treatment strategies.

In this study, we established an overlap extension PCR method to amplify the HPV16 whole genome, which is particularly suitable for low viral load samples. This approach effectively avoids the nonspecific amplification and primer competition issues commonly associated with RCA,<sup>12</sup> thereby significantly improving amplification specificity and stability. The full-length HPV16 genome amplified using this method—unlike Cobas 4800 and similar commercial kits—can be used for NGS to comprehensively analyze



**Fig. 4. Agarose electrophoresis images of Thermo Fisher (a, b) and Vazyme (c, d) DNA polymerases at different annealing temperatures.** a: Amplification of Fragment 1; Lanes 01–04 correspond to 57.2°C, 58.6°C, 59.5°C, and 60°C in sequence. b: Amplification of Fragment 2; Lanes 02–05 correspond to 60°C, 59.5°C, 58.6°C, and 57.2°C in sequence; Lane 01 is the no-template control. c/d: Amplification of Fragment 1/2, respectively; Lanes 02–05 correspond to 57.2°C, 58.6°C, 59.5°C, and 60°C in sequence; Lane 01 is the no-template control. M: 15 kb DNA marker.

genomic sequence variations, including single-nucleotide variants, insertions/deletions, and gene recombination. These analyses offer a more comprehensive technical basis for in-depth investigation into the evolution, oncogenic mechanisms, and drug resistance development of HPV16.

Notably, certain regions of the HPV16 genome (e.g., the E4 gene region) contain high guanine–cytosine (GC) content and complex secondary structures, posing additional challenges for whole-genome amplification.<sup>13</sup> To ensure amplification efficiency, we specifically addressed these technical difficulties during method optimization. First, we selected DNA polymerases with high elongation capacity and thermostability. Second, primer design avoided regions with extremely high GC content and strong sec-

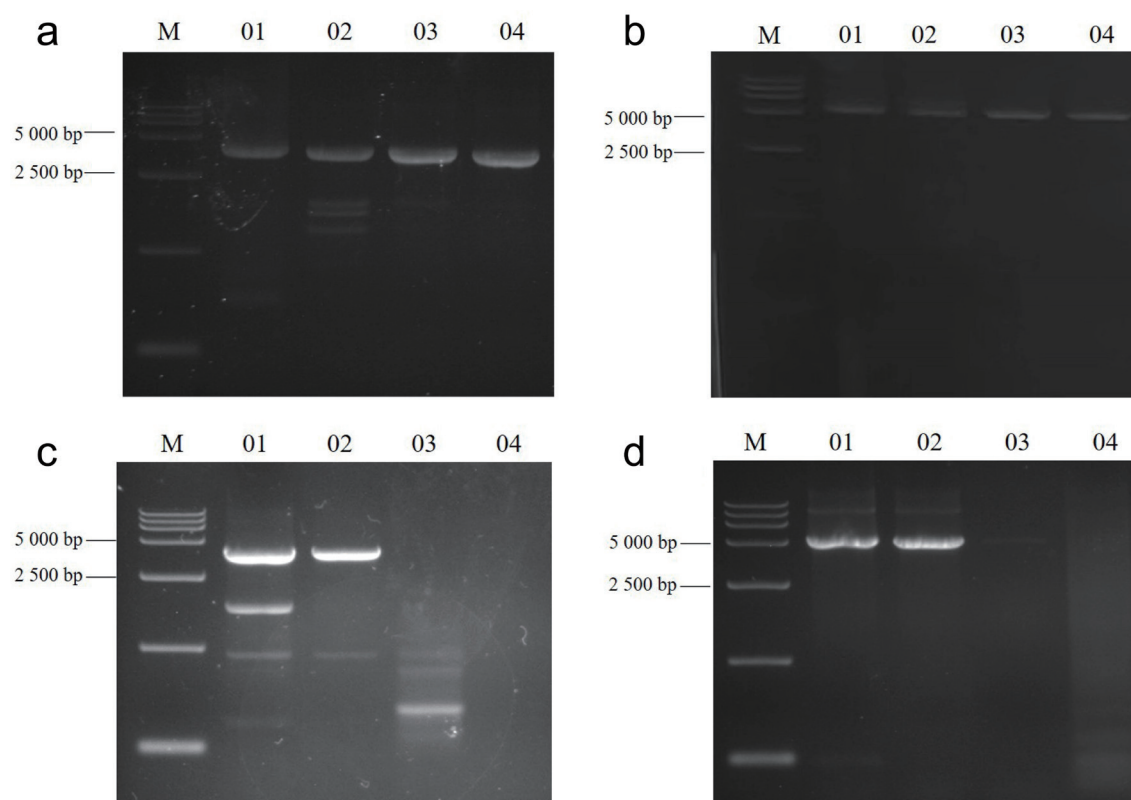
ondary structure formation to ensure primer binding efficiency and amplification specificity. Furthermore, systematic optimization of key parameters (e.g., annealing temperature and extension time) significantly improved amplification efficiency for these complex genomic regions. These optimizations not only resolved technical challenges but also established a foundation for subsequent polymerase performance comparisons.

