Vietnam National University, Hanoi

International School

==========

M.A Thesis Proposal

GENOME SEQUENCE ANALYSIS OF HUMAN PAPILLOMAVIRUS TYPE 11 (HPV-11) INFECTING VIETNAMESE MEN

HOANG DUC ANH

Field: Biomedical Engineering Technology

Code: 8520212

Hanoi – 2026

Vietnam National University, Hanoi

International School

==========

M.A Thesis Proposal

GENOME SEQUENCE ANALYSIS OF HUMAN PAPILLOMAVIRUS TYPE 11 (HPV-11) INFECTING VIETNAMESE MEN

HOANG DUC ANH

Field: Biomedical Engineering Technology

Code: 8520212

Supervisor: Dr. Than Van Thai

Hanoi – 2026

CERTIFICATE OF ORIGINALITY

I, the undersigned, hereby certify my authority for the study project report entitled "***Genome sequence analysis of Human papillomavirus type 11 infecting Vietnamese men***" submitted in partial fulfillment of the requirements for the degree of Master of Biomedical Engineering Technology. Except where the reference is indicated, no other person's work has been used without due acknowledgement in the text of the thesis.

Hanoi, 2026

Hoang Duc Anh

Approved by

SUPERVISOR

(Signature and full name)

Than Van Thai

Date:……

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Dr. Than Van Thai, a lecturer at the Faculty of Applied Sciences and a member of the Interdisciplinary Research Group on Biomedicine and Health at the International School, Vietnam National University, Hanoi, for his direct guidance and companionship throughout my research journey.

I would also like to extend my deepest gratitude to Assoc. Prof. Dr. Chu Dinh Toi, Dean of the Faculty of Applied Sciences and Head of the Interdisciplinary Research Group on Biomedicine and Health at the International School, Vietnam National University, Hanoi, for his dedicated support, valuable insights, and guidance in helping me complete this thesis.

Furthermore, I would like to express my gratitude to my family, friends, and colleagues at the Interdisciplinary Research Group on Biomedicine and Health, International School, Vietnam National University, Hanoi. Their constant support, encouragement, and invaluable assistance have been a source of great strength, helping me overcome numerous challenges throughout my academic and professional journey.

Finally, I would like to express my sincere appreciation to the faculty members of the Faculty of Applied Sciences and the Office of Graduate Studies at the International School, Vietnam National University, Hanoi, for providing favorable conditions and continuous support during my studies and the completion of this thesis.

Hanoi, 2026

Graduate Student

TABLE OF CONTENTS

ABSTRACT

Human Papillomavirus (HPV) is currently the most common sexually transmitted virus, causing various types of cancer, including cervical cancer, head and neck squamous cell carcinoma (HNSCC), and anal cancer. Human Papillomavirus type 11 (HPV-11) is one of the primary causative agents of benign epithelial lesions, including condylomata acuminata (genital warts) and recurrent respiratory papillomatosis, severely affecting men's quality of life. However, genomic data on circulating HPV-11 variants in Vietnamese men remains highly limited, posing challenges for epidemiological surveillance as well as the development of diagnostic and preventive strategies. This study aims to comprehensively analyze the whole-genome sequences of HPV-11 strains isolated from Vietnamese men to elucidate their phylogenetic relationships, and characteristic mutational profiles. Clinical samples were collected from male patients diagnosed with HPV-11 infection. Following DNA extraction, the viral genome was amplified via polymerase chain reaction (PCR). The amplified products were then subjected to Sanger sequencing. Finally, bioinformatics analysis, including sequence alignment and phylogenetic construction was performed using BioEdit, MEGA, and DNASTAR software. In this study, a set of specific primers was designed to amplify the full-length viral genome, yielding the complete sequence of the Vietnamese HPV-11 strain with a total length of 7,932 bp. Phylogenetic analysis revealed that the Vietnamese HPV-11 strains belonged to lineage A, sublineage A2, demonstrating a close genetic relationship with Asian strains from China and Thailand, as well as European strains from Slovenia and Hungary. Furthermore, mutational profiling identified several nonsynonymous mutations compared to Chinese and European references, specifically: 6 in the E6 gene, 3 in the E7 gene, 6 in the L1 gene, and 8 in the L2 gene. These findings provide detailed whole-genome characteristics of HPV-11 isolates from Vietnamese men, thereby expanding the global papillomavirus genetic database.

LIST OF ABBREVIATIONS

ABBREVIATION	DEFINITION
HPV	*Human Papillomavirus*
HNSCC	Head and Neck Squamous Cell Carcinoma
HR-HPV	High-risk Human papillomavirus
LR-HPV	Low-risk Human papillomavirus
ORFs	Open Reading Frames
URR	Upstream Regulatory Region
LCR	Long Control Region
ORI	Origin of replication
EGFR	Epidermal Growth Factor Receptor
VEGF	Vascular Endothelial Growth Factor
RRP	Recurrent Respiratory Papillomatosis
WGS	Whole-Genome Sequencing

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1. INTRODUCTION

Human Papillomavirus (HPV) belongs to the Papillomaviridae family is the most prevalent sexually transmitted virus today, impacting the lives of both men and women worldwide. HPV is primarily transmitted through skin-to-skin contact or contact between the skin and the mucosa of the cervix and mouth. HPV infection is the underlying cause of various cancers, including cervical cancer, head and neck squamous cell carcinoma (HNSCC), and anal cancer [1].

Based on differences in DNA sequences, scientists have identified more than 200 HPV types to date. They are primarily classified into three main groups: high-risk HPV (HR-HPV) types including HPV 16, 18, 31, 33, 35, 45, 52, and 58, are strongly associated with cancers of the cervix, anus, vagina, vulva, penis, and oropharynx [2, 3]. potentially carcinogenic types; and low-risk HPV (LR-HPV) types, such as HPV-6 and HPV-11, cause of genital warts and recurrent respiratory papillomatosis [4]. Although LR-HPV genotypes have a low probability of causing cancer, the resulting lesions can severely affect patients' well-being and psychological health. Furthermore, they incur substantial medical expenses and impose a tremendous burden on the public healthcare system [5].

Human papillomavirus type 11 (HPV-11) is classified under the genus Alphapapillomavirus, species Alphapapillomavirus 10. HPV-11 is one of the most prevalent strains within the LR-HPV group, primarily causing genital warts and laryngeal papillomatosis in both men and women. Although most lesions induced by HPV-11 do not progress into malignancy, recent studies have reported its presence in 0–5.5% of penile cancers and 0–87.5% of laryngeal cancers [6]. In men, HPV-11 often presents without clear symptoms, leading to late diagnosis and an increased risk of transmission. Furthermore, it causes lesions in the genital areas that result in pain, itching, and discomfort, significantly affecting daily activities and sexual health. These lesions can spread and recur frequently, leading to a substantial impact on the physical health, psychological well-being, and overall quality of life of infected individuals [5].

In Vietnam, HPV infection in men, specifically HPV-11, receives less attention than in women. Educational and communication activities regarding transmission methods, prevention through vaccination, and especially the risk of progression to various forms of cancer, have not been emphasized for the male population. Current

HPV-related research primarily focuses on epidemiological studies and surveys, with very few studies on molecular biology and, notably, the full-length genome sequencing of HPV-11. Therefore, to gain deeper insights into the genetic diversity of HPV-11 among Vietnamese men, we conducted a study titled "*Genome sequence analysis of human papillomavirus type 11 (HPV-11) infecting Vietnamese men.*" To achieve this general objective, the study set out two specific objectives:

- Objective 1: To determine the complete genome of HPV-11 and establish its classification.
- Objective 2: To analyze genomic characteristics and identify variants of the sequenced genome.

CHAPTER 2. LITERATURE REVIEW

2.1. Review of previous studies

The total prevalence of HPV virus globally is 31% - 50.8% (95% CI 27-35). The average HPV infection rate in men worldwide is approximately 7.9%, ranging from 3.5 - 45% depending on age and country. Of which, the infection rate of LR types (mainly HPV-11, HPV-6) is from 2.3% to 23.9% [7].

A systematic review of 9,342 men in sub-Saharan African countries demonstrated that among the common HPV strains, the individual prevalence rates were 57.0% for HPV-6, 27.4% for HPV-16, 20.5% for HPV-18, and 14.7% for HPV-11 [8]. A retrospective study from 2016-2022 on 3,681 male patients in Liaocheng, China reported the common low-risk genotypes as HPV-6 (56.99%), HPV-11 (23.79%) and HPV-43 (6.37%) [9]. A study in Turkey (2023) by Nilgun Tekkesin et al., on 1304 men, showed that the prevalence of common HPV genotypes in the country was 5.7% (in HR strains such as HPV-16) and 20,08% (in LR strains such as HPV-11) [10]. From this, it can be concluded that HPV-11 is the most common virus strain in the LR- HPV group in men globally.

In Vietnam, HPV infection in men poses a significant risk of transmission to women through sexual activity. Despite this, the issue of HPV infection in men in our country receives very little attention compared to women. A few studies have indicated that the percentage of male patients with HPV who seek treatment at men's health clinics is approximately 25-34%. This includes the study by Lac Thi Kim Ngan and colleagues in 2025, which studied the clinical characteristics of HPV strains and the treatment results of genital warts at Can Tho City Dermatology Hospital in 2020. The cross-sectional descriptive study was conducted with 109 patients diagnosed with genital warts at Can Tho Dermatology Hospital. The study showed that HPV was detected in 89% of patients. LR- HPV types such as HPV-11 (50.5%) and HPV-6 (47.4%) were the most common. HR- HPV types such as HPV-51 (30.9%) and HPV-52 (20.6%) were also common, and there was an association between HPV genotype and sex [11]. Another study by Nguyen Hoai Bac et al. in 2022 on 941 male patients with sexually transmitted diseases found that 191 patients had HPV virus DNA, most commonly in the age group from 20 to 39 years old. The circulating genotypes include two LR genotypes that were mainly detected: HPV-11 (36.7%) and HPV-6 (21.4%) [12]. From this, it can be seen that HPV-11 is one of the most common

LR- HPV virus types in Vietnamese men with sexually transmitted diseases. On the other hand, current studies related to HPV mainly focus on epidemiological studies and evaluation surveys, while very few studies are related to molecular biology and especially the complete genome sequencing of HPV-11.

2.2. Review of theoretical background

2.2.1. Morphology and structure of HPV

Human Papillomavirus is a non-enveloped virus belonging to the Papillomaviridae family, exhibiting helical symmetry. HPV particles have a diameter of 52–55 nm, with a capsid composed of 72 capsomeres. Each capsomere is formed by two proteins: L1 and L2 (these proteins serve as the antigenic components used in specific immune responses). These structural proteins are all encoded by the virus itself: The major capsid protein (L1) has a size of about 55 kDa and accounts for about 80% of the total viral protein. The secondary capsid protein (L2) has a size of about 70 kDa [13]. The capsomeres diagram of the virus particle is shown in Figure 1 [14].

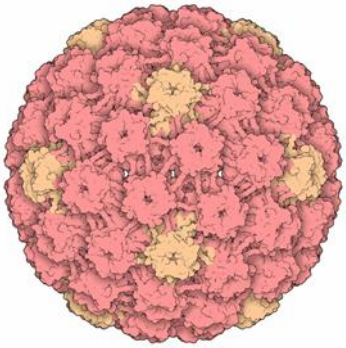


Figure 1: diagram of the HPV virus particles capsomeres

HPV has genetic material consisting of a double-helix DNA strand approximately 8,000 base pairs (bp) in length. The viral genome accounts for about 12% of the total viral particle weight, with a guanosine and cytosine (G-C) content of 42%. The HPV DNA structure comprises 8 to 10 open reading frames (ORFs) (depending on the specific strain), which are divided into three critical regions:

- The Upstream Regulatory Region (URR), also known as the Long Control Region (LCR), contains non-coding DNA functions to regulate DNA replication and transcription. This is the most hypervariable region, accounting for about 10% of the genome length. The URR region sequence includes:

- Enhancer sequences: Binding sites for transcription factors such as AP-1, NF1, Oct-1, TEF1, TEF2, YY1, etc.
- Promoter: Includes the TATA structure and the start region for RNA synthesis transcription.
- Origin of replication (ORI): Contains activation subunits and some silencing genes.

- The Early Region: Consists of 6 genes, designated E1, E2, E4, E5, E6, and E7, and their respective ORFs. The products of this region are functional proteins that facilitate viral DNA replication, induce cell proliferation, and cause cellular transformation, leading to cell immortality.
- The Late Region: Consists of 2 genes that synthesize the L1 and L2 proteins, which are the structural capsid proteins of the virus.

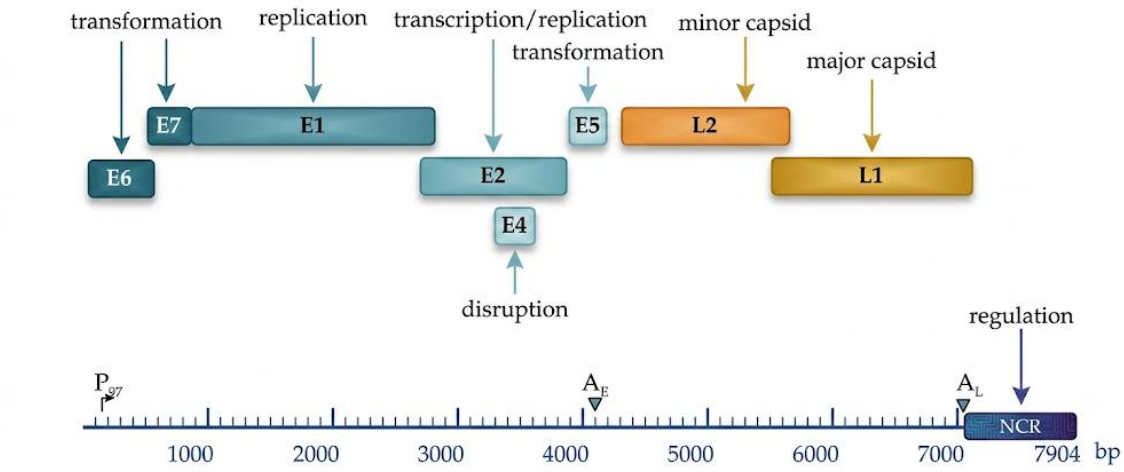


Figure 2: Schematic Diagram of HPV Gene Organization and Protein Function [15].

E1-E7: HPV genes; E1 proteins interfere with the replication process, E2 proteins interfere with transcription and replication, while E5, E6, and E7 proteins are involved in the transformation process.

2.2.2. Functions and products of HPV component genes

- **Function of the E1 gene**

The E1 gene (along with the L1 gene) is one of the two most conserved gene regions of the HPV virus, encoding the functional E1 protein, which is essential for DNA and plasmid replication. The E1 protein recognizes and binds to the replication origin

(ori). However, the E1 protein has a relatively weak affinity for DNA. It requires the assistance of the E2 protein (a transcriptional regulatory protein) to form the E1-E2 complex, which helps E1 bind firmly and accurately to the origin [16, 17]. The most important function of the E1 gene is to perform DNA splitting (helicase) and help the viral gene strands unfold during replication, which is process that depends on adenosine triphosphate (ATP) [18]. In addition, the E1 gene has also been shown to interact with human DNA polymerase to initiate the replication of new viral DNA strands [19]. In addition to its traditional functions in viral replication, the E1 protein has been identified as a potent inhibitor of the host immune response. This interaction not only facilitates the viral life cycle and sustains viral persistence but also drives tumor initiation and progression [20].

- **Function of the E2 gene**

Besides its function in viral DNA replication, the E2 gene plays a key role in transcription, as well as in regulating transcription and maintaining the viral gene sequence outside the chromosome. The transcriptional regulatory function of the E2 gene is achieved through binding to E2BSs in the viral gene sequence that have an affinity for the E2 gene, and these related sites determine the effectiveness of the E2 gene in the transcription process. In the gene sequence of HR- HPV, the E2 gene is capable of inhibiting transcription from factors that promote the early expression of viral genes; therefore, when HR-HPV genes enter the host chromosome, it increases the likelihood of expressing the oncogenes E6 and E7 [21].

- **Function of the E4 gene**

The E4 gene codes for the E4 protein. The synthesis of the E4 protein is regulated by Protein Kinase A [22]. Furthermore, research on the conserved LLXLL region at positions 12–16 of the HPV-16 E4 protein indicates its important role in cell keratinization through its function of synthesizing cytoplasmic fibers [23]. Beyond cytoskeleton reorganization, the E4 protein plays a pivotal role in cell cycle regulation and the facilitation of viral DNA replication. Specifically, E4 has been shown to induce G_2 arrest in host cells, primarily by preventing the translocation of the active Cyclin B1-CDK1 complex into the nucleus [24]. In addition, the E4 protein is also capable of directly binding to and disrupting the mitochondrial function of the host cell. After binding to and collapsing the cytokeratin network, the E4 protein utilizes a specific leucine cluster-containing region at its N-terminus to localize and adhere

to the mitochondria. This binding severs the connection between the mitochondria and microtubules, causing them to cluster into a large mass next to the cell nucleus. More severely, this accumulation leads to the loss of mitochondrial membrane potential and directly triggers the cellular apoptotic pathway. Biologically, this dual destruction mechanism simultaneously disrupting the keratinocyte cytoskeleton and promoting cell apoptosis is a key strategy of HPV to weaken the host cell structure in the upper epidermal layers, thereby facilitating the efficient release and dissemination of new viral particles into the environment [25].

- **Function of the E5 gene**

The E5 gene codes for the E5 protein, which is characterized as a small, hydrophobic protein localized in the Golgi membrane and endoplasmic reticulum. The E5 protein is considered essential for viral entry and persistence in the host cell. By activating the Epidermal Growth Factor Receptor (EGFR) and upregulating Vascular Endothelial Growth Factor (VEGF), the E5 protein is capable of initiating several cellular pathways. In addition, the E5 protein has a role in inhibiting apoptosis (programmed cell death). Apoptotic receptors are activated by TNF and Fas receptor by Fas ligand. One study demonstrated that the E5 protein reduces the activity of FasL and Fas by inhibiting caspase 8. The E5 protein binds to B cell protein associated protein 31 (Bap 31) to form an immune complex, inhibiting the degradation of Bap 31 protein, which leads to the suppression of programmed cell death. The E5 protein also activates transmembrane lipoprotein A4, which interacts with Bap 31, leading to increased cell division (mitogenesis). The mechanism of E5 protein action is detailed in **Figure 3** [26].

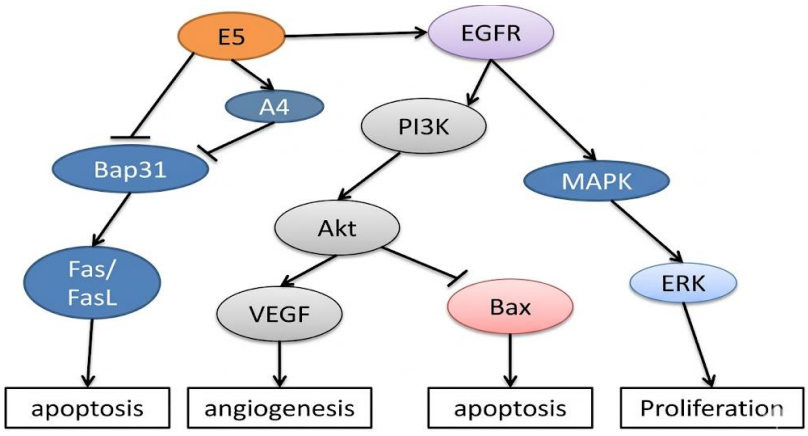


Figure 3: Effects of E5 protein on cell signaling pathways

- **Function of the E6 gene [27]**

The E6 gene codes for the E6 protein, which consists of approximately 150 amino acids forming the Cys-X-X-Cys structure that binds and regulates zinc (Zn). It codes for the first ORF (Open Reading Frame) in the HPV gene cluster and is one of the virus's main oncoproteins (cancer-causing proteins). The E6 gene has three main specific functions:

(1) E6 protein in HR-HPV types binds, either alone or in combination with the E7 protein, causing strong and continuous host cell proliferation and leading to cell immortalization.

(2) Interacts with p53 through the E6 and E6 and AP ligation, leading to the degradation of p53, which is a highly important transcription factor and tumor suppressor. Additionally, E6 binds to PDZ proteins, causing the degradation of these proteins.

(3) Binds to the ras gene during the process of cell immortalization, which helps stimulate the growth of NIH3T3 and activates the E2 promoter of Adenovirus.

- **Function of the E7 gene**

The E7 protein is coded by the E7 gene and is composed of 98 amino acids featuring two critical conserved regions that play an important role in the mechanism of cellular carcinogenesis. The first conserved gene region, CR1, spans amino acids 37-49 and participates in the replication process and p53 degradation. The CR2 conserved region, located between amino acids 116-137, is involved in phosphorylation and p53 binding. The functional activity of E7 in the carcinogenic mechanism is due to the following reasons: (1) The E7 protein has a conserved N-terminus region and a zinc-binding domain at the C-terminus, which facilitates tighter binding to E6, supporting each other in the cell immortalization mechanism; (2) E7 contains the pocket protein binding motif (LXCXE), which helps E7 bind to tumor suppressor genes (such as pRb) or bind to two other pocket proteins, p107 and p130, leading to the release of a large number of tumor suppressive factors. The E7 protein is known for its ability to interact with the Retinoblastoma protein (pRb), the PI3K/Akt pathway, and miRNA. The function of pRb is to bind to the E2F1-3 transcription factors to prevent the function of E3E1-3. When the E7 protein binds to pRb, it prevents pRb from binding to E2F1-3. The liberated E2F1-3 can then bind to the cell cycle regulatory factor DP-

1, DP-2, shortening the G1-S phase, increasing genomic instability, and reducing apoptosis. The E7 protein also reduces the stability of pRb by interfering with intracellular protein degradation via the cullin ubiquitin complex. Furthermore, the E7 protein binds directly to E2, promoting E2F's activity in driving the cell into the S phase and inhibiting the transcriptional activity of E2F6, thereby prolonging the S phase in HPV-infected cells. Through different pathways, the E7 protein interacts with the pRb tumor suppressor protein, shortening the cell's G1-S phase repair stage, facilitating the accumulation of mutations during each mitotic division. Free E2F transcription stimulates the transcription process, extending the lifespan of the cell. A diagram of the E7 protein and the location of its conserved regions is shown in Figure 4 [28].

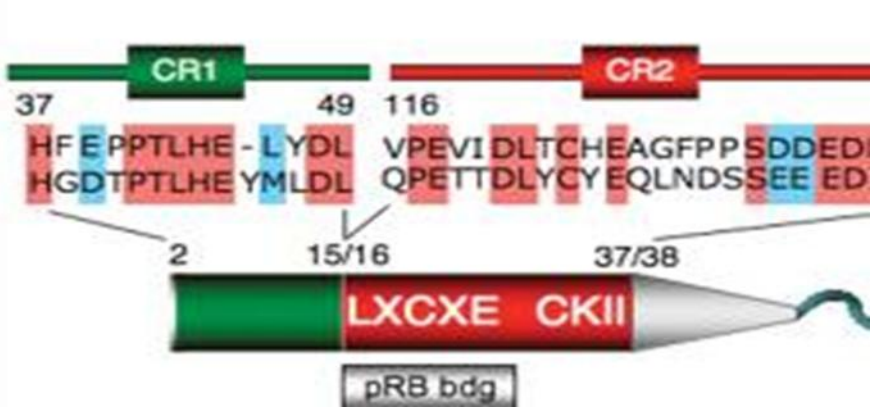


Figure 4: Diagram of the E7 Protein and the Location of its Conserved Gene Regions

- **Functions of the L1 and L2 genes**

The L1 and L2 genes are two genes located in the late coding region, encoding the major and minor capsid proteins. The L1 gene is considered the most conserved region of the virus and is used for detection as well as a criterion in the classification of Papillomaviruses. Inside the L1 gene is a double-stranded, circular DNA chromosome, linked to histones, and the L1 protein has a size of 55 kDa. When L1 interacts with the specific receptor on the cell surface, heparan sulfate proteoglycans (HSPG), it stimulates the initiation of HPV entry into the host cell [29]. This leads to the degradation of amino acids on the L2 capsid by furin-a cellular protease-thereby helping L2 be released from endosomal phagocytosis [30]. The cleavage of L2 by furin facilitates the virus's attachment to a second receptor on the epithelial cell surface [31]. Furthermore, L2 plays a crucial role in escaping local immune system

control. It reduces the phosphorylation of Protein Kinase B (Akt), decreasing Akt activity, which subsequently affects the PI3K-Akt pathway, a key process of Langerhans cells, helping HPV escape immune system surveillance. Thus, it can be seen that the L1 protein is involved in identifying and binding to receptors on the cell surface. The binding of HPV L1 to the receptor activates immune cells and cleaves some amino acids of the L2 protein. L2 inhibits immune cells, providing a favorable condition for HPV to survive and enter the host cell.

2.2.3. HPV Classification

- **Classification based on nucleotide sequence similarity of E6, E7, and L1 genes**

Some viruses, such as hepatitis, HIV, and others, are taxonomically classified based on serology. For HPV, genotype determination is based on the degree of similarity in the nucleotide composition and the degree of homology between the amino acid components of the E6, E7, and L1 gene sequences [32]. A genotype is considered new if at least 10% of its E6, E7, L1 gene regions differ from previously known genotypes. A subtype within a genotype is categorized as a new subclassification when its genome has a difference of 2-10% compared to the known genotype. If the subtype has a difference of 1-2% in the coding region or 5% in the non-coding region, it is called a variant [33]. According to a study by Chen et al. in 2011, a classification criteria set for HPV was published, specifically: if the difference in the entire HPV genome is 1-10%, it is called a lineage, and it is called a sublineage if the difference is 0.5-1% [34]. In 2013, Bzhalava et al. successfully constructed an HPV classification tree for 198 HPV types based on the L1 gene, detailed in Figure 5 [35].

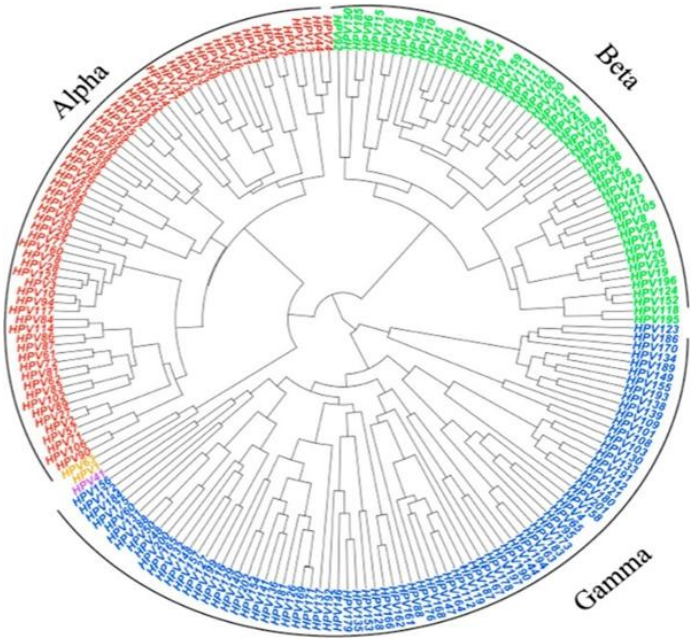


Figure 5: Phylogenetic tree constructed based on the L1 gene sequence of 198 HPV types

HPV is categorized into 16 genera, named after the letters of the Greek alphabet α- π. HPV is a member of five specific genera: alpha (α)-, beta (β)-, gamma (γ)-, mu (μ)-, and nu (ν)-papillomavirus. Each genus possesses its own specific characteristics and biological features, which are detailed in Table 1.

Table 1: Characteristics of HPV types

Genus	Characteristics
Alpha - papillomavirus	+ Primarily infects the mucous membranes of the mouth or anogenital region in humans and primates. + Recent epidemiological data suggest a closer association between specific types and HR types. + Divided into LR and HR types. HR types: HPV-16, -18, -31, -33, -35, -39, -45, -51, -52, -56, -58, and -59. LR types: HPV-6, -11, -40, -42, -43, -44, -54, -61, -70, -72, -81.
Beta - papillomavirus	+ Includes HR and LR skin lesions in humans.

	+ Infections exist in a latent form in the general population, activated under conditions of immunosuppression. + Associated with the development of nasopharyngeal cancer (or pharyngeal cancer). + Often associated with asymptomatic infection in immunocompetent hosts, but can proliferate in immunosuppressed hosts and in patients with *Epidermodysplasia Verruciformis* (EV). Also referred to as EV-HPV types due to the strong relationship with EV disease. + Persistent infection is considered a risk for the development of skin cancer, especially in immunocompromised individuals. + Common types: HPV-5, -9.
Gamma-papillomavirus	+ Causes skin lesions in humans- commonly found in skin and oral mucosa. + Histologically distinguishable by species- specific cytoplasmic inclusion bodies. + Common types: HPV-50, -60, -88.
Mu- papillomavirus	+ Causes benign skin lesions—histologically distinguishable by typical species-specific intracytoplasmic inclusion bodies. + Not associated with cancer. + Types: HPV-1, -63, -204.
Nu - papillomavirus	+ Causes benign and malignant skin lesions. + DNA detected in skin cancer. + Type: HPV-41.

- **Classification of HPV based on its oncogenic potential**

Another method to classify HPV is based on their pathogenicity or oncogenic potential. To date, over 200 genotypes have been identified, of which approximately 40 types are transmissible through sexual contact. HPV is divided into three groups: LR-HPV, HR-HPV, and oncogenic HPV [36] (Table 1). LR- HPV causes genital and skin warts, while HR-HPV causes cancers of the head and neck (including the oral cavity, tonsils, and throat) and anogenital cancers (including cervix, anus, vulva, vagina, and penis). HPV-16 and -18 are the two most common types causing HR-HPV- related malignancies, accounting for approximately 70% of cervical cancer cases globally. Studies show that HPV-6 and -11 are the most common LR-HPV types [37]. HPV invades the mucosa and is often found in the genital organs of both sexes. For example, HPV-6 and -11 are specifically detected in genital warts, whereas HPV-16 and HPV-18 are found in anogenital cancer cells.

- **Classification of HPV based on disease site (tropism or adaptation of HPV to target cells)**

Based on the site of infection, HPV is divided into three groups:

(1) Keratinized Epithelium-adapted HPV Group: HPV in this group is capable of infecting the skin, forming common warts (HPV-2, -4, -26, -27, -29, -57), flat warts (HPV-1, -2, -4), and Butcher's warts (HPV-7). Lesions typically appear on the face, neck, hands, and feet.

(2) Non-genital Mucosal Cell-adapted Group: This includes HPV capable of causing disease in the oral and oropharyngeal mucosa (HPV-6, -11, -13, -32), leading to Recurrent Respiratory Papillomatosis (RRP).

(3) Genital Mucosal Cell-adapted Group: This HPV group causes diseases in the genital tract, such as genital warts (HPV-6, -11, -42, -43, -44, -54), cervical cancer, penile cancer, vulvar cancer, and vaginal cancer.

Thus, HPV can be classified according to various criteria depending on the dominant characteristic of that criterion. From this, it can be seen that HPV is a virus type with high diversity.

2.2.4. Pathobiological characteristics of HPV

- **Pathogenesis of HPV**

Sexual contact is the primary route of HPV transmission, and males play a significant role in this epidemiological chain, serving as both a source carrier and a source of transmission. The HPV virus is relatively stable and can remain infectious for a long time in humid environments. HPV often develops into a chronic, persistent infection. The general pathological feature of these disorders is epithelial dysplasia, encompassing intraepithelial neoplasia grades 1, 2, and 3, which are comparable to mild, moderate, and severe degrees. In cytology, low-grade squamous intraepithelial lesions (LSIL) are considered Grade 1 intraepithelial neoplasia, with only minor changes in the differentiation pattern. These lesions are typically eliminated by the immune system in less than a year. High-grade squamous intraepithelial lesions (HSIL) cause Grade 2 and 3 intraepithelial neoplasia. Grade 3 intraepithelial neoplasia can be defined as carcinoma in situ. Malignant progression of HR-HPV-related lesions typically occurs when other threats are present, such as an inadequate immune response and/or prolonged latency following genetic abnormalities in the host cell DNA [38, 39] (Figure 6).

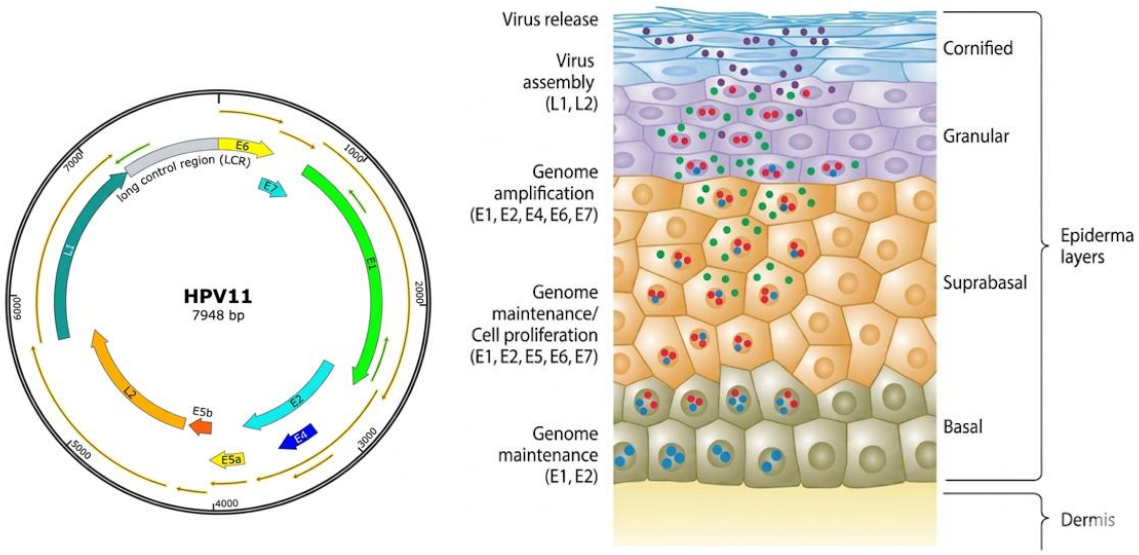


Figure 6: The human papillomavirus (HPV) life cycle and the stroma.

- **Mechanism of pathogenesis**

HPV is primarily transmitted through sexual contact. Replication of the HPV genome occurs in the nucleus of the host cell and proceeds according to the following steps:

(1) HPV genome entry into the host cell chain: For the LR- HPV group, the HPV genome enters the host genome chain in an episomal form (circular DNA outside the host chromosome). In this form, the E2 gene coding region is not altered. However, the concentration of E2 protein increases along with the HPV DNA replication process, leading to cell proliferation and inhibition of E6 and E7 activity. The inhibition of E6 and E7 leads to the activation of the p53 pathway and the inhibition of pRB tumor formation. Consequently, this results in the proliferation of a large number of cells, but still under the control of p53 and pRB. When the HPV gene integrates into the host chromosome, it disrupts the E2 gene and releases the inhibitory control over E6 and E7 activity [40].

(2) Cell immortalization: E6 and E7 proteins from HR- HPV genotypes also have the ability to combine with ras. The ras protein is an intracellular signaling molecule; when ras is activated, it promotes cell growth, differentiation, and survival. The gene coding for the ras protein is considered the first oncogene discovered. The mechanism by which the E6 protein causes cell immortalization is evidenced by its ability to inactivate p53, express hTERT (human telomerase reverse transcriptase), and increase telomerase activity [40].

(3) Host cell genomic instability: Abnormal mitosis can be caused by the E6 and E7 proteins of "high-risk" genotypes but not typically seen in "low-risk" genotypes, leading to the loss of alleles in certain genes, where these 17 genes are related to the onset and progression of cancer. E6 causes genomic instability due to its ability to inhibit p53 function, resulting in the disruption of normal DNA repair processes and consequently causing gene mutation. E7 causes genomic instability through the inactivation of pRB and also by its ability to affect centrosome synthesis, consequently causing alterations in DNA segregation during cell division [41].

(4) Altered response to DNA damage: The E6 and E7 genes can cause the body to lose its ability to respond to DNA damage. When DNA damage occurs, the body responds by activating p53, which produces a protein that regulates the resting period between two cell cycles. E6 and E7 have the ability to inhibit the cell cycle arrest process regulated by p53. E6 only binds to and inactivates p53, but E7 not only causes

dysfunction of the cell cycle regulator pRb but also inactivates p21, an inhibitor of the cyclin-dependent kinase enzyme, which is a necessary factor produced by p53 activation [42].

(5) Cell proliferation and differentiation: HPV replicates in coordination with the differentiation process of basal cells in the form of episomes. Simultaneously, it replicates within suprabasal cells that have exited the host cell cycle, thanks to the role of HPV E6 and E7 in reprogramming and maintaining DNA synthesis in infected keratinocytes [43].

2.2.5. Transmission routes and HPV prevention methods

- **Transmission routes:**

HPV can be transmitted directly through the skin and mucous membranes from an infected individual to a healthy person, with sexual transmission being the most common. Both homosexual and heterosexual activities are primary causes of direct HPV transmission via the genital, oral, and anal routes.

HPV can also be transmitted from skin-to-skin, skin-to-mucosa, or mucosa-to-mucosa through wart secretions, saliva, or contaminated personal items such as towels and clothing carrying the virus from an infected individual.

Furthermore, HPV is transmitted from mother to child during the perinatal period. HPV-infected secretions from the maternal genital tract are directly transmitted to the neonatal ocular, oral, and respiratory mucosa. This is a known cause of persistent respiratory conditions due to HPV, such as Recurrent Respiratory Papillomatosis (RRP).

- **Prevention methods:**

Condom use: Unprotected sexual intercourse is the primary transmission route for HPV. Men serve as both a reservoir and an intermediate source of infection, making them a critical component of the HPV epidemiological chain [44]. Therefore, the use of condoms during sexual activity is essential to prevent HPV transmission.

Smoking cessation: Smoking is a known independent risk factor for HPV infection. A study has shown that the prevalence of HPV detected in men was 68.2% for smokers compared to 63.2% for non-smokers. This indicates that men with a history of smoking are more likely to be infected with HPV than those who do not smoke.

Consequently, timely smoking cessation can play a particularly vital role in HPV prevention among men.

Male circumcision (MC): Numerous studies have confirmed the role of MC in reducing the prevalence of various HPV strains. By limiting the risk of viral entry through epithelial micro-lesions, this procedure helps diminish both viral shedding and the persistence of the virus within the body. Consequently, circumcision is considered one of the leading effective interventions to eliminate oncogenic HPV types and prevent associated infections [45].

HPV Vaccine: To date, a total of 107 countries has implemented HPV vaccination programs. However, HPV vaccination is not solely a female issue; men are also at risk for various diseases caused by HPV infection. Furthermore, HPV vaccination in men provides significant benefits to women by reducing the risk of cervical cancer. Therefore, the HPV vaccine is undoubtedly necessary and should be encouraged for both males and females.

Asymptomatic HPV infection in men is considered a major cause of continuous transmission to female partners. For this reason, HPV infection in men increases the risk of cervical cancer in women. HPV infection also poses a risk of genital warts, penile cancer, and anal cancer in men. Consequently, HPV screening for men is regarded as highly essential.

2.2.6. The role of HPV whole-genome sequencing research

Previously, traditional methods for identifying and classifying HPV variants mostly relied on the amplification and sequencing of a short fragment within the L1 or E6/E7 gene regions [46]. However, this approach exhibits several limitations in the era of precision medicine. For HPV, particularly LR types like HPV-11, researching the whole-genome sequence (WGS) of HPV offers powerful opportunities for applications in epidemiology, public health, and clinical diagnostics [47].

First, from a molecular epidemiology perspective, a novel HPV type is defined by an L1 gene sequence diversity of more than 10% compared to established types. At the intra-type variant level, WGS serves as the unique tool to precisely distinguish lineages (differing by 1.0% to 10% across the entire genome) and sublineages (differing by 0.5% to 1.0% across the entire genome). This is based on the reality that strains of the same type are closely related and their sequence heterogeneity is best

synthesized across the complete genome; newly evolved variant genomes exhibit mutations that are not always uniformly distributed throughout the genome [48]. Furthermore, whole-genome sequencing also assists in elucidating the association between single nucleotide polymorphisms (SNPs) or specific point mutations in geographically dependent genes and the severity of clinical phenotypes. Finally, the accumulation of HPV WGS data alongside research on the pathogenesis of HPV11 in international databases serves as a core foundation for optimizing the efficacy of current multivalent vaccines such as Gardasil-4 and Gardasil-9 [49]. Concurrently, it provides highly accurate sequence information for the design of primers and probes in next-generation molecular diagnostic assays, thereby minimizing false-negative results.

CHAPTER 3. MATERIALS AND METHODS

3.1. Materials

Table 2: Biological products and chemicals

No	Biological Products and Chemicals
1	Swabs (FLOQSwab R100 and FLOQSwab U80, Copan Diagnostics Inc., Murrieta, CA)
2	HPV Collection Medium (Inactivated Sample Transport Kit) (ABT, Vietnam)
3	Toppure® Genomic DNA Extraction Kit (ABT, Vietnam)
4	FIREPOL® MASTER MIX WITH 7.5 MM MGCL2, 5X (Solis BioDyne, Tartu, Estonia)
5	6X GelRed Loading Buffer with TriColor (ABT, Vietnam)
6	BigDye® Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher, USA)
7	Agarose (ABT, Vietnam)

3.2. Equipment and instruments

Table 3: Equipment and instruments

No	Equipment
1	Biosafety Level 2 Laboratory (BSL-2 Lab)
2	Biosafety cabinet (BSC)
3	PCR cabinet (or Molecular Biology cabinet)
4	Benchtop Centrifuge (Eppendorf, Germany)
5	Mupid exU 17 Electrophoresis System (Takara Bio, Japan)
6	PCR Cycler (Eppendorf, Germany)
7	UVP Transilluminator (Gel Documentation System) (Analytik Jena)
8	-80°C Ultra-low Temperature Freezer (Eppendorf, Germany)
9	-20°C Freezer
10	2-8°C Refrigerator (or 2-8°C Fridge)
11	Autoclave (Sterilizer)
12	Pipettes and Pipette Tips from 0.1 µL – 1000 µL (Eppendorf, Germany)
13	Sterilized PCR tubes, Eppendorf tubes, and Falcon tubes

3.3. Research methods

3.3.1. Sample collection and storage

Cell samples or lesions from the genital area (papillomas) were collected using nylon-flocked swabs (FLOQSwab R100 and FLOQSwab U80, Copan Diagnostics Inc., Murrieta, CA). The specimens were then preserved in the viral inactivation transport medium, HPV Collection Medium (ABT, Vietnam), and stored at -80°C in an ultra-low temperature freezer (Eppendorf, Germany) until use.

3.3.2. DNA extraction

The HPV DNA extraction process was performed using the Toppure® Genomic DNA Extraction Kit (ABT, Vietnam). First, a tissue sample weighing no more than 0.5 g was placed into a 1.5 mL tube, followed by the addition of 200 µL of Tissue Lysis (TL) buffer and 20 µL of Proteinase K. The sample was vortexed, incubated at 72°C for 10–15 minutes, and then homogenized using a pestle until completely dissolved. Next, a mixture containing 200 µL of the sample, 200 µL of Cell Lysis (CL) buffer, and 20 µL of Proteinase K was thoroughly mixed and incubated at 72°C for 20 minutes. The entire mixture was transferred to a silica column, supplemented with 200 µL of ethanol, and centrifuged at 13,000 rpm for 1 minute. The column was retained, while the supernatant was discarded. The column washing step was carried out by sequentially adding 500 µL of Wash buffer 1 (WB1) and 500 µL of Wash buffer 2 (WB2), with each step followed by centrifugation at 13,000 rpm for 1 minute. After drying the membrane by centrifuging for an additional minute at the same speed, the column was transferred to a new 1.5 mL collection tube. To elute the purified DNA, 50 µL of Elute buffer (EB) was applied directly to the center of the membrane, and the column was centrifuged at 13,000 rpm for 1 minute. The final target DNA product was stored in a -20°C freezer for subsequent experiments.

3.3.3. HPV diagnostic method based on detection Primer Pair MY11/MY09

The L1 gene segment of HPV virus was amplified using the detection primer pair My11/My09 with a length of 450 bp and the thermal cycle was standardized. The procedure and primer pair sequences are described in **Table 4** below:

Table 4: Nucleotide sequence of primer pair MY11/MY09

Primer	Nucleotide Sequence (5'-3')	Size (bp)
MY11(+)	GCMCAGGGWCATAAYAATGG	450 bp
MY09(-)	CGTCCMARRGGAWACTGATC	

Table 5: Components and thermal cycling conditions for HPV L1 gene amplification reaction using MY11/MY09 primers

Components	Reaction Volume	Stage	Temperature	Time
MY11	0.5 µl	Denaturation	95°C	5 minutes
MY09	0.5 µl	Annealing	95°C 55°C 72°C	1 minute 1 minute 30 seconds 1 minute
Master Mix	4 µl	Extension	72°C	10 minutes
DEPC water	13 µl	Storage	4°C	

(My11: forward primer; My09: reverse primer; M: A/C; R: A/G; W: A/T; Y: C/T)

3.3.4. Method for designing primers for whole-genome amplification of HPV-11

- Step 1: Library preparation

- Collect reference strains for HPV-11 from the NCBI GenBank database.

- These sequences are compiled into a library for use in analyzing and identifying conserved regions.

- Step 2: Analysis of target region characteristics

- The full HPV-11 genome consists of nine DNA segments, including the genes: L1, L2, E6, E7, E1, E2, E4, E5α and E5β.

- Determine the location and length of the target amplification segments based on the genome sequence.

- Step 3: Primer design

- Utilize the Primer-BLAST software to design specific primer pairs that amplify the complete gene sequences of HPV-11.

- Based on several mandatory criteria, such as: length, temperature (melting temperature), G-C ratio, etc...

- Step 4: Experimental validation and optimization

- Use the BLAST tool to check the specificity of the primers across the entire HPV genome and other genomes, ensuring they do not bind to unintended sites.

3.3.5. Polymerase chain reaction (PCR)-based amplification method

After extraction, the obtained DNA will be amplified using the PCR method with specific primer pairs and the corresponding thermal cycling program. The composition of the PCR reaction is described in Table 6.

Table 6: Composition of the PCR Reaction

Components	Volume
Forward primer	0.5 µl
Reverse primer	0.5 µl
Master mix	4 µl
DEPC water	13 µl
Total	18 µl

The thermal cycling program for the PCR reaction is as follows:

Table 7: Temperature and Duration in the PCR Reaction

Stage	Temperature (°C)	Duration	Cycles
Initial Denaturation	95	5 min	1
Denaturation	95	1 min	30
Annealing	56	1 min	
Extension	72	1 min	
Final Extension	72	1 min	1
Hold	4	10 min	1

3.3.6. Agarose gel electrophoresis

The PCR products were stained with 6X GelRed Loading Buffer with TriColor (ABT, Vietnam), and then separated by electrophoresis on a 1.5% agarose gel (ABT, Vietnam) using the Mupid exU 17 electrophoresis system (Takara Bio, Japan). The electrophoresis results were visualized using a UVP Transilluminator (Analytik Jena).

3.3.7. Sanger sequencing

PCR products with clear bands, and a length corresponding to the pre-determined size (prior to primer design), will be used as templates for Sanger sequencing. Sanger sequencing was performed at Apical Scientific Sdn. Bhd. (Firstbase, Malaysia), using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher, USA) on an ABI PRISM 3730xl Genetic Analyzer (Applied Biosystems, USA). Each sequence segment was analyzed in both the forward and reverse directions, and any nucleotide discrepancies will be compared against the reference sequence in the NCBI database.

3.3.8. Nucleotide sequence analysis

The contiguous DNA sequences (contigs) obtained from forward and reverse sequencing were assembled and edited using Segman 5.0 software (DNASTAR, Madison, WI, USA). The NCBI BLAST tool was used to compare and align the study sample's sequence with the published sequence library of various types on GenBank. The processed sequence data of the HPV type and the data of other types were analyzed using Mega 6.0 software.

3.3.9. Phylogenetic tree construction

The complete nucleotide sequences of each gene from the study sample and the reference sequence library on NCBI were used for phylogenetic analysis. The first gene sequence found for each genotype (the prototype sequence) was selected for comparison with the sequences being studied in Vietnam. Information regarding the gene sequences of the selected genotypes was classified into lineages and sublineages for the purpose of comparison and phylogenetic tree construction in Vietnam. The sequences under investigation in Vietnam were aligned with the prototype sequence and the corresponding lineage and sublineage sequences of each genotype. The phylogenetic tree was constructed using MEGA 6.0 software based on the maximum likelihood method with a Bootstrap value of 1000 replicates, and was built based on the differences in single nucleotide polymorphisms (SNP).

3.4. Ethical considerations

The clinical specimens used in this study were obtained with the consent of the Department of Andrology at Hanoi Medical University Hospital and with the formal approval of the patients. Patients reserved the right to refuse specimen collection and to decline participation in the study. All patient information was kept strictly confidential.

CHAPTER 4. RESULT AND DISCUSSION

4.1. Detecting HPV DNA in samples through PCR results

The PCR amplification using the consensus primer set MY11/MY09 successfully yielded distinct and highly specific DNA fragments of approximately 450 bp in the positive channels for clinical samples 175 and 182 (Figure 7). This size aligns precisely with the expected amplicon target of the highly conserved L1 open reading frame (ORF) of mucosal HPV genotypes. The high intensity and sharpness of the bands on the 1.5% agarose gel indicate a robust viral load and efficient amplification kinetics without noticeable non-specific product formation. Notably, the corresponding negative control lanes (-) for both samples showed a complete absence of bands or primer-dimer interference. This confirms the high specificity of the primer set under the optimized thermocycling conditions and rules out potential cross-contamination during the DNA extraction and PCR preparation stages. Consequently, these results provide a reliable and contamination-free molecular foundation, confirming the presence of HPV DNA in samples 175 and 182, which validates them for subsequent downstream genotyping and sequence analysis.

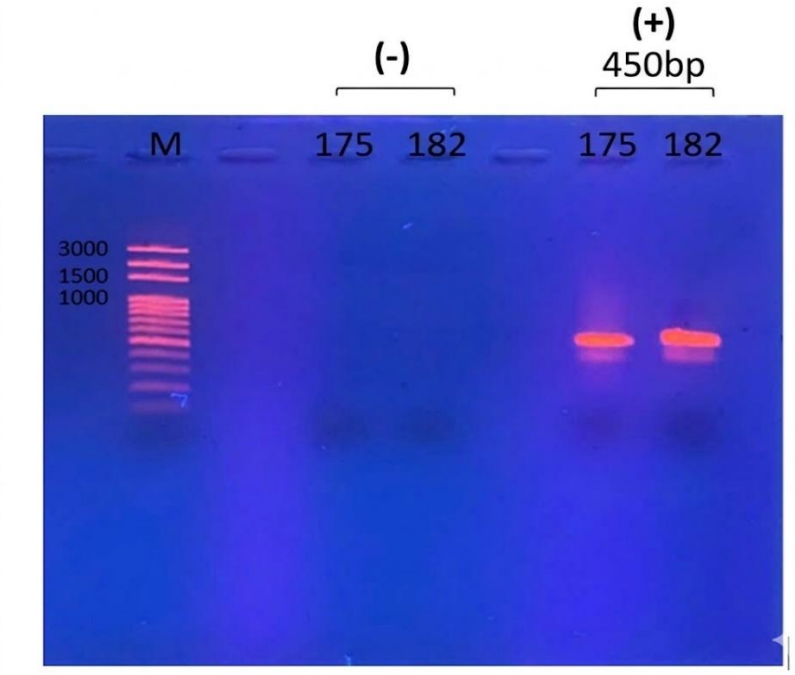


Figure 7: Electrophoresis results of PCR products for HPV diagnosis

M: 100bp marker; (+): Amplification of HPV-11 gene using primer pair MY11/MY09; 175, 182: Specimen containing HPV-11 strain.

4.2. The amplification results for the HPV-11 genes used in the study

4.2.1. Information about the primers used in the amplification reaction

The computational design and experimental validation of the seven primer pairs (Table 8) demonstrate an exceptionally robust and strategic approach for the complete physical coverage of the circular double-stranded DNA genome of HPV-11 (about 8000 bp). An evaluation of the primer design highlights several key strengths that contributed directly to the high-quality experimental outcomes:

- Optimized overlapping strategy: The physical boundaries of the designed amplicons (ranging from 1107 bp to 1405 bp establish a highly efficient overlapping layout. For instance, the amplicon generated by F1/R1 (7798 bp to 1076 bp) seamlessly overlaps with the F2/R2 product (911 bp to 2315 bp). This continuous overlapping architecture across all seven fragments ensures that no genomic junctions or highly variable regions are lost during subsequent sequence assembly, which is a critical requirement for accurate comparative genomics of HPV-11.
- Uniform and optimized product sizes: The designed amplicon lengths are strictly confined within a narrow and highly uniform range, varying only from 1107 bp to 1405 bp. In whole-genome PCR strategies, maintaining highly consistent fragment sizes is critical to prevent preferential amplification of shorter fragments, which often leads to uneven yields across different genomic regions. By standardizing the amplicon sizes, we successfully synchronized the extension rate and elongation duration of the Taq DNA polymerase across all seven reactions. This uniformity guarantees balanced, high-yield amplification for every single genomic segment under identical extension time parameters, while also streamlining the downstream agarose gel analysis and subsequent multiplex purification steps.
- Synchronized melting temperatures (T_m) for single-plate run: All primer pairs exhibit outstanding thermodynamic homogeneity, with their melting temperatures (Tm) tightly clustered between 58°C and 61°C, and predominantly centered at 60°C. Minimizing (Tm) variance to less than 3°C is a major technical advantage; it allows all seven independent PCR assays to be executed simultaneously on a single 96-well plate using a unified, single-step annealing temperature (typically set at 55-58°C). This highly synchronized thermal profile prevents the common pitfalls of PCR

multiplexing, such as non-specific amplification at lower temperatures or failure of primer annealing at higher temperatures. Consequently, this design dramatically reduces experimental runtime, minimizes well-to-well pipetting variations, and ensures highly reproducible, high-throughput amplification of the entire HPV-11 genome.

Collectively, these results demonstrate that the newly designed primer set is highly reliable, specific, and fully optimized for down-stream whole-genome sequencing and phylogenetic analysis of local HPV-11 strains.

Table 8: Information on primers for HPV-11 genomic amplification

Primer name	Position	Sequence (5'-3')	Product size (bp)	T_m (°C)
F1-HPV-11	7798-7817	TCGGTTACCCACACCCTACA		60
			1300	
R1-HPV-11	1076-1048	TGGCGTCTAACCGTGGACTT		61
F2-HPV-11	911-930	GTGGAAGCACAGGCATTGTT		59
			1405	
R2-HPV-11	2315-2296	ATGGTTGTGTGGCATCATCC		58
F3-HPV-11	2168-2187	CCTGACACTGGGAAGTCGTG		60
			1203	
R3-HPV-11	3370-3351	AGGGTGTTGTTAGTGGACGG		60
F4-HPV-11	3255-3274	ACTACATACACCCCCGCACA		60
			1353	
R4-HPV-11	4607-4590	GGGCAACAGGCTCCACAA		60
F5-HPV-11	4547-4566	AAGCCTGCTATTACTGGGGG		60
			1178	
R5-HPV-11	5724-5705	TGGATACAGGGTTGGGAGGA		60
F6-HPV-11	5788-5807	GACTCCTTGCTGTGGGACAT		60
			1373	
R6-HPV-11	7160-7141	GGTACGTTTTCGTTTGGGGG		60
F7-HPV-11	6937-6956	CCTGTCAGAAACCCACACCT		60
			1107	
R7-HPV-11	110-91	GCATTCCTGCAAAACACGCA		60

4.2.2. Results of gene amplification using PCR reaction

After identifying HPV-positive samples using the PCR method, one sample confirmed as HPV-11 was selected for PCR amplification of the complete genome sequence, using the primer pairs designed as described in Table 7.

The positive control sample (+), used the MY11/MY09 detection primer pair to identify the presence of HPV viral DNA with a length of approximately 450 bp, all bands of the positive control appeared. The negative control sample did not show any bands, ensuring that the product was not contaminated by foreign matter during the PCR process. The estimated length of the HPV-11 genome is over 7 kbp; therefore, seven DNA fragments were amplified using specific primer pairs, and the expected PCR product sizes are as follows: HPV-11F1/R1 is 1300 bp, HPV-11F2/R2 is 1405 bp, HPV-11F3/R3 is 1203 bp, HPV-11F4/R4 is 1353 bp, HPV-11F5/R5 is 1178 bp, HPV-11F6/R6 is 1373 bp, HPV-11F7/R7 is 1107 bp. Amplification results showed that the designed primer pair had high sensitivity and specificity for the experimental sample. The PCR products were identified on a 1.5% agarose gel with sizes approximately equivalent to the expected sizes (Figure 8).

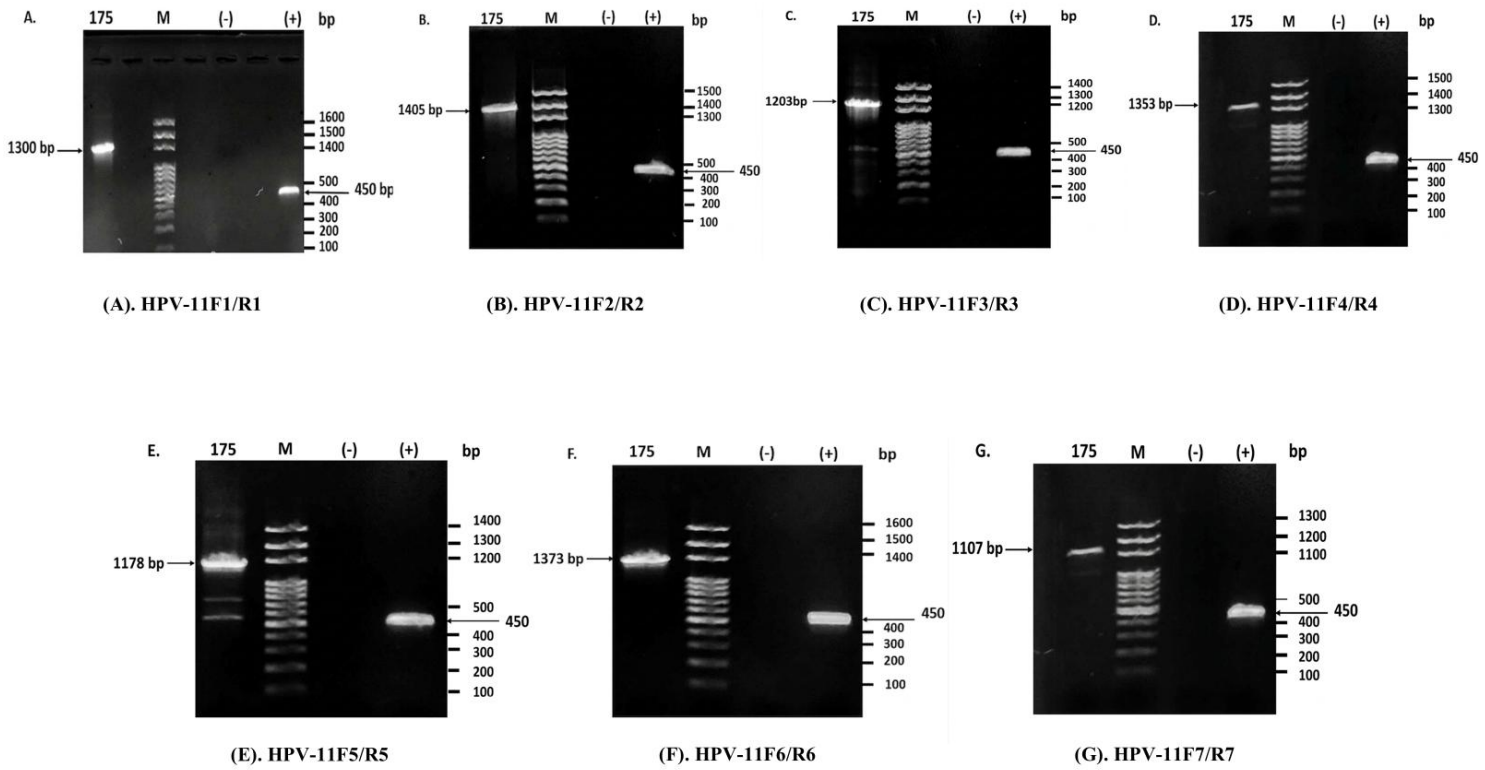


Figure 8: PCR product of HPV-11 gene amplification

M: 100bp marker; (+): Amplification of HPV-11 gene using primer pair MY11/MY09; 175: Specimen containing HPV-11 genotype. From left to right, top to bottom, the amplified PCR products with primer pairs are shown: (A). HPV-11F1/R1, (B). HPV-11F2/R2, (C.) HPV-11F3/R3, (D). HPV-11F4/R4, (E). HPV-11F5/R5, (F). HPV-11F6/R6, (G). HPV-11F7/R7

4.3. Results of functional gene sequencing analysis

4.3.1. L1 gene

The full length of the gene encoding the L1 protein of the studied HPV-11 strain is 1506 bp, and the L1 protein has a size of 501 amino acids. These genetic analyses indicated that the L1 gene of our HPV-11 strain has a high degree of nucleotide similarity (ranging from 99.1-99.9%) to the reference strains (Table 9).

Specifically, one strain shared the highest nucleotide identity with our study strain (99.9%): FN907962. Strains with a 99.8% similarity level include JQ773409, HE611263, JN644141, LN833185, MK313763, MN788368, LN833184, JQ773408, MK463914, FR872717, JQ773411, JQ773412, and MK463921. Those at 99.7% include MK313765, MK313767, FN870021, M14119, KU298879, MK463916, and FN907963, while those at 99.6% include LN833161, HE574702, LN833165, and LN833169. Additionally, EU918768 and LN833187 exhibited identities of 99.5% and 99.4%, respectively. The lowest identity was observed in strain LN833183 at 99.1%.

Table 9: Nucleotide sequence identity of the HPV-11 L1 gene with reference strains

IS.HPV11.24_81																																						
HE574702/JO-RRP_2/A2	99.6%																																					
EU918768/LZod45/2009/C	99.5%	99.6%																																				
FN870021/A86/A3	99.7%	99.8%	99.8%																																			
MK313767/CAC1/A2	99.7%	99.8%	99.6%	99.8%																																		
LN833184/A2	99.8%	99.8%	99.7%	99.9%	99.9%																																	
MK313765/JO-RRP8/A2	99.7%	99.8%	99.6%	99.8%	100.0%	99.9%																																
MN788368/JO-RRP10/A2	99.8%	99.8%	99.7%	99.9%	99.9%	100.0%	99.9%																															
JQ773409/CU17/2012/Tha	99.8%	99.8%	99.7%	99.9%	99.9%	100.0%	99.9%	100.0%																														
MK463921/HPV11-gv-1110	99.8%	99.7%	99.6%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%																													
FR872717/2011/Hungary/I	99.8%	99.7%	99.6%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%																												
FN907963/CS20/2011/Slo	99.7%	99.8%	99.8%	100.0%	99.8%	99.9%	99.8%	99.9%	99.9%	99.8%	99.8%																											
HE611263/LP220/A2	99.8%	99.8%	99.7%	99.9%	99.9%	100.0%	99.9%	100.0%	100.0%	99.8%	99.8%	99.9%																										
FN907962/CAC86	99.9%	99.7%	99.6%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%																									
MK463916/HPV11-gv-110	99.7%	99.6%	99.5%	99.7%	99.7%	99.8%	99.7%	99.8%	99.8%	99.9%	99.8%	99.7%	99.8%	99.8%																								
JQ773411/CU19/A2	99.8%	99.7%	99.6%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	100.0%	99.8%	99.8%	99.8%	99.8%																							
KU298879/2016/Brazil/83	99.7%	99.8%	99.6%	99.8%	99.8%	99.9%	99.8%	99.9%	99.9%	99.8%	99.8%	99.8%	99.9%	99.8%	99.7%	99.8%																						
JQ773408/CU16/A2	99.8%	99.8%	99.6%	99.8%	99.8%	99.9%	99.8%	99.9%	99.9%	99.9%	99.9%	99.8%	99.9%	99.9%	99.8%	99.9%	99.8%																					
LN833161/A2	99.6%	99.7%	99.6%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.7%	99.7%	99.8%	99.8%	99.7%	99.6%	99.7%	99.8%	99.8%																				
MK463914/HPV11-gv-110	99.8%	99.7%	99.6%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.9%	99.7%																			
LN833185/A2	99.8%	99.8%	99.7%	99.9%	99.9%	100.0%	99.9%	100.0%	100.0%	99.8%	99.8%	99.9%	100.0%	99.8%	99.8%	99.8%	99.9%	99.9%	99.8%	99.8%																		
LN833165/A4	99.6%	99.6%	99.6%	99.8%	99.7%	99.8%	99.7%	99.8%	99.8%	99.6%	99.6%	99.8%	99.8%	99.6%	99.6%	99.6%	99.7%	99.7%	99.6%	99.6%	99.8%																	
MK313763/JO-RRP2/A2	99.8%	99.8%	99.7%	99.9%	99.9%	100.0%	99.9%	100.0%	100.0%	99.8%	99.8%	99.9%	100.0%	99.8%	99.8%	99.8%	99.9%	99.9%	99.8%	99.8%	100.0%	99.8%																
M14119/1986/Germany/A1	99.7%	99.8%	99.8%	100.0%	99.8%	99.9%	99.8%	99.9%	99.9%	99.8%	99.8%	100.0%	99.9%	99.8%	99.7%	99.8%	99.8%	99.8%	99.8%	99.8%	99.9%	99.8%	99.9%															
JN644141/GUMC-AJ/2013/	99.8%	99.8%	99.7%	99.9%	99.9%	100.0%	99.9%	100.0%	100.0%	99.8%	99.8%	99.9%	100.0%	99.8%	99.8%	99.8%	99.9%	99.9%	99.8%	99.8%	100.0%	99.8%	100.0%	99.9%														
LN833169/A3	99.6%	99.6%	99.6%	99.8%	99.7%	99.8%	99.7%	99.8%	99.8%	99.6%	99.6%	99.8%	99.8%	99.6%	99.6%	99.6%	99.7%	99.7%	99.6%	99.6%	99.8%	100.0%	99.8%	99.8%	99.8%													
LN833187/A4	99.4%	99.5%	99.5%	99.7%	99.6%	99.6%	99.6%	99.6%	99.6%	99.5%	99.5%	99.7%	99.6%	99.5%	99.4%	99.5%	99.6%	99.6%	99.5%	99.5%	99.6%	99.8%	99.6%	99.7%	99.6%	99.8%												
JQ773412/CU20/A2	99.8%	99.7%	99.6%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	100.0%	99.8%	99.8%	99.8%	99.8%	100.0%	99.8%	99.9%	99.7%	99.8%	99.8%	99.6%	99.8%	99.8%	99.8%	99.6%	99.5%										
LN833183	99.1%	99.3%	99.2%	99.4%	99.2%	99.3%	99.2%	99.3%	99.3%	99.2%	99.2%	99.4%	99.3%	99.2%	99.1%	99.2%	99.2%	99.2%	99.2%	99.2%	99.3%	99.2%	99.3%	99.4%	99.3%	99.2%	99.1%	99.2%									

Following the analysis of the sequencing results, we compared the nucleotide variations of the studied sequence with the prototype sequence and several HPV-11 sequences published in the NCBI GenBank database. The results revealed 16 nucleotide substitution sites, among which 6 specific nucleotide changes led to amino acid alterations, including: 526^{Ser} (A/T), 528^{Ser} (C/T), 716^{Tyr} (G/A), 1061^{Ser} (T/C), 1361^{Glu} (C/A), and 1427^{Ala} (T/C) (Table 10).

At position 526, an A to T substitution was observed compared to the prototype sequence, resulting in an amino acid change from Threonine to Serine. At position 528, a C-to-T change relative to the prototype sequences led to a Threonine to Serine amino acid substitution. At position 716, a G to A variation was recorded in the studied sequence, altering the amino acid from Cysteine to Tyrosine. At position 1061, a T to C transition compared to the prototype sequence converted the amino acid Leucine to Serine. At position 1361, a C to A substitution in the studied sequence resulted in an amino acid alteration from Alanine to Glutamic acid. Finally, position 1427 exhibited a T to C change that substituted the amino acid Valine with Alanine.

Table 10: Distribution of nucleotide and amino acid variations in the L1 gene sequence of the studied HPV-11 strain

Isolate name	Nation/ Year	Type	Nucleotide (amino acid) locations within the L1 gene															
			159	**258**	**303**	**526**	**528**	**716**	**744**	**837**	**1061**	**1140**	**1179**	**1275**	**1361**	**1362**	**1427**	**1467**
M14119	Germany/1986	HPV11/REF	-	C *Ser*	G *Gly*	-	-	-	A *Arg*	C *Asn*	-	-	-	-	-	-	-	-
EU918768	China/2009	HPV11	-	C *Ser*	G *Gly*	-	**-**	**G** ***Cys***	A *Arg*	C *Asn*	-	C *Ser*	-	-	-	G *Glu*	-	-
LN833183	Slovenia 2016	HPV11	G *Lys*	C *Ser*	G *Gly*	**A** ***Thr***	**C** ***Thr***	-	A *Arg*	C *Asn*	**T** ***Leu***	-	A *Ser*	A *Gln*	**C** ***Ala***	-	**T** ***Val***	C *Ser*
IS.HPV-11.24 81	**Vietnam/ 2025**	**HPV-11**	**A** **Lys**	**T** ***Ser***	**A** ***Gly***	**T** ***Ser***	**T** ***Ser***	**A** ***Tyr***	**G** ***Arg***	**T** ***Asn***	**C** ***Ser***	**T** ***Ser***	**T** ***Ser***	**G** ***Gln***	**A** ***Glu***	**A** ***Glu***	**C** ***Ala***	**T** ***Ser***

**Ala: Alanine; Arg: Arginine; Leu: Leucine; Ser: Serine; Asn: Asparagine; Gln: Glutamine; Glu: Glutamic acid; Tyr: Tyrosine; Thr: Threonine; Gly: Glycine; Cys: Cysteine; Val: Valine; Lys: Lysine.*

Phylogenetic analysis based on the full-length L1 gene nucleotide sequence revealed that the study sample from Vietnam (designated IS.HPV-11.24-81-L1) belongs to sublineage A2, supported by a 64% bootstrap value. Within this sublineage, the investigated strain is genetically close to the reference strain FN907962, sharing a major evolutionary cluster with global HPV-11 strains circulating in Thailand (JQ773409), Hungary (FR872717), the USA (JN644141), and Brazil (KU298879). This clustering indicates that, similar to other countries worldwide, sublineage A2 is one of the predominantly circulating sublineages among male patients infected with HPV-11 in Vietnam. Concurrently, the phylogenetic tree demonstrates a clear separation between the study sample belonging to sublineage A2 and other sublineages, such as A1 (prototype strain M14119/Germany), A3 (LN833169), and A4 (LN833165, LN833187), supported by a bootstrap value of 83%. These data confirm the distinct characteristics and relative conservation of the HPV-11 L1 gene, providing an important scientific foundation for evaluating the efficacy of next-generation vaccines in Vietnam (Figure 10).

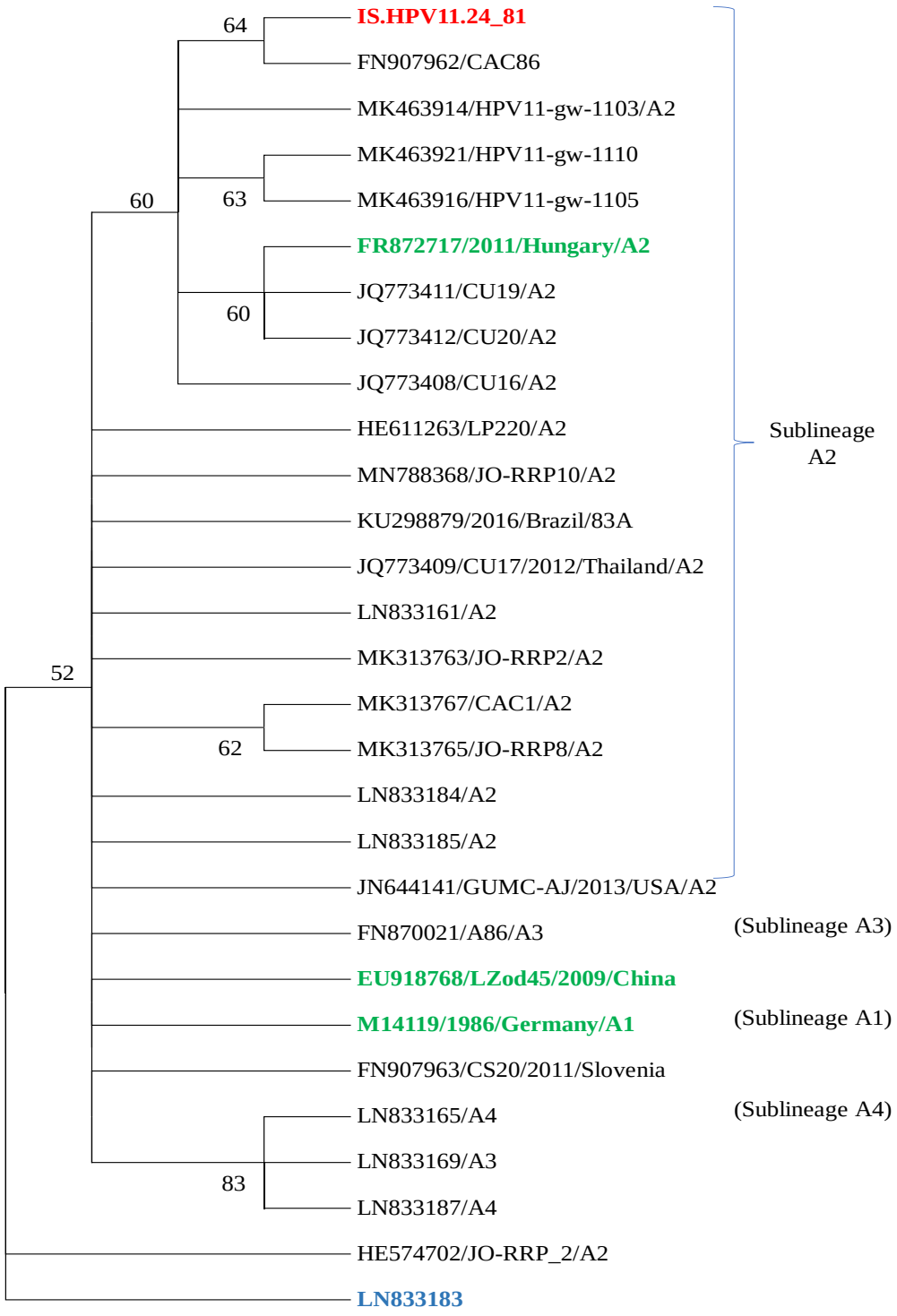


Figure 9: Phylogenetic tree of the L1 gene of HPV-11 strains

4.3.2. L2 gene

The full-length L2 gene of the studied HPV-11 strain is 1368 bp, encoding a 451-amino-acid L2 protein. These genetic analyses indicated that the L2 gene of our studied HPV-11 strain (sample IS.HPV-11.24-81-L2) shared high nucleotide identity with the reference strains, ranging from 98.4% to 100%. Further details are presented in Table 11.

Specifically, the L2 gene of the studied HPV-11 strain shared 100% nucleotide identity with strains JQ773409, LN833161, MK313765, MK313767, HE574702, HE611263, JN644141, LN833185, MK313763, MN788368, LN833184, FR872717, JQ773411, JQ773412, and KU298879, indicating that these sequences are completely identical at the nucleotide level. Furthermore, sample IS.HPV-11.24-81-L2 exhibited a 99.7% similarity level with strains JQ773408 and EU918768. The studied sample also showed 99.5% similarity with strains LN833187 and LN833165, and 99.3% with LN833169. Finally, the lowest similarity (98.4%) was observed in strain LN833183.

Table 11: Nucleotide sequence identity of the HPV-11 L2 gene with reference strains

IS.HPV11.24_81																													
JQ773409/CU17/2012/T	100.0%																												
LN833161/A2	100.0%	100.0%																											
MK313765/JO-RRP8/A	100.0%	100.0%	100.0%																										
MK313767/CAC1/A2	100.0%	100.0%	100.0%	100.0%																									
HE574702/JO-RRP_2/	100.0%	100.0%	100.0%	100.0%	100.0%																								
HE611263/LP220/A2	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%																							
JN644141/GUMC-AJ/2(	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%																						
LN833185/A2	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%																					
MK313763/JO-RRP2/A	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%																				
MN788368/JO-RRP10/	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%																			
LN833184/A2	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%																		
JQ773408/CU16/A2	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%																	
MK463914/HPV11-gw-11	99.9%	99.9%	99.9%	99.9%	99.9%	99.9%	99.9%	99.9%	99.9%	99.9%	99.9%	99.9%	99.7%																
FR872717/2011/Hungar	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	99.7%	99.9%															
JQ773411/CU19/A2	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	99.7%	99.9%	100.0%														
JQ773412/CU20/A2	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	99.7%	99.9%	100.0%	100.0%													
EU918768/LZod45/200	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.5%	99.7%	99.7%	99.7%	99.7%												
LN833187/A4	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.3%	99.4%	99.5%	99.5%	99.5%	99.4%											
LN833165/A4	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.4%	99.7%	99.6%	99.6%	99.6%	99.4%	99.5%										
LN833169/A3	99.4%	99.4%	99.4%	99.4%	99.4%	99.4%	99.4%	99.4%	99.4%	99.4%	99.4%	99.4%	99.2%	99.5%	99.4%	99.4%	99.4%	99.2%	99.4%	99.8%									
FN870021/A86/A3	99.9%	99.9%	99.9%	99.9%	99.9%	99.9%	99.9%	99.9%	99.9%	99.9%	99.9%	99.9%	99.7%	99.8%	99.9%	99.9%	99.9%	99.8%	99.5%	99.5%	99.4%								
M14119/1986/Germany/	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.6%	99.7%	99.8%	99.8%	99.8%	99.7%	99.4%	99.4%	99.3%	99.9%							
KU298879/2016/Brazil/	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	99.7%	99.9%	100.0%	100.0%	100.0%	99.7%	99.5%	99.6%	99.4%	99.9%	99.8%						
MK463921/HPV11-gw-11	99.9%	99.9%	99.9%	99.9%	99.9%	99.9%	99.9%	99.9%	99.9%	99.9%	99.9%	99.9%	99.7%	99.8%	99.9%	99.9%	99.9%	99.7%	99.4%	99.5%	99.4%	99.8%	99.7%	99.9%					
MK463916/HPV11-gw-11	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.6%	99.7%	99.8%	99.8%	99.8%	99.6%	99.4%	99.4%	99.3%	99.7%	99.7%	99.8%	99.7%				
FN907962/CAC86	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%	99.6%	99.7%	99.8%	99.8%	99.8%	99.6%	99.4%	99.4%	99.3%	99.7%	99.7%	99.8%	99.7%	99.7%			
LN833183/B	98.7%	98.7%	98.7%	98.7%	98.7%	98.7%	98.7%	98.7%	98.7%	98.7%	98.7%	98.7%	98.5%	98.6%	98.7%	98.7%	98.7%	98.5%	98.7%	98.7%	98.6%	98.6%	98.6%	98.7%	98.6%	98.6%	98.6%		

Following the analysis of the sequencing results, we compared the nucleotide variations of the studied sequence with the prototype sequence and several HPV-11 sequences published in the NCBI GenBank database. The results revealed 19 nucleotide substitution sites, among which 8 specific nucleotide changes led to amino acid alterations, including: 308Ile (C/T), 573Asp (A/C), 643Ala (A/G), 874Leu (A/C), 973Val (A/G), 1000Leu (A/C), 1054Phe (C/T), and 1102Leu (T/C) (Table 12).

At position 308, a C to T substitution compared to the prototype sequence results in a Threonine to Isoleucine amino acid change. Position 573 exhibits an A to C alteration relative to the prototype sequences, leading to an amino acid substitution from Glutamic acid to Aspartic acid. At position 643, an A to G substitution is observed in the study sequence, changing the amino acid from Threonine to Alanine. At position 874, a C to T change compared to the prototype sequence converts Isoleucine to Leucine. At position 973, an A to G alteration in the study sequence modifies the amino acid from Isoleucine to Valine. Position 1000 shows an A to C substitution that replaces Methionine with Leucine. At position 1054, a C to T change corresponds to an amino acid substitution from Leucine to Phenylalanine. Finally, position 1102 displays a T to C alteration compared to the prototype sequences, resulting in a Phenylalanine to Leucine substitution.

Table 12: Distribution of nucleotide and amino acid variations in the L2 gene sequences of the studied HPV-11 strains

Isolate name	Nation/ Year	Type	Nucleotide (amino acid) locations within the L2 gene																		
			36	**168**	**231**	**308**	**324**	**336**	**378**	**399**	**405**	**471**	**573**	**643**	**741**	**819**	**874**	**973**	**1000**	**1054**	**1102**
M14119	Germany 1986	HPV11 /REF	-	-	C *Pro*	-	-	-	-	-	-	C *Pro*	-	-	-	-	-	-	-	-	-
EU918768	China 2009	HPV11	-	-	C *Pro*	**C** ***Thr***	-	-	C *Gly*	-	-	-	-	-	-	-	-	-	-	-	-
LN833183	Slovenia 2016	HPV11	G *Ser*	G *Leu*	-	-	A *Glu*	C *Ile*	-	G *Ser*	A *Ser*	-	**A** ***Glu***	**A** ***Thr***	G *Thr*	G *Glu*	**A** ***Ile***	**A** ***Ile***	**A** ***Met***	**C** ***Leu***	**T** ***Phe***
IS.HPV-11.24 81	**Vietnam 2025**	**HPV-11**	**A** ***Ser***	**A** ***Leu***	**T** ***Pro***	**T** ***Ile***	**G** ***Glu***	**T** ***Ile***	**T** ***Gly***	**T** ***Ser***	**G** ***Ser***	**A** ***Pro***	**C** ***Asp***	**G** ***Ala***	**T** ***Thr***	**A** ***Glu***	**C** ***Leu***	**G** ***Val***	**C** ***Leu***	**T** ***Phe***	**C** ***Leu***

**Ala: Alanine; Asp: Aspartic acid; Leu: Leucine; Ile: Isoleucine; Ser: Serine; Phe: Phenylalanine; Glu: Glutamic acid; Thr: Threonine; Pro: Proline; Gly: Glycine; Val: Valine.*

Regarding the L2 gene region, the phylogenetic tree further confirmed that the study sample IS.HPV-11.24-81-L2 belongs to sublineage A2, with a bootstrap value of 54%. However, unlike the well-defined branching pattern observed in the L1 gene, the strains within sublineage A2 in the L2 phylogenetic tree exhibited a polytomy structure. Within this structure, the Vietnamese study sample clustered together and showed no divergence from the reference strains from Thailand (JQ773409), the USA (JN644141), Hungary (FR872717), or Brazil (KU298879).

This result reflects an exceptionally high degree of genetic conservation of the L2 gene within sublineage A2, where nucleotide variations are relatively limited, allowing isolates from diverse geographical regions to maintain substantial sequence similarity. Additionally, the L2 gene phylogenetic tree recorded independent clustering of strains belonging to sublineages A1, A3, and A4 in the lower branches (with bootstrap values of 66% and 93%, respectively). The taxonomic consistency of sublineage A2 across both the L1 and L2 gene regions provides a robust dataset, confirming the characteristics of the HPV-11 strain circulating among the study subjects in Vietnam.

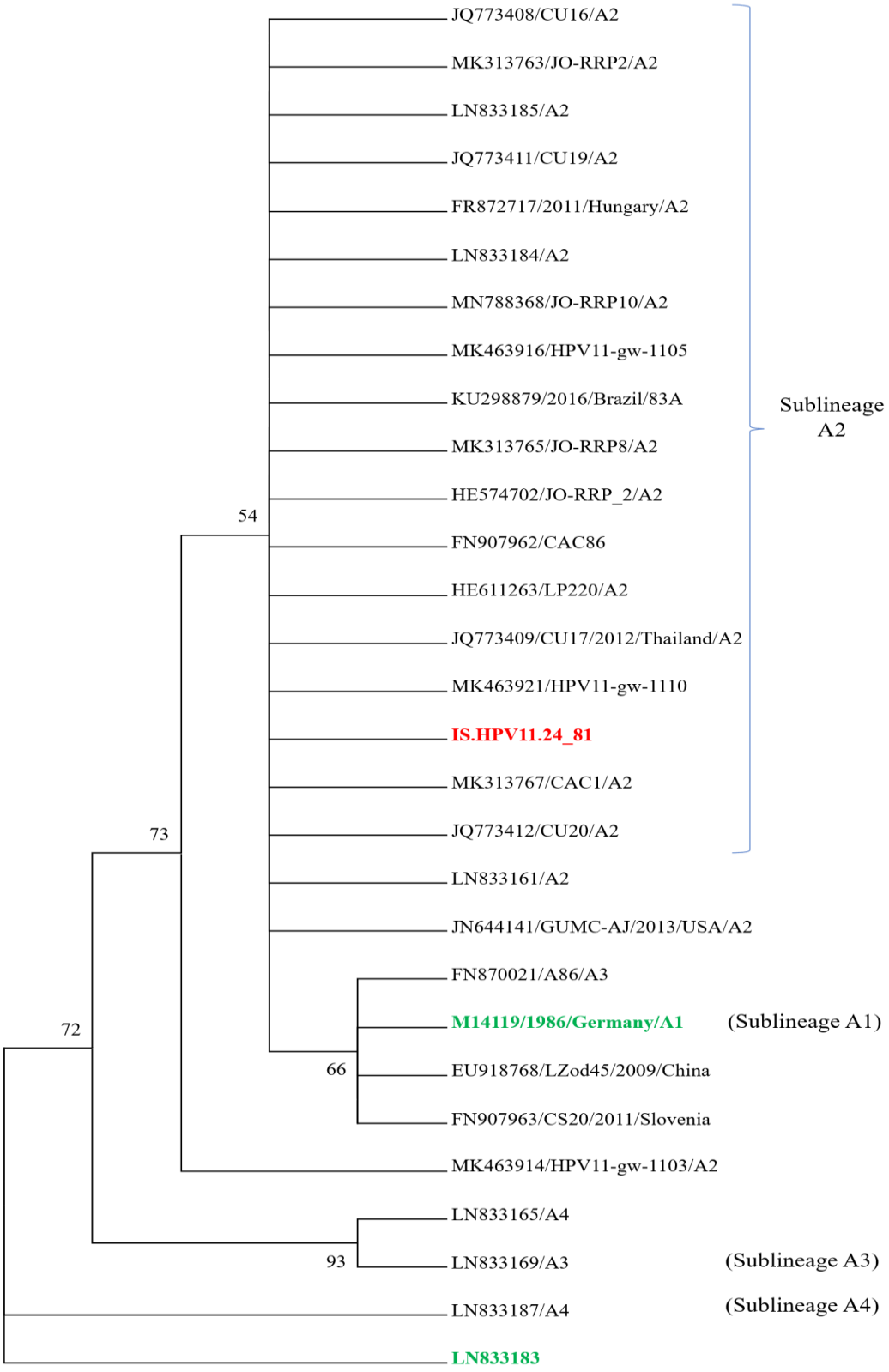


Figure 10: Phylogenetic tree of the L2 gene of HPV-11 strains

4.3.3. E6, E7 gene

The full-length E6 protein-coding gene of the studied HPV-11 strain is 453 bp, encoding an E6 protein of 150 amino acids. These genetic analyses indicated that the E6 gene of the studied HPV-11 strain shares high nucleotide similarity with the reference strains, ranging from 98.6% to 100%. Specifically, the E6 gene of the studied HPV-11 strain exhibits 100% nucleotide similarity with strains JQ773409, LN833161, MK313765, MK313767, HE574702, HE611263, JN644141, LN833185, MK313763, MN788368, LN833184, JQ773408, MK463914, FR872717, KU298879, MK463921, and MK463916, and 99.7% with strains JQ773411 and JQ773412. The study sample shows lower similarity (99.5%) with strains EU918768, M14119, and FN907963. Strain FN907962 displays a similarity rate of 99.3%, whereas two strains, LN833165 and LN833183, share 99.1% similarity. The lowest similarity rate of 98.6% is observed with strain LN833169. Details are presented in Table 13.

The full-length E7 protein-coding gene of the studied HPV-11 strain is 297 bp, encoding an E7 protein of 98 amino acids. Genetic similarity analyses indicated that the E7 gene of the studied HPV-11 strain shares 100% nucleotide identity with strains JQ773409, LN833161, MK313765, MK313767, HE574702, HE611263, JN644141, LN833185, MK313763, MN788368, LN833184, JQ773408, MK463914, FR872717, JQ773411, JQ773412, KU298879, MK463921, MK463916, and FN907962. Therefore, these sequences are completely identical at the nucleotide level. Furthermore, the study sample IS.HPV-11.24-81-E7 displays a lower similarity rate of 98.9% with strains EU918768, LN833187, FN870021, M14119, and FN907963. Finally, the lowest similarity rate of 97.9% is observed with strains LN833165, LN833169, and LN833183. Details are presented in Table 14.

Table 13: Nucleotide sequence identity of the HPV-11 E6 gene with reference strains

IS.HPV11.24_81																												
JQ773409/CU17/2012/Thailand	100%																											
LN833161/A2	100%	100%																										
MK313765/JO-RRP8/A2	100%	100%	100%																									
MK313767/CAC1/A2	100%	100%	100%	100%																								
HE574702/JO-RRP_2/A2	100%	100%	100%	100%	100%																							
HE611263/LP220/A2	100%	100%	100%	100%	100%	100%																						
JN644141/GUMC-AJ/A2	100%	100%	100%	100%	100%	100%	100%																					
LN833185/A2	100%	100%	100%	100%	100%	100%	100%	100%																				
MK313763/JO-RRP2/A2	100%	100%	100%	100%	100%	100%	100%	100%	100%																			
MN788368/JO-RRP10/A2	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%																		
LN833184/A2	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%																	
JQ773408/CU16/A2	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%																
MK463914/HPV11-gw-1103/A2	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%															
FR872717/A2	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%														
JQ773411/CU19/A2	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%													
JQ773412/CU20/A2	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	100%												
EU918768/LZod45	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.3%	99.3%											
LN833187/A4	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.5%	99.5%	99.7%										
LN833165/A4	99.1%	99.1%	99.1%	99.1%	99.1%	99.1%	99.1%	99.1%	99.1%	99.1%	99.1%	99.1%	99.1%	99.1%	99.1%	98.8%	98.8%	99.1%	99.3%									
LN833169/A3	98.6%	98.6%	98.6%	98.6%	98.6%	98.6%	98.6%	98.6%	98.6%	98.6%	98.6%	98.6%	98.6%	98.6%	98.6%	98.4%	98.4%	98.6%	98.8%	99.5%								
FN870021/A86/A3	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.5%	99.5%	99.7%	100%	99.3%	98.8%							
M14119/A1	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.3%	99.3%	99.5%	99.7%	99.1%	98.6%	99.7%						
KU298879/83A	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	99.7%	99.7%	99.5%	99.7%	99.1%	98.6%	99.7%	99.5%					
MK463921/HPV11-gw-1110	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	99.7%	99.7%	99.5%	99.7%	99.1%	98.6%	99.7%	99.5%	100%				
MK463916/HPV11-gw-1105	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	99.7%	99.7%	99.5%	99.7%	99.1%	98.6%	99.7%	99.5%	100%	100%			
FN907963/CS20	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.3%	99.3%	99.5%	99.7%	99.1%	98.6%	99.7%	100%	99.5%	99.5%	99.5%		
FN907962/CAC86	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.1%	99.1%	98.8%	99.1%	98.4%	98.0%	99.1%	98.8%	99.3%	99.3%	99.3%	98.8%	
LN833183	99.1%	99.1%	99.1%	99.1%	99.1%	99.1%	99.1%	99.1%	99.1%	99.1%	99.1%	99.1%	99.1%	99.1%	99.1%	98.8%	98.8%	99.1%	99.3%	98.6%	98.2%	99.3%	99.1%	99.1%	99.1%	99.1%	99.1%	98.8%

Table 14: Nucleotide sequence identity of the HPV-11 E7 gene with reference strains

IS.HPV11.24_81																																	
JQ773409/CU17/2012/	100.0%																																
LN833161/A2	100.0%	100.0%																															
MK313765/JO-RRP8/.	100.0%	100.0%	100.0%																														
MK313767/CAC1/A2	100.0%	100.0%	100.0%	100.0%																													
HE574702/JO-RRP_2	100.0%	100.0%	100.0%	100.0%	100.0%																												
HE611263/LP220/A2	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%																											
JN644141/GUMC-AJ/2	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%																										
LN833185/A2	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%																									
MK313763/JO-RRP2/.	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%																								
MN788368/JO-RRP10	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%																							
LN833184/A2	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%																						
JQ773408/CU16/A2	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%																					
MK463914/HPV11-gw-'	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%																				
FR872717/2011/Hunga	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%																			
JQ773411/CU19/A2	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%																		
JQ773412/CU20/A2	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%																	
EU918768/LZod45/20(	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%																
LN833187/A4	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	100.0%															
LN833165/A4	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.6%	99.6%														
LN833169/A3	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.3%	99.6%	99.6%	100.0%													
FN870021/A86/A3	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	100.0%	100.0%	99.6%	99.6%												
M14119/1986/Germany	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	100.0%	100.0%	99.6%	99.6%	100.0%											
KU298879/2016/Brazil	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	99.6%	99.6%	99.3%	99.3%	99.6%	99.6%										
MK463921/HPV11-gw-'	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	99.6%	99.6%	99.3%	99.3%	99.6%	99.6%	100.0%									
MK463916/HPV11-gw-'	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	99.6%	99.6%	99.3%	99.3%	99.6%	99.6%	100.0%	100.0%								
FN907963/CS20/2011/	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	100.0%	100.0%	99.6%	99.6%	100.0%	100.0%	99.6%	99.6%	99.6%							
FN907962/CAC86	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	99.6%	99.6%	99.3%	99.3%	99.6%	99.6%	100.0%	100.0%	100.0%	99.6%						
LN833183	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	99.3%	99.3%	98.9%	98.9%	99.3%	99.3%	98.9%	98.9%	98.9%	99.3%	98.9%					

Following the analysis of the sequencing results, nucleotide variations in the study sequences were compared with the prototype sequences and selected HPV-11 sequences published in the NCBI GenBank database.

For the E6 gene region of the studied HPV-11 strain, comparison with the reference strains revealed 10 nucleotide substitution positions, including 4 synonymous substitutions and 6 non-synonymous substitutions. Specifically, the 4 nucleotide substitutions that do not alter the amino acid sequence include: 36^{Ser} (T/C), 79^{Leu} (T/C), 279^{Thr} (C/T), and 291^{Ile} (A/T). The 6 nucleotide substitutions compared to the prototype sequence that lead to amino acid changes include: 194^{Ala} (T/C), 282^{Asn} (G/T), 285^{Glu} (C/A), 297^{Lys} (C/A), 298^{Val} (T/G), and 365^{Gly} (A/G). Details are presented in Table 15.

For the E7 gene region of the studied HPV-11 strain, the results revealed 4 nucleotide substitution positions, including 3 non-synonymous substitutions and 1 synonymous substitution. Specifically, the 3 nucleotide substitutions resulting in amino acid alterations include: 34^{Val} (T/G), 133^{Ser} (G/T), and 232^{Gln} (A/C). At position 34, a T to G substitution compared to the prototype sequence leads to an amino acid change from Leucine to Valine. At position 133, a G to T substitution relative to the prototype sequence results in an amino acid alteration from Alanine to Serine. At position 232, an A to C substitution compared to the prototype sequence leads to an amino acid change from Lysine to Glutamine. Additionally, 1 nucleotide substitution that does not lead to an amino acid change is located at position 37^{Leu} (T/C), exhibiting a T to C substitution compared to the prototype sequence. Details are presented in Table 15.

Table 15: Distribution of nucleotide and amino acid substitutions in the E6 and E7 genes of the studied HPV-11 strain compared to reference strains

Isolate name	Nation/ Year	Type	Nucleotide (amino acid) locations within the E6 gene										Nucleotide (amino acid) locations within the E7 gene			
			36	79	194	279	282	285	291	297	298	365	34	37	133	232
M14119	Germany/1986	HPV11/REF	T *Ser*	-	-	C *Thr*	-	-	-	-	-	-	-	-	**G** ***Ala***	-
LN833169	Slovenia/2016	HPV-11	-	-	-	C *Thr*	**G** ***Lys***	**C** ***Asp***	A Ile	**C** ***Asn***	**T** ***Leu***	-	-	-	**G** ***Ala***	**A** ***Lys***
LN833183	Slovenia/2016	HPV-11	-	T *Leu*	**T** ***Val***	C *Thr*	**-**	-	-	-	-	**A** ***Glu***	**T** ***Leu***	T *Leu*	**G** ***Ala***	-
IS.HPV-11.24 81	**Vietnam/ 2025**	**HPV-11**	**C** ***Ser***	**C** ***Leu***	**C** ***Ala***	**T** ***Thr***	**T** ***Asn***	**A** ***Glu***	**T** ***Ile***	**A** ***Lys***	**G** ***Val***	**G** ***Gly***	**G** ***Val***	**C** ***Leu***	**T** ***Ser***	**C** ***Gln***

**Ala: Alanine; Asn: Asparagine; Asp: Aspartic acid; Gln: Glutamine; Glu: Glutamic acid; Gly: Glycine; Ile: Isoleucine; Leu: Leucine; Lys: Lysine; Ser: Serine; Thr: Threonine; Val: Valine.*

The results of this study identified highly variable and conserved regions within the HPV-11 genome. As widely accepted in previous studies, the primary oncoproteins E6 and E7 play a fundamental role in cell transformation and proliferation. Consequently, differences in oncogenic risk among strains and lineages are predominantly determined by the sequence variability of the E6 and E7 genes. Furthermore, E6 and E7 are essential HPV protein products designed to target the tumor suppressor proteins p53 and retinoblastoma (Rb). The degradation of pRb can initiate abnormal cell replication, while the inhibition of p53 can cause these abnormally replicating cells to lose control. Additionally, the HPV genome can only replicate in tandem with the replication of the host genome. Therefore, E6 and E7 are particularly critical in the HPV life cycle [50]. For the E6 gene region of the studied HPV-11 strain, amino acid comparison with the reference strains revealed 10 nucleotide substitution positions, including 6 non-synonymous substitutions and 4 synonymous substitutions. At position 194, a T to C substitution compared to the prototype sequence leads to an amino acid change from Valine to Alanine. At position 282, a G to T substitution relative to the prototype sequence results in an amino acid alteration from Lysine to Asparagine. Position 285 exhibits a C to A change compared to the prototype sequence, leading to an amino acid substitution from Aspartic acid to Glutamic acid. At position 297, a C to A substitution is observed in the study sequence, changing the amino acid from Asparagine to Lysine. At position 298, a T to G substitution results in an amino acid change from Leucine to Valine. Finally, at position 365, an A to G substitution leads to an amino acid alteration from Glutamic acid to Glycine. These positions within the prototype HPV-11 E6 may be located in the central region between the two internal zinc-binding motifs (Zn-motifs), which plays a critical role in the stability of the E6 protein [50].

Overall, the HPV-11 E6 sequence was demonstrated to be highly conserved in the studied sample. Similarly, the E7 gene of the studied strain exhibits 3 nucleotide substitutions that result in amino acid alterations. Specifically, the 3 nucleotide substitution positions leading to amino acid changes include 34^{Val} (T/G), 133^{Ser} (G/T), and 232^{Gln} (A/C). At position 34, a T to G substitution compared to the prototype sequence leads to an amino acid change from Leucine to Valine. At position 133, a G to T substitution relative to the prototype sequence results in an amino acid alteration from Alanine to Serine. At position 232, an A to C substitution compared to the prototype sequence is observed, changing the amino acid from Lysine to

Glutamine. The fact that both sequences are so highly conserved supports the hypothesis that they play an essential role in the structure and function of HPV-11 [51, 52], Consistent with a previous study by Dartmann [53]. This reflects that their roles are highly regulated, and their functions may tolerate little to no structural changes. Consequently, these genes are likely promising targets for HPV primer design and diagnostic detection.

Regarding the E6 and E7 gene regions, the phylogenetic tree analysis further demonstrated taxonomic consistency by placing the Vietnamese study sample, IS.HPV-11.24-81-E6, E7, into sublineage A2, supported by a bootstrap value of 63% (E7 gene). Consistent with the molecular epidemiological trends observed in the L2 gene, the A2 sublineage E6 and E7 gene regions displayed a typical polytomy structure. Within this structure, the study sample exhibited high nucleotide sequence similarity and showed no distinct evolutionary divergence from the vast majority of global reference strains from diverse geographical areas, including Thailand (JQ773409), the USA (JN644141), Hungary (FR872717), and Brazil (KU298879). This result reflects natural selection operating to maintain maximum structural and functional conservation of the E6 and E7 gene regions in the low-risk genotype HPV-11.

On the other hand, the E6 and E7 gene phylogenetic trees also recorded epidemiological mixing and incomplete clustering among the reference strains belonging to sublineages A1, A3, and A4 in the lower branches, typified by the co-branching of LN833165/A4 and LN833169/A3 with bootstrap values up to 95% for the E6 gene and 65% for the E7 gene. This finding reaffirms that the early-region functional genes (E6, E7) exhibit a slow evolutionary rate and high conservation, making them less valuable for deep resolution of sublineages compared to the structural capsid genes (L1, L2). Nevertheless, the consistent classification of the IS.HPV-11.24-81-E6, E7 sample into sublineage A2 provides robust experimental scientific evidence, confirming that the HPV-11 strain circulating among the male patients in this study possesses a homogenous, stable genome and belongs to the globally prevalent sublineage A2. Details are presented in Figures 12 and 13.

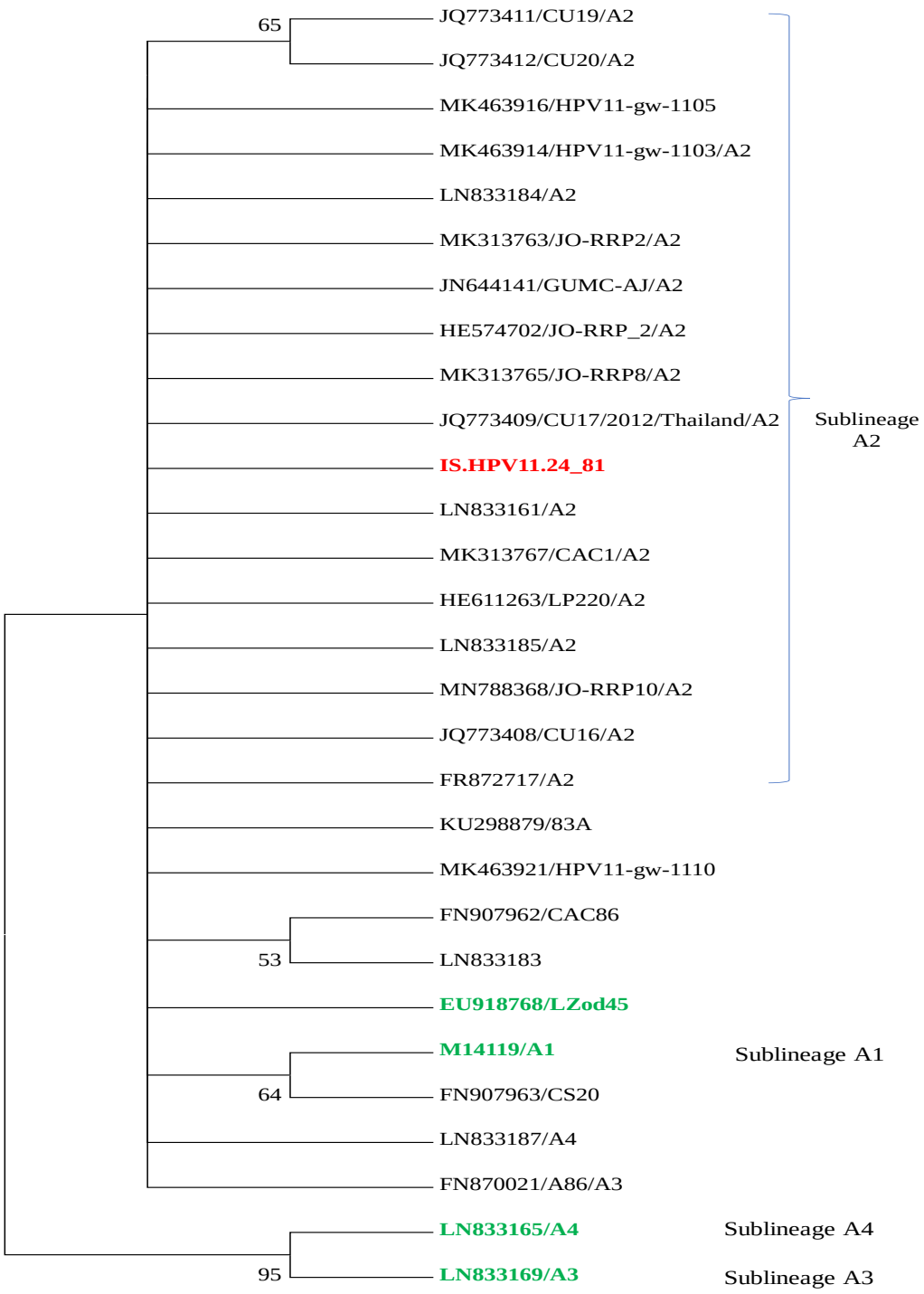


Figure 11: Phylogenetic tree of the E6 gene of HPV-11 strains

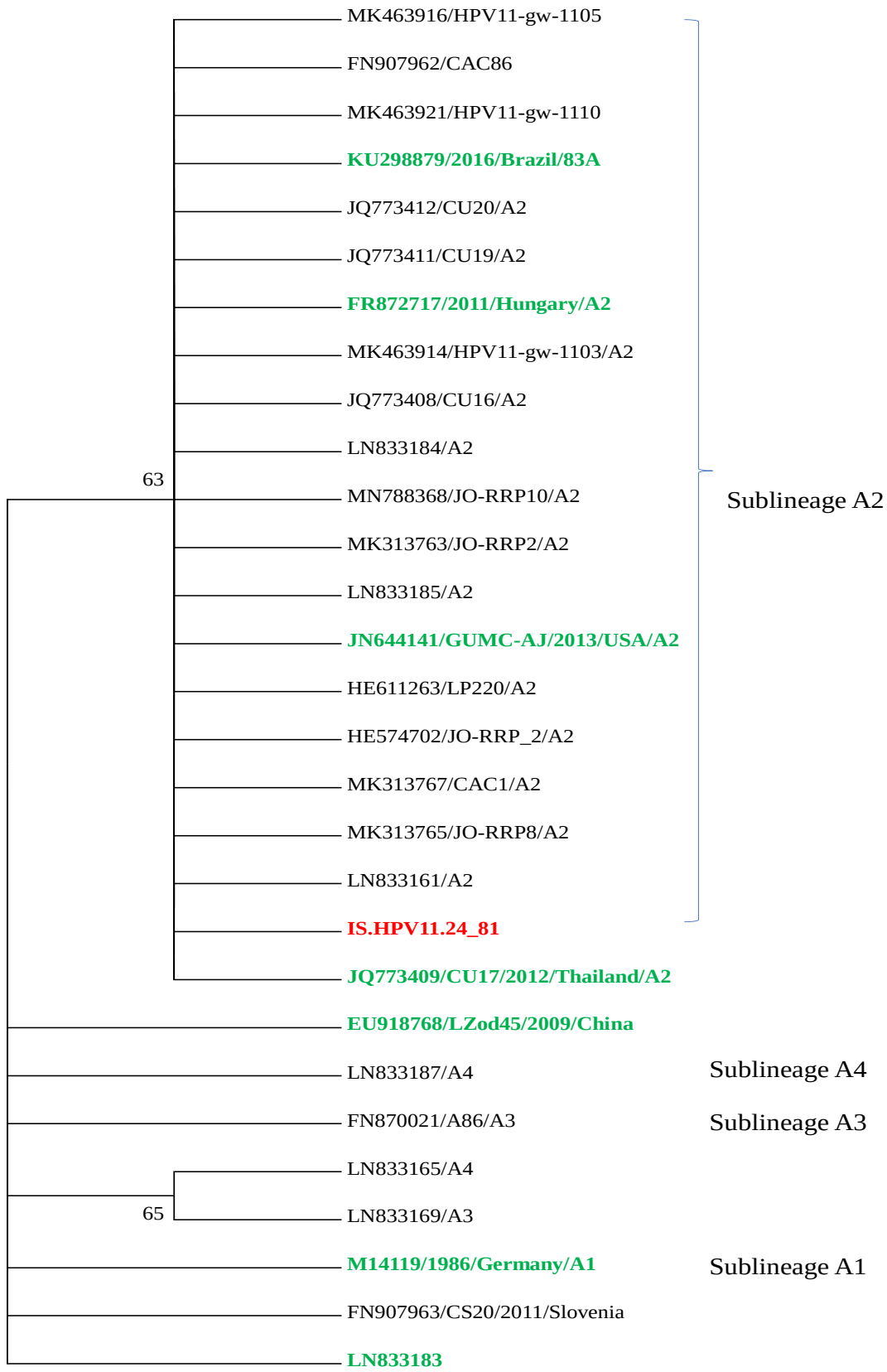


Figure 12: Phylogenetic tree of the E7 gene of HPV-11 strains

4.3.4. E5α, E5β gene

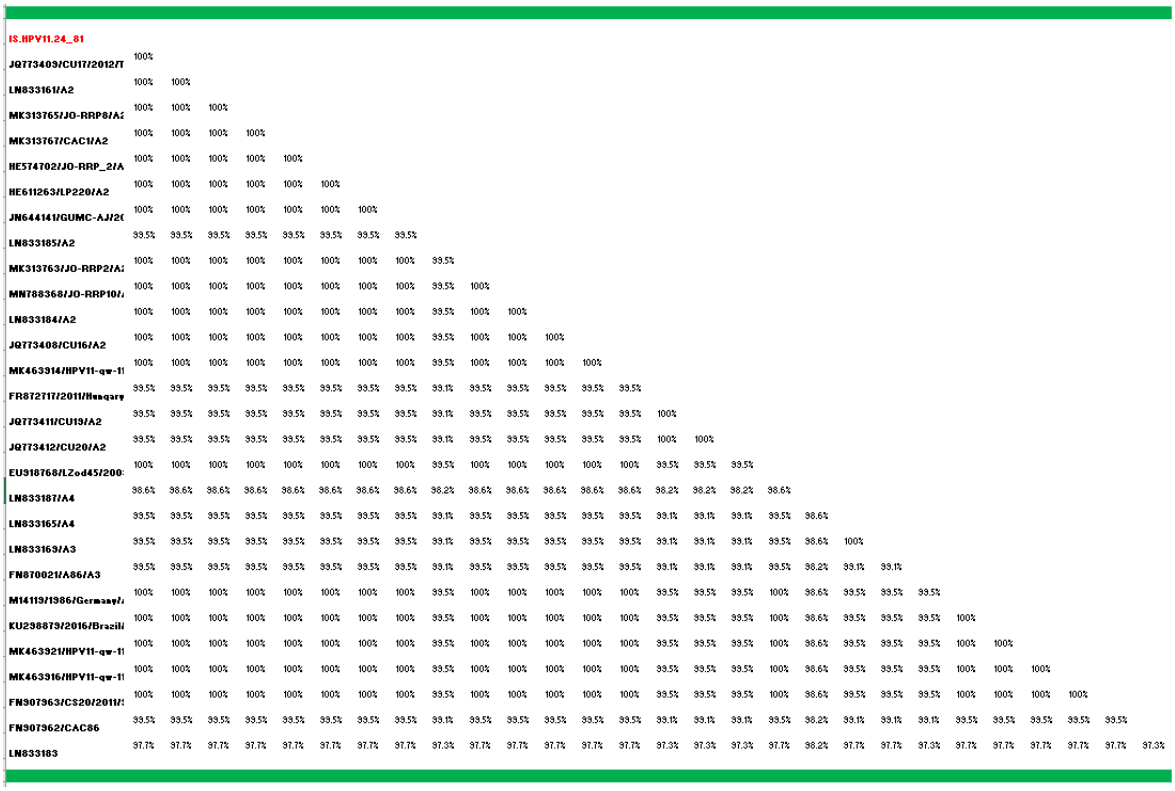
The full-length E5α protein-coding gene of the studied HPV-11 strain is 276 bp, encoding an E5α protein of 91 amino acids. Genetic analyses indicated that the E5α gene of the studied HPV-11 strain shares 100% nucleotide identity with strains JQ773409, LN833161, MK313765, MK313767, HE574702, HE611263, JN644141, LN833185, MK313763, MN788368, LN833184, JQ773408, MK463914, FR872717, JQ773411, JQ773412, KU298879, MK463921, MK463916, and FN907962. Therefore, these sequences are identical at the nucleotide level. Furthermore, the study sample IS.HPV-11.24-81-E5α displays a similarity rate of 99.2% with strains EU918768, FN870021, M14119, and FN907963. The studied strain also shares 98.9% similarity with strains LN833187, LN833165, and LN833169. Finally, the lowest similarity rate of 98.5% is observed with strain LN833183. Details are presented in Table 16.

Table 16: Nucleotide sequence identity of the HPV-11 E5α gene with reference strains

IS.HPV11.24_81																												
JQ773409/CU17/2012/T	100%																											
LN833161/A2	100%	100%																										
MK313765/JO-RRP8/A2	100%	100%	100%																									
MK313767/CAC1/A2	100%	100%	100%	100%																								
HE574702/JO-RRP_2/A	100%	100%	100%	100%	100%																							
HE611263/LP220/A2	100%	100%	100%	100%	100%	100%																						
JN644141/GUMC-AJ/2(	100%	100%	100%	100%	100%	100%	100%																					
LN833185/A2	100%	100%	100%	100%	100%	100%	100%	100%																				
MK313763/JO-RRP2/A2	100%	100%	100%	100%	100%	100%	100%	100%	100%																			
MN788368/JO-RRP10/	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%																		
LN833184/A2	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%																	
JQ773408/CU16/A2	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%																
MK463914/HPV11-qw-11	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%															
FR872717/2011/Hungary	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%														
JQ773411/CU19/A2	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%													
JQ773412/CU20/A2	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%												
EU918768/LZod45/200	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%											
LN833187/A4	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.1%										
LN833165/A4	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.5%									
LN833169/A3	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.9%	98.5%	100%								
FN870021/A86/A3	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	100%	98.1%	98.9%	98.9%							
M14119/1986/Germany/	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	100%	98.1%	98.9%	98.9%	100%						
KU298879/2016/Brazil/	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	99.2%	98.9%	98.9%	98.9%	99.2%	99.2%					
MK463921/HPV11-qw-11	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	99.2%	98.9%	98.9%	98.9%	99.2%	99.2%	100%				
MK463916/HPV11-qw-11	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	99.2%	98.9%	98.9%	98.9%	99.2%	99.2%	100%	100%			
FN907963/CS20/2011/	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	99.2%	100%	98.1%	98.9%	98.9%	100%	100%	99.2%	99.2%	99.2%		
FN907962/CAC86	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	99.2%	98.9%	98.9%	98.9%	99.2%	99.2%	100%	100%	100%	99.2%	
LN833183/B	98.5%	98.5%	98.5%	98.5%	98.5%	98.5%	98.5%	98.5%	98.5%	98.5%	98.5%	98.5%	98.5%	98.5%	98.5%	98.5%	98.5%	97.8%	98.1%	98.1%	98.1%	97.8%	97.8%	98.5%	98.5%	98.5%	97.8%	98.5%

The full-length E5β protein-coding gene of the studied HPV-11 strain is 225 bp, encoding an E5β protein of 74 amino acids. Genetic analyses indicated that the E5β gene of the studied HPV-11 strain shares 100% nucleotide identity with strains JQ773409, LN833161, MK313765, MK313767, HE574702, HE611263, JN644141, MK313763, MN788368, LN833184, JQ773408, MK463914, EU918768, M14119, KU298879, MK463921, MK463916, and FN907963. In other words, these sequences are completely identical at the nucleotide level. Furthermore, the study sample IS.HPV-11.24-81-E5β displays a similarity rate of 99.5% with strains LN833185, FR872717, JQ773411, JQ773412, LN833165, LN833169, FN870021, and FN907962. The studied strain shares a lower similarity rate of 98.6% with strain LN833187, and exhibits the lowest similarity rate with strain LN833183. Details are presented in Table 17.

Table 17: Nucleotide sequence identity of the HPV-11 E5β gene with reference strains



IS.HPV11.24_81																														
JQ773409/CU17/2012/T	100%																													
LN833161/A2	100%	100%																												
MK313765/JO-RRP8/A2	100%	100%	100%																											
MK313767/CAC1/A2	100%	100%	100%	100%																										
HE574702/JO-RRP_2/A	100%	100%	100%	100%	100%																									
HE611263/LP220/A2	100%	100%	100%	100%	100%	100%																								
JN644141/GUMC-AJ/2(	100%	100%	100%	100%	100%	100%	100%																							
LN833185/A2	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%																						
MK313763/JO-RRP2/A2	100%	100%	100%	100%	100%	100%	100%	100%	99.5%																					
MN788368/JO-RRP10/	100%	100%	100%	100%	100%	100%	100%	100%	99.5%	100%																				
LN833184/A2	100%	100%	100%	100%	100%	100%	100%	100%	99.5%	100%	100%																			
JQ773408/CU16/A2	100%	100%	100%	100%	100%	100%	100%	100%	99.5%	100%	100%	100%																		
MK463914/HPV11-qw-11	100%	100%	100%	100%	100%	100%	100%	100%	99.5%	100%	100%	100%	100%																	
FR872717/2011/Hungary	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.1%	99.5%	99.5%	99.5%	99.5%	99.5%																
JQ773411/CU19/A2	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.1%	99.5%	99.5%	99.5%	99.5%	99.5%	100%															
JQ773412/CU20/A2	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.1%	99.5%	99.5%	99.5%	99.5%	99.5%	100%	100%														
EU918768/LZod45/200	100%	100%	100%	100%	100%	100%	100%	100%	99.5%	100%	100%	100%	100%	100%	99.5%	99.5%	99.5%													
LN833187/A4	98.6%	98.6%	98.6%	98.6%	98.6%	98.6%	98.6%	98.6%	98.2%	98.6%	98.6%	98.6%	98.6%	98.6%	98.2%	98.2%	98.2%	98.6%												
LN833165/A4	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.1%	99.5%	99.5%	99.5%	99.5%	99.5%	99.1%	99.1%	99.1%	99.5%	98.6%											
LN833163/A3	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.1%	99.5%	99.5%	99.5%	99.5%	99.5%	99.1%	99.1%	99.1%	99.5%	98.6%	100%										
FN870021/A86/A3	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.1%	99.5%	99.5%	99.5%	99.5%	99.5%	99.1%	99.1%	99.1%	99.5%	98.2%	99.1%	99.1%									
M14119/1986/Germany/	100%	100%	100%	100%	100%	100%	100%	100%	99.5%	100%	100%	100%	100%	100%	99.5%	99.5%	99.5%	100%	98.6%	99.5%	99.5%	99.5%								
KU298879/2016/Brazil	100%	100%	100%	100%	100%	100%	100%	100%	99.5%	100%	100%	100%	100%	100%	99.5%	99.5%	99.5%	100%	98.6%	99.5%	99.5%	99.5%	100%							
MK463921/HPV11-qw-11	100%	100%	100%	100%	100%	100%	100%	100%	99.5%	100%	100%	100%	100%	100%	99.5%	99.5%	99.5%	100%	98.6%	99.5%	99.5%	99.5%	100%	100%						
MK463916/HPV11-qw-11	100%	100%	100%	100%	100%	100%	100%	100%	99.5%	100%	100%	100%	100%	100%	99.5%	99.5%	99.5%	100%	98.6%	99.5%	99.5%	99.5%	100%	100%	100%					
FN907963/CS20/2011/S	100%	100%	100%	100%	100%	100%	100%	100%	99.5%	100%	100%	100%	100%	100%	99.5%	99.5%	99.5%	100%	98.6%	99.5%	99.5%	99.5%	100%	100%	100%	100%				
FN907962/CAC86	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.5%	99.1%	99.5%	99.5%	99.5%	99.5%	99.5%	99.1%	99.1%	99.1%	99.5%	98.2%	99.1%	99.1%	99.1%	99.5%	99.5%	99.5%	99.5%	99.5%			
LN833183	97.7%	97.7%	97.7%	97.7%	97.7%	97.7%	97.7%	97.7%	97.3%	97.7%	97.7%	97.7%	97.7%	97.7%	97.3%	97.3%	97.3%	97.7%	98.2%	97.7%	97.7%	97.3%	97.7%	97.7%	97.7%	97.7%	97.7%	97.3%		

Following the analysis of the sequencing results, nucleotide variations in the study sequence were compared with the prototype sequences and selected HPV-11 sequences published in the NCBI GenBank database.

For the E5α gene region of the HPV-11 strain, the results revealed 6 nucleotide substitution positions, including 4 non-synonymous substitutions and 2 synonymous substitutions. Specifically, the 4 nucleotide substitutions resulting in amino acid alterations include 82^{Phe} (A/T), 115^{Asp} (C/G), 121^{Leu} (G/C), and 163^{Leu} (A/T). At position 82, an A to T substitution compared to the prototype sequence leads to an amino acid change from Isoleucine to Phenylalanine. At position 115, a C to G substitution relative to the prototype sequence results in an amino acid alteration from Histidine to Aspartic acid. At position 121, a G to C substitution compared to the prototype sequence leads to an amino acid change from Valine to Leucine. At position 163, an A to T substitution relative to the prototype sequence results in an amino acid alteration from Methionine to Leucine. Details are presented in Table 18.

Table 18: Distribution of nucleotide and amino acid substitutions in the E5α genes of the studied HPV-11 strain compared to reference strains

Isolate name	Nation/ Year	Type	Nucleotide (amino acid) locations within the E5α gene					
			82	108	115	121	123	163
M14119	Germany/ 1986	HPV-11	**A** **Ile**	-	-	**G** **Val**	-	-
EU918768	China/ 2009	HPV-11	**A** **Ile**	-	-	**G** **Val**	-	-
LN833187	Slovenia/ 2016	HPV-11	-	-	**C** **His**	-	-	-
LN833169	Slovenia/ 2016	HPV-11	-	-	-	**G** **Val**	-	**A** **Met**
FN907963	Slovenia/ 2011	HPV-11	**A** **Ile**	-	-	**G** **Val**	-	-
LN833183	Slovenia/ 2016	HPV-11	-	G Leu	-		T Leu	
IS.HPV-11. 24_81_E5A	**Vietnam/ 2024**	**HPV-11**	**T** **Phe**	**A** **Leu**	**G** **Asp**	**C** **Leu**	**A** **Leu**	**T** **Leu**

For the E5β gene region of HPV-11, the results revealed 8 nucleotide substitution positions, including 7 non-synonymous substitutions and 1 synonymous substitution. Specifically, the 7 nucleotide substitutions resulting in amino acid alterations include 21^{His} (G/C), 82^{Leu} (T/C), 127^{Lys} (C/A), 129^{Lys} (C/A), 167^{Thr} (A/C), 180^{Val} (T/G), and 188^{Ser} (T/C). The single nucleotide substitution that does not lead to an amino acid change is located at position 195^{Gly} (A/T). At position 21, a G to C substitution compared to the prototype sequence leads to an amino acid change from Glutamine to Histidine. At position 82, a T to C substitution relative to the prototype sequence results in an amino acid alteration from Phenylalanine to Leucine. At position 127, a C to A substitution compared to the prototype sequence leads to an amino acid change from Glutamine to Lysine. At position 129, a C to A substitution relative to the prototype sequence results in an amino acid alteration from Asparagine to Lysine. At

position 167, an A to C substitution compared to the prototype sequence leads to an amino acid change from Asparagine to Threonine. At position 180, a T to G substitution relative to the prototype sequence results in an amino acid alteration from Leucine to Valine. At position 188, a T to C substitution compared to the prototype sequence leads to an amino acid change from Leucine to Serine. Details are presented in Table 19.

Table 19: Distribution of nucleotide and amino acid substitutions in the E5β genes of the studied HPV-11 strain compared to reference strains

Isolate Name	Nation/ Year	Type	Nucleotide (amino acid) locations within the E5β gene							
			21	82	127	129	167	180	188	195
LN833165	Slovenia/ 2016	HPV-11	T *His*	-	-	-	-	-	-	-
LN833169	Slovenia/ 2016	HPV-11	T *His*	-	-	-	-	-	-	-
LN833185	Slovenia/ 2016	HPV-11	-	-	-	-	-	T *Val*	-	-
LN833187	Slovenia/ 2016	HPV-11	**G** ***Gln***	-	-	-	-	-	**T** ***Leu***	A *Gly*
FN870021	Slovenia/ 2012	HPV-11	-	-	**C** ***Gln***	-	-	-	-	-
LN188183	Slovenia/ 2016	HPV-11	**G** ***Gln***	**T** ***Phe***	-	**C** ***Asn***	**A** ***Asn***	**T** ***Leu***	-	-
IS.HPV-11. 24_81_E5B	**Vietnam/ 2024**	**HPV-11**	**C** ***His***	**C** ***Leu***	**A** ***Lys***	**A** ***Lys***	**C** ***Thr***	**G** ***Val***	**C** ***Ser***	**T** ***Gly***

**Ala: Alanine; Arg: Arginine; Asp: Aspartic acid; Leu: Leucine; Ile: Isoleucine; Ser: Serine; Phe: Phenylalanine; Asn: Asparagine; Gln: Glutamine; Glu: Glutamic acid; Tyr: Tyrosine; Thr: Threonine; Pro: Proline; Gly: Glycine; Cys: Cysteine; Val: Valine; Ser: Serine; His: Histidine; Lys: Lysine*

The HPV-11 E5α and E5β study samples were found to cluster within sublineage A2, grouping together with reference strains such as JQ773409 (Thailand), KU298879 (Brazil), and FR872717 (Hungary). Based on these genetic relationships, the HPV-11 E5α and E5β study samples exhibit an exceptionally high level of similarity, with virtually no significant evolutionary divergence compared to other global strains within sublineage A2. Details are presented in Figures 14 and 15.

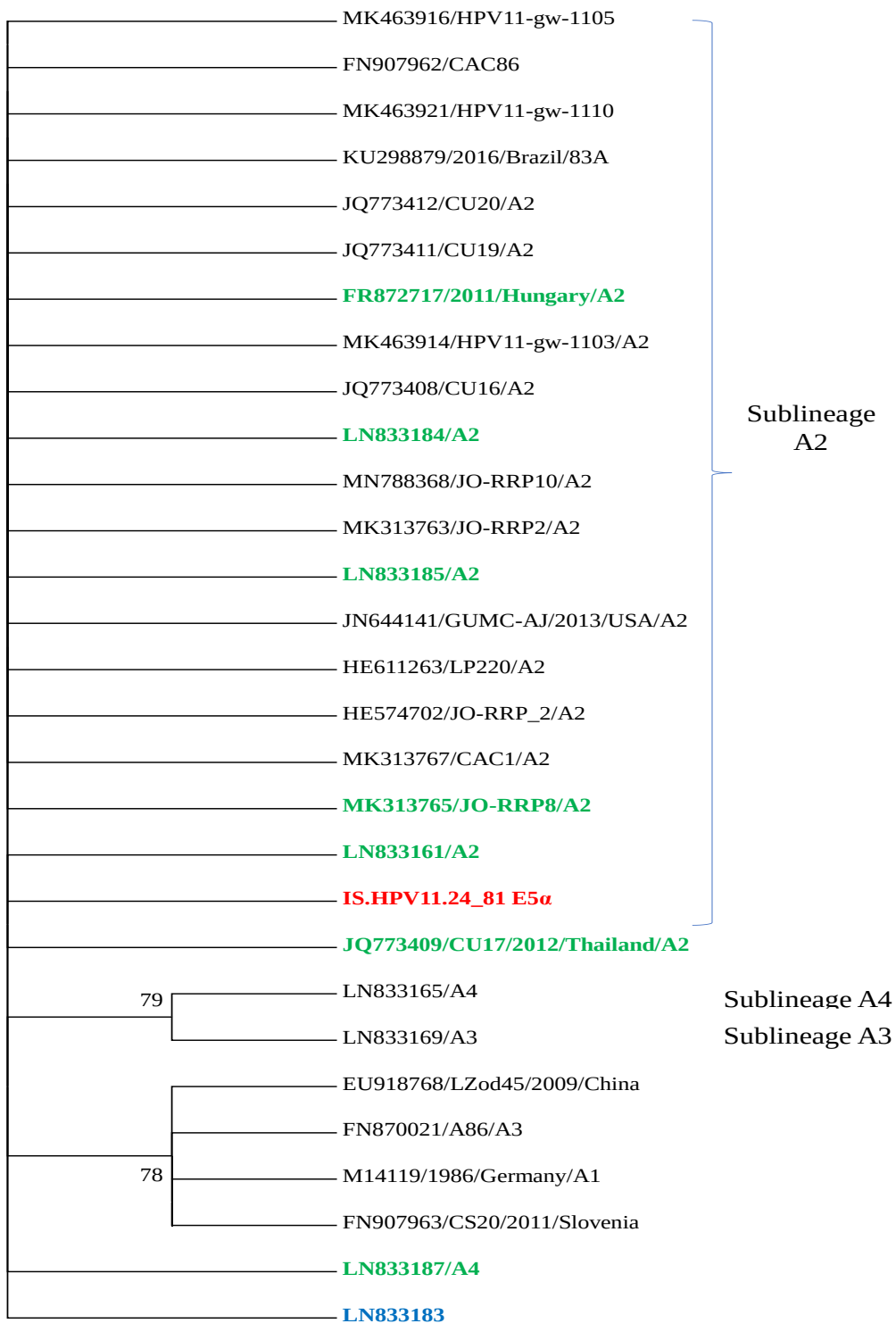


Figure 13: Phylogenetic tree of the E5α gene of HPV-11 strains

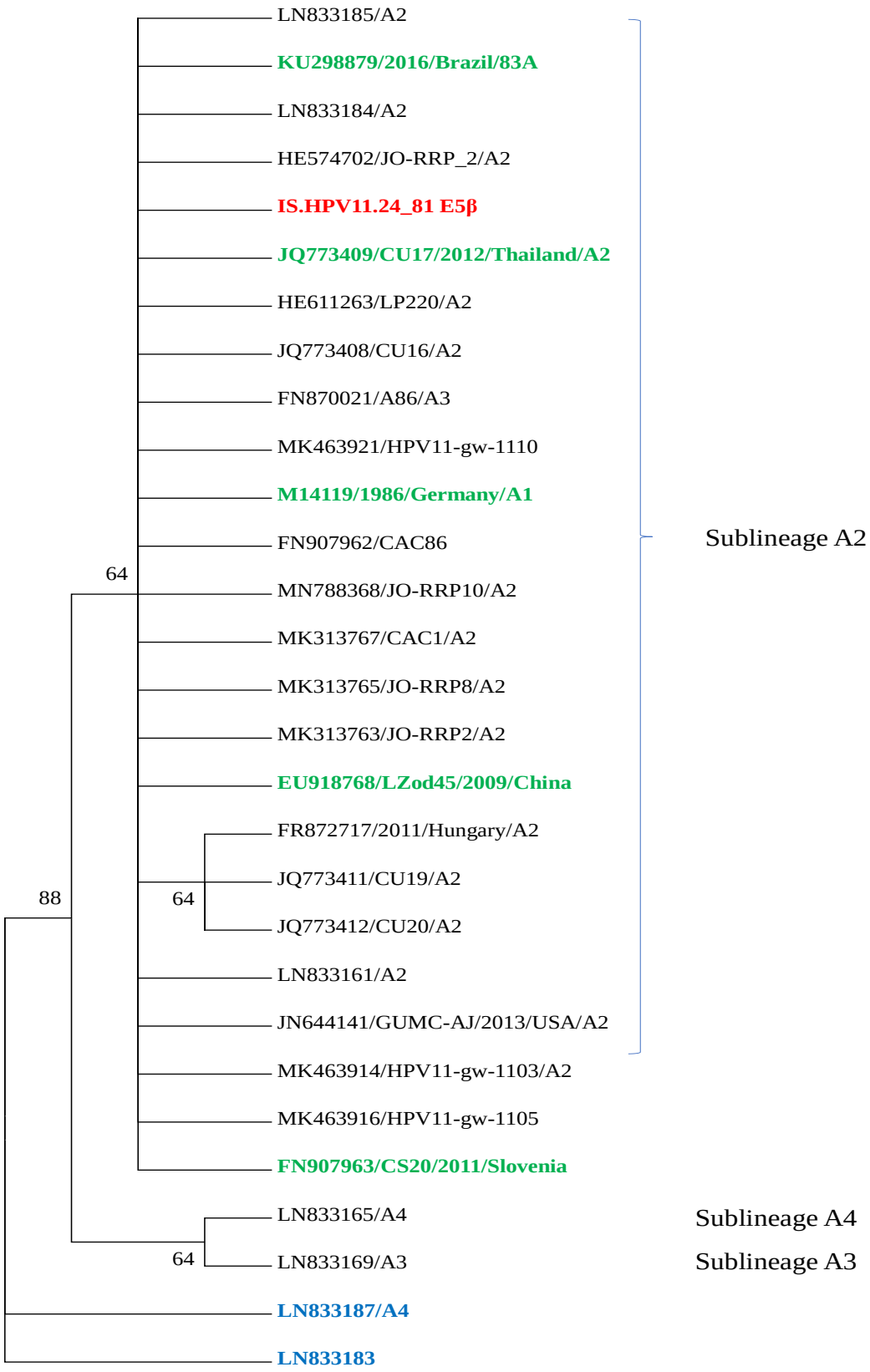


Figure 14: Phylogenetic tree of the E5β gene of HPV-11 strains

4.3.5. E1, E2, E4 Gene

The full-length E1 protein-coding gene of the studied HPV-11 strain is 1,950 bp, encoding an E1 protein of 649 amino acids. Genetic analyses indicated that the E1 gene of the studied HPV-11 strain exhibits high nucleotide similarity ranging from 99.3% to 100% compared to the reference strains. Specifically, the E1 gene of the studied HPV-11 strain shares 100% nucleotide identity with strains JQ773408, MK463914, FR872717/2011, JQ773411, JQ773412, and MK463916. The study sample displays a lower nucleotide similarity rate of 99.9% with strains LN833187, KU298879, MK463921, and FN907963, and a rate of 99.8% with strains JQ773409, LN833161, MK313765, MK313767, HE611263, JN644141, LN833185, MN788368, LN833184, EU918768, LN833165, FN870021, M14119, and FN907962. The study sample shares 99.7% similarity with strains HE574702, MK313763, and LN833169, while the lowest similarity rate of 99.3% is observed with strain LN833183. Details are presented in Table 20.

The full-length E2 protein-coding gene of the studied HPV-11 strain is 1,104 bp, encoding an E2 protein of 367 amino acids. Genetic analyses indicated that the E2 gene of the studied HPV-11 strain exhibits high nucleotide similarity ranging from 98.5% to 99.9% compared to the reference strains. Specifically, the E2 gene of the studied HPV-11 strain shares a 99.9% nucleotide similarity rate with strains MK463921, MK463916, and FN907962. The study sample exhibits the lowest similarity rate with strain LN833183. Details are presented in Table 20.

The full-length E4 protein-coding gene of the studied HPV-11 strain is 327 bp, encoding an E4 protein of 108 amino acids. Genetic analyses indicated that the E4 gene of the studied HPV-11 strain exhibits high nucleotide similarity ranging from 98.1% to 100% compared to the reference strains. Specifically, the E4 gene of the studied HPV-11 strain shares 100% nucleotide identity with strains JQ773408, MK463914, FR872717, JQ773411, JQ773412, LN833187, MK463921, MK463916, and FN907962. The study sample exhibits the lowest similarity rate of 98.1% with strain LN833183. Details are presented in Table 20.

Table 20: Distribution of nucleotide and amino acid substitutions in the E1, E2, E4 genes of the studied HPV-11 strain compared to reference strains

Isolate name	Nucleotide similarity levels of the investigated strain IS.HPV-11.24.81 compared to reference strains		
	E1 Gene	**E2 Gene**	**E4 Gene**
JQ773409	99.8%	99.7%	99.6%
LN833161	99.8%	99.7%	99.6%
MK313765	99.8%	99.7%	99.6%
MK313767	99.8%	99.7%	99.6%
HE574702	99.7%	99.7%	99.6%
HE611263	99.8%	99.7%	99.6%
JN644141	99.8%	99.7%	99.6%
LN833185	99.8%	99.7%	99.6%
MK313763	99.7%	99.7%	99.6%
MN788368	99.8%	99.7%	99.6%
LN833184	99.8%	99.7%	99.6%
JQ773408	100.0%	99.8%	100.0%
MK463914	100.0%	99.8%	100.0%
FR872717	100.0%	99.8%	100.0%
JQ773411	100.0%	99.8%	100.0%
JQ773412	100.0%	99.8%	100.0%
EU918768	99.8%	99.2%	99.3%
LN833187	99.9%	98.9%	100.0%
LN833165	99.8%	99.2%	99.3%
LN833169	99.7%	99.3%	99.6%
FN870021	99.8%	99.2%	99.3%

Isolate name	Nucleotide similarity levels of the investigated strain IS.HPV-11.24.81 compared to reference strains		
	E1 Gene	**E2 Gene**	**E4 Gene**
M14119	99.8%	99.2%	99.3%
KU298879	99.9%	99.7%	99.6%
MK463921	99.9%	99.9%	100.0%
MK463916	100.0%	99.9%	100.0%
FN907963	99.9%	99.2%	99.3%
FN907962	99.8%	99.9%	100.0%
LN833183	99.3%	98.5%	98.1%

The HPV-11 E1, E2, and E4 study samples exhibited a relatively high similarity to variants from Hungary (FR872717, HE574702), Thailand (JQ773408, JQ773411), and Brazil (KU298879), all belonging to sublineage A2. Consequently, based on these genetic relationships, it is evident that the HPV-11 E1, E2, and E4 study samples are closely related to global HPV-11 strains from Asia, Europe, and the Americas. Details are presented in Figures 16, 17, and 18.

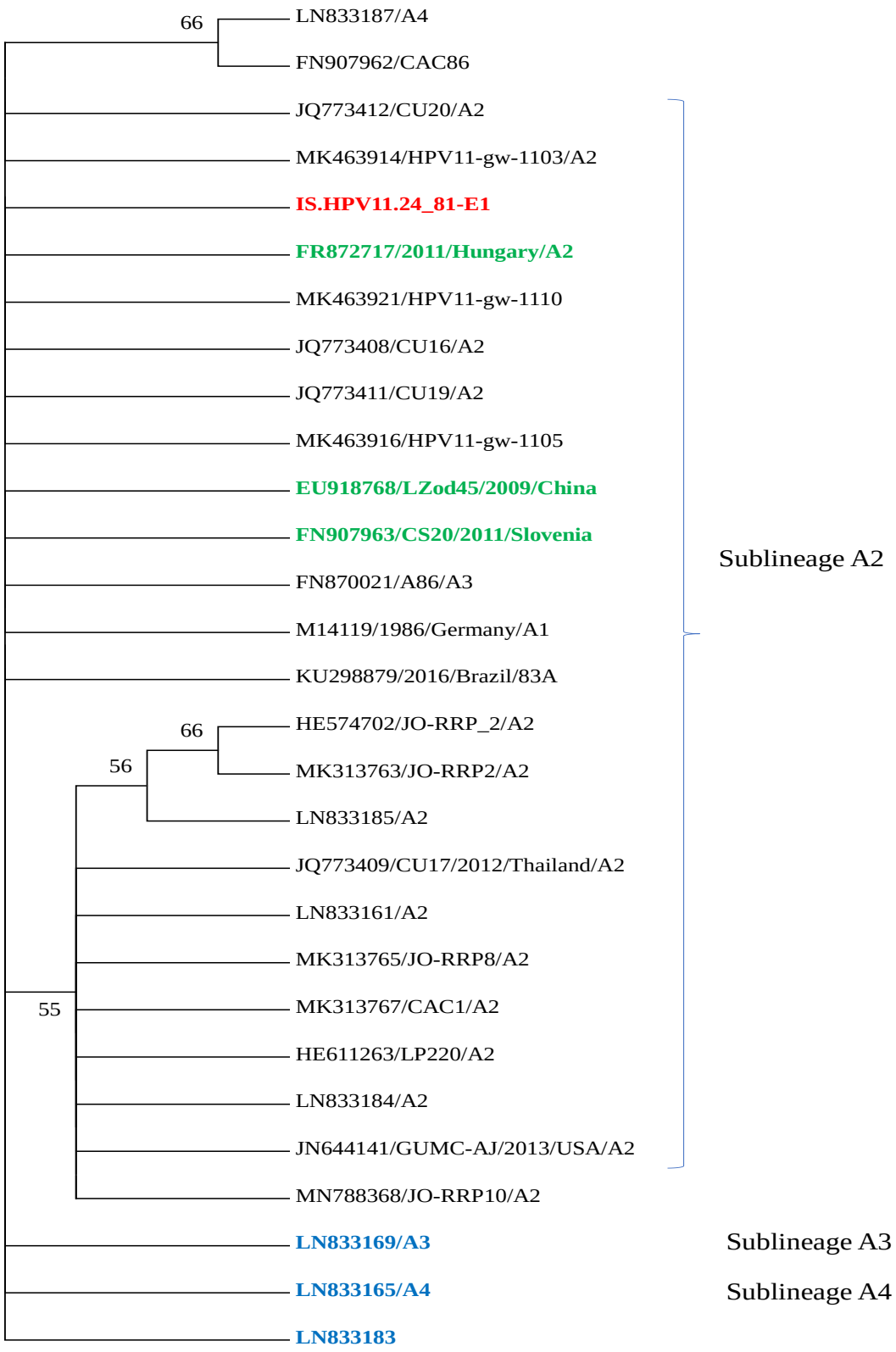


Figure 15: Phylogenetic tree of the E1 gene of HPV-11 strains

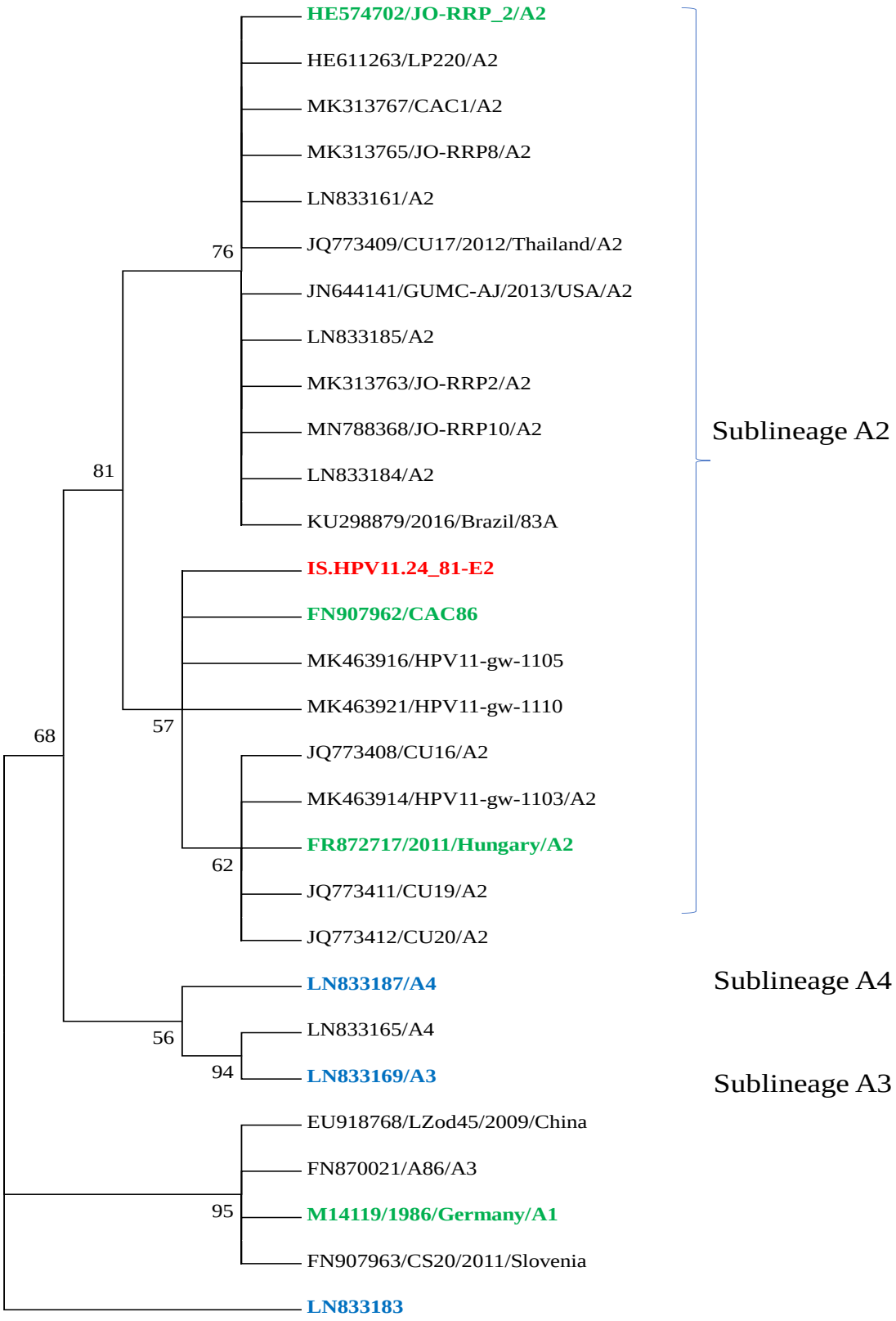


Figure 16: Phylogenetic tree of the E2 gene of HPV-11 strains

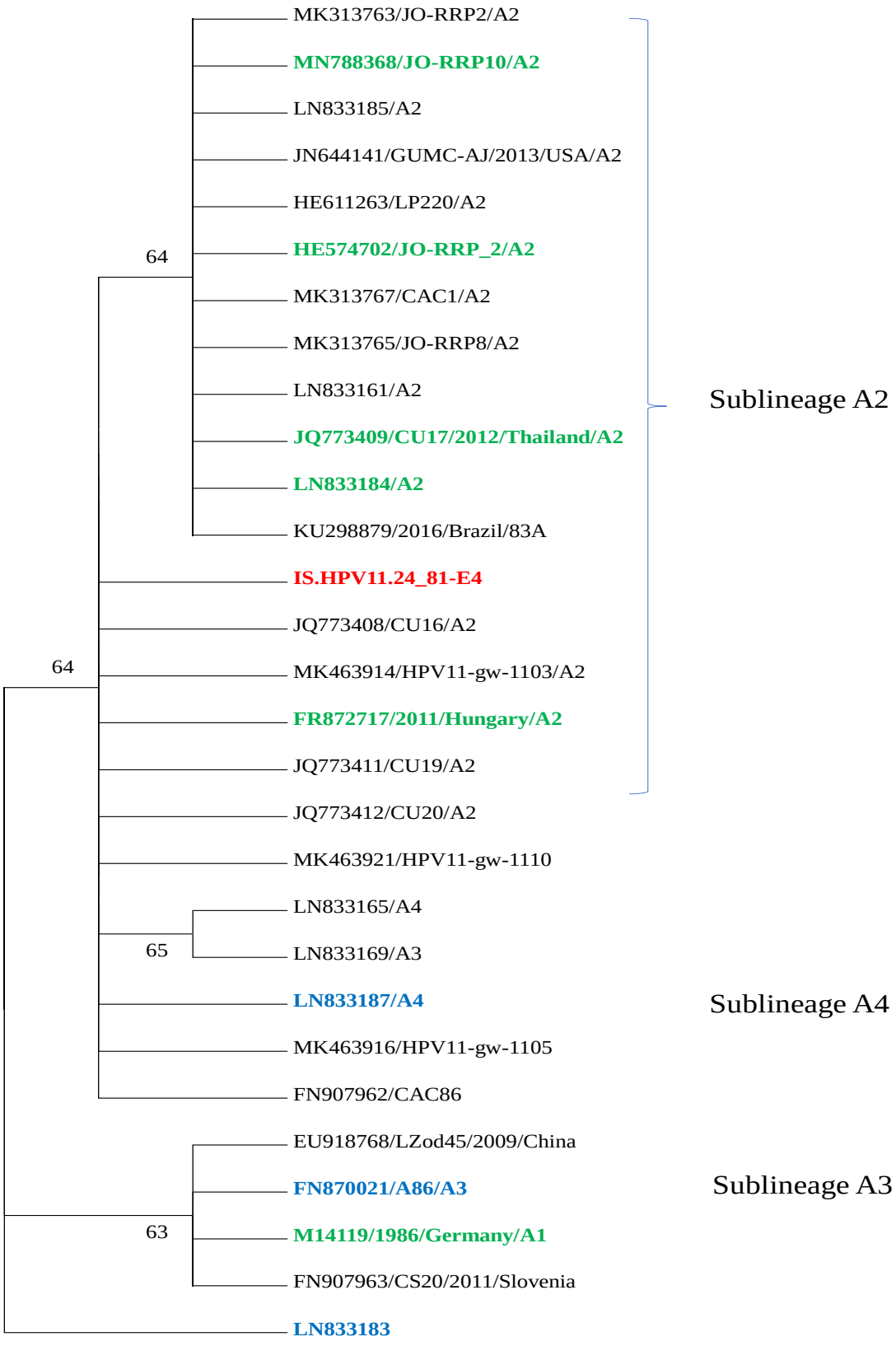


Figure 17: Phylogenetic tree of the E4 gene of HPV-11 strains

The results of the construction and analysis of phylogenetic trees across all gene regions of the Vietnamese study sample, IS.HPV-11.24-81, demonstrated an absolute 100% consensus. All phylogenetic topologies consistently identified the study sample as belonging to Sublineage A2. This uniformity across the entire genome length, from the early to late gene regions, provides robust empirical evidence confirming the pristine nature and high genetic stability of this viral strain. Furthermore, it completely rules out the possibility of cross-recombination events between sublineages within the investigated HPV-11 sample.

Across all phylogenetic trees, the Vietnamese study sample consistently clustered tightly and exhibited near-absolute nucleotide sequence identity (ranging from 99.5% to 100%) with international reference strains representing sublineage A2. These include strains predominantly circulating in Asia (Thailand JQ773409), Europe (Hungary FR872717, Slovenia FN907963), the Americas (the USA JN644141, Brazil KU298879), as well as strains directly linked to global recurrent respiratory papillomatosis (JO-RRP). Evolutionarily, the phylogenetic topologies reflect distinct natural selection pressures operating on individual gene regions. While structural and regulatory genes (L1, L2, E2) possess higher polymorphism, yielding well-defined branching patterns with robust bootstrap values supporting major nodes (ranging from 79% to 94%), the functional and early-region genes (E1, E4, E5A, E5B, E6, E7) maintain a characteristic flat, star-like polytomy structure with branch lengths approaching zero. This outcome demonstrates stringent purifying selection aimed at maintaining absolute nucleotide conservation to preserve the core biological functions of the low-risk HPV-11 virus.

Despite achieving comprehensive sequencing and phylogenetic analyses across multiple core target gene regions, this study is subject to certain limitations. Due to objective constraints, sequencing was performed on only a single representative isolate (IS.HPV-11.24-81) obtained from a male patient attending the Department of Andrology at Hanoi Medical University Hospital; thus, it does not fully capture the comprehensive epidemiological picture of HPV-11 across different geographical regions or on a national scale. Another limitation stems from the biological characteristics of HPV-11, which exhibits a relatively low nucleotide sequence variation (typically less than 1–2% between strains); consequently, no novel sublineage was identified within our study sample.

CHAPTER 5. CONCLUSIONS AND RECOMMENDATIONS

5.1. Conclusions

Based on the results of the genome sequence analysis of Human Papillomavirus type 11 (HPV-11) infecting Vietnamese men, we draw the following conclusions:

- For the first time in Vietnam, the full-length genome (7,932 bp) of an HPV-11 strain (IS.HPV-11.24-81) isolated from male patients was sequenced and comprehensively analyzed.
- Characteristic nucleotide mutations and amino acid changes were identified across both the early gene regions (E1, E2, E4, E5α, E5β, E6, E7) and late gene regions (L1, L2).
- A system of 7 specific primer pairs was successfully designed and optimized to amplify the complete HPV-11 genome.

5.2. Recommendations

- Scale up the research with a larger sample cohort to evaluate the efficiency and specificity of the designed primers in comparison with the reference primers.
- Expand the scope of research to include other HPV genotypes to obtain a more objective and comprehensive assessment.
- Although the evolutionary variations observed in the studied HPV-11 strain relative to the reference strains are minor, they may potentially influence the biological characteristics and pathogenicity of the virus. Therefore, further in-depth studies are required to evaluate their potential impact on the infectivity, replication, and disease progression of HPV-11 in Vietnamese men.

REFERENCES

1. Crosbie, E.J., et al., *Human papillomavirus and cervical cancer.* Lancet, 2013. **382**(9895): p. 889-99.
2. Wei, F., et al., *Causal attribution of human papillomavirus genotypes to invasive cervical cancer worldwide: a systematic analysis of the global literature.* Lancet, 2024. **404**(10451): p. 435-444.
3. Wolf, J., et al., *Human papillomavirus infection: Epidemiology, biology, host interactions, cancer development, prevention, and therapeutics.* Rev Med Virol, 2024. **34**(3): p. e2537.
4. Muñoz