The sensitivity assessment revealed that both DNA polymerases achieved amplification success rates exceeding 75% in samples with  $\geq 200$  copies/ $\mu\text{L}$ , with no statistical difference detected ( $P = 1.000$ ). This underscores the method’s stability for high-viral-load samples. In contrast, for samples with  $< 200$  copies/ $\mu\text{L}$ , Thermo Fisher DNA polymerase demonstrated a significantly higher success rate (62.5% vs 0%,  $P = 0.037$ ), highlighting its superiority in low-viral-load scenarios. This advantage likely stems from its stronger exonuclease proofreading activity, which maintains stable amplification even with low template quantities. Additionally, its excellent extension efficiency allows for longer DNA chain synthesis per template binding, reducing polymerase dissociation frequency and thereby enhancing amplification efficiency for low-copy templates.<sup>14</sup> In contrast, Vazyme DNA polymerase had lower sensitivity (20 copies/ $\mu\text{L}$ ) but performed well in specificity and stability for routine clinical sample detection. These findings provide important guidance for polymerase selection in different applications: Thermo Fisher polymerase is preferable for processing low-copy-number samples (e.g., whole-genome sequencing preprocessing), while Vazyme DNA polymerase offers a cost-effective alternative for routine clinical testing.

**Table 2. Effects of Thermo Fisher and Vazyme DNA polymerases at different temperatures**

Annealing temperature (°C)	57.2°C	58.6°C	59.5°C	60°C
Thermo Fisher				
Fragment one	++	+	++	+++
Fragment two	+	+++	+++	+++
Vazyme				
Fragment one	+	+++	++	++
Fragment two	++	+++	+	++

“+”, low; “++”, medium; “+++”, high.



**Fig. 5. Amplification results of two DNA polymerases on HPV16 samples with different concentrations (representative electrophoresis image of four replicate experiments; the band pattern is consistent with the overall trend of all repeats).** a/b: Fragment 1/2 amplified by Thermo Fisher DNA polymerase; Lanes 01–04 correspond to template copy numbers 2, 20, 200, and 2,000 copies/μL in sequence. c/d: Fragment 1/2 amplified by Vazyme DNA polymerase; Lanes 01–04 correspond to template copy numbers 2,000, 200, 20, and 2 copies/μL in sequence. M: 15 kb DNA marker. HPV, human papillomavirus.

In terms of clinical applications, our method shows significant technical advantages. First, it requires only conventional PCR equipment and completes HPV16 whole-genome amplification within 4 h while maintaining high compatibility with mainstream NGS platforms like Illumina. In terms of cost-effectiveness, the per-sample testing cost is lower than commercial kits, providing an economical solution for large-scale clinical screening.<sup>15–17</sup> More importantly, the complete genome data obtained through this method can be directly applied to multiple fields, including clinical genotyping, oncogenic mutation analysis, and vaccine escape monitoring, offering crucial evidence for personalized diagnosis and treatment. Establishing standardized, automated high-throughput detection protocols in the future will further enhance the clinical value of this method.

### Limitations

While the HPV16 whole-genome amplification method developed in this study shows several advantages, some limitations require improvement. First, validation was performed only using cervical exfoliated cell samples, without including samples from other anatomical sites (e.g., oropharynx, anus), which may affect its broad applicability. Additionally, the samples did not systematically cover the full disease spectrum from HPV infection to cancer development, limiting the evaluation of its performance across different disease stages. Future studies should validate the method using more diverse sample types (e.g., oral/anal swabs) and clinical cohorts (e.g., patients with cervical lesions at different stages) to assess its potential for diagnostic and precision medicine applications. Second, although amplification conditions for GC-rich

**Table 3. The effective amplification success rates and 95% confidence intervals of two DNA polymerases in HPV16 samples of different concentrations**

Sample concentration	Thermo Fisher DNA polymerase	Vazyme DNA polymerase
Effective amplification success rate (95% confidence interval)		
2,000 copies/μL	100% (47.8–100%)	100% (47.8–100%)
200 copies/μL	100% (47.8–100%)	75% (24.9–96.8%)
20 copies/μL	75% (24.9–96.8%)	0% (0–39.3%)
2 copies/μL	50% (11.6–88.4%)	0% (0–39.3%)

The success rate of effective amplification = number of effective amplifications/total number of repeats (four repeats for each concentration). The 95% confidence interval was calculated using the Wilson scoring method. HPV, human papillomavirus.

regions were optimized, efficiency for strains with extreme GC content still needs improvement. Developing novel thermostable DNA polymerases or optimizing buffer systems may help address this limitation in future work.

## Conclusions

The overlapping extension PCR method established in this study demonstrates high sensitivity and specificity for amplifying HPV16 in cervical exfoliated cells. This method provides a robust technical foundation for genomic research, with potential applications in whole-genome sequencing and variation analysis. Additionally, this study provides insights for the optimization of nested PCR amplification systems and serves as a reference for other studies involving the amplification of long DNA fragments.

## Acknowledgments

We thank all the participants from the Obstetrics and Gynecology Department of Peking University First Hospital for providing us with the specimens.

## Funding

This study was supported by the Natural Science Foundation of Hebei Province (No. H2022209048) and the National Key Research and Development Program (No. 121000004008239033–2021YFC2301103/05).

## Conflict of interest

The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. Informed consent was obtained for this publication.

## Author contributions

Study conception and design (ZC, XZ, SL), main data analysis, and manuscript draft (ZC, HL, SZ, HY, XZ, SL). All authors contributed to the article and approved the final version.

## Ethical statement

This study was conducted in accordance with the ethical standards of the Declaration of Helsinki (as revised in 2024) and was approved by the Ethics Committee of North China University of Science and Technology (Approval Number: 20250616). Informed consent was obtained from all patients, and the confidentiality of patient information was ensured through data anonymization.

## Data sharing statement

After publication, all major data sheets will be available upon request.

## References

- [1] de Villiers EM, Fauquet C, Broker TR, Bernard HU, zur Hausen H. Classification of papillomaviruses. *Virology* 2004;324(1):17–27. doi:

Chen Z. *et al*: A method for amplifying the full genome of HPV16

- 10.1016/j.virol.2004.03.033, PMID:15183049.
- [2] Yu L, Majerciak V, Zheng ZM. HPV16 and HPV18 Genome Structure, Expression, and Post-Transcriptional Regulation. *Int J Mol Sci* 2022;23(9):4943. doi:10.3390/ijms23094943, PMID:35563334.
- [3] Zhang Y, Qiu K, Ren J, Zhao Y, Cheng P. Roles of human papillomavirus in cancers: oncogenic mechanisms and clinical use. *Signal Transduct Target Ther* 2025;10(1):44. doi:10.1038/s41392-024-02083-w, PMID:39856040.
- [4] Biglari S, Moghaddam AS, Tabatabaiefar MA, Sherkat R, Youssefian L, Saeidian AH, *et al*. Monogenic etiologies of persistent human papillomavirus infections: A comprehensive systematic review. *Genet Med* 2024;26(2):101028. doi:10.1016/j.gim.2023.101028, PMID:37978863.
- [5] Levi AW, Bernstein JI, Hui P, Duch K, Schofield K, Chheng DC. A Comparison of the Roche Cobas HPV Test With the Hybrid Capture 2 Test for the Detection of High-Risk Human Papillomavirus Genotypes. *Arch Pathol Lab Med* 2016;140(2):153–157. doi:10.5858/arpa.2015-0027-OA, PMID:26910220.
- [6] Agorastos T, Chatzistamatiou K, Katsamagkas T, Koliopoulos G, Daponte A, Constantinidis T, *et al*. Primary screening for cervical cancer based on high-risk human papillomavirus (HPV) detection and HPV 16 and HPV 18 genotyping, in comparison to cytology. *PLoS One* 2015;10(3):e0119755. doi:10.1371/journal.pone.0119755, PMID:25793281.
- [7] Karunanathie H, Kee PS, Ng SF, Kennedy MA, Chua EW. PCR enhancers: Types, mechanisms, and applications in long-range PCR. *Biochimie* 2022;197:130–143. doi:10.1016/j.biochi.2022.02.009, PMID:35231536.
- [8] Tian W, Li P, He W, Liu C, Li Z. Rolling circle extension-actuated loop-mediated isothermal amplification (RCA-LAMP) for ultrasensitive detection of microRNAs. *Biosens Bioelectron* 2019;128:17–22. doi:10.1016/j.bios.2018.12.041, PMID:30616213.
- [9] Han J, Song Y, Lee J, Nyamzaya K, Kim E. Rolling circle amplification for next-generation molecular diagnostics, genome analysis, and spatial transcriptome profiling. *Nanoscale* 2025;17(37):21345–21378. doi:10.1039/d5nr03032c, PMID:40917009.
- [10] Jin J, Li S, Huang H, Li J, Lyu Y, Ran Y, *et al*. Development of human papillomavirus and its detection methods (Review). *Exp Ther Med* 2024;28(4):382. doi:10.3892/etm.2024.12671, PMID:39161614.
- [11] Unckell F, Strecek RE, Sapp M. Generation and neutralization of pseudovirions of human papillomavirus type 33. *J Virol* 1997;71(4):2934–2939. doi:10.1128/JVI.71.4.2934-2939.1997, PMID:9060652.
- [12] Inoue J, Shigemori Y, Mikawa T. Improvements of rolling circle amplification (RCA) efficiency and accuracy using *Thermus thermophilus* SSB mutant protein. *Nucleic Acids Res* 2006;34(9):e69. doi:10.1093/nar/gkl350, PMID:16707659.
- [13] Roberts S, Ashmole I, Gibson LJ, Rookes SM, Barton GJ, Gallimore PH. Mutational analysis of human papillomavirus E4 proteins: identification of structural features important in the formation of cytoplasmic E4/cytokeratin networks in epithelial cells. *J Virol* 1994;68(10):6432–6445. doi:10.1128/JVI.68.10.6432-6445.1994, PMID:7521917.
- [14] Laatri S, El Khayari S, Qriouet Z. Exploring the molecular aspect and updating evolutionary approaches to the DNA polymerase enzymes for biotechnological needs: A comprehensive review. *Int J Biol Macromol* 2024;276(Pt 2):133924. doi:10.1016/j.ijbiomac.2024.133924, PMID:39033894.
- [15] Lorenzi AT, Fregnani JH, Possati-Resende JC, Neto CS, Villa LL, Longatto-Filho A. Self-collection for high-risk HPV detection in Brazilian women using the careHPV™ test. *Gynecol Oncol* 2013;131(1):131–134. doi:10.1016/j.ygyno.2013.07.092, PMID:23880151.
- [16] Gautam A, Gedda MR, Rai M, Sundar S, Chakravarty J. Human Papillomavirus Genome based Detection and Typing: A Holistic Molecular Approach. *Curr Mol Med* 2019;19(4):237–246. doi:10.2174/1566524019666190405120441, PMID:30950352.
- [17] Casas CPR, Albuquerque RCR, Loureiro RB, Gollner AM, Freitas MG, Duque GPDN, *et al*. Cervical cancer screening in low- and middle-income countries: A systematic review of economic evaluation studies. *Clinics (Sao Paulo)* 2022;77:100080. doi:10.1016/j.clin-sp.2022.100080, PMID:35905574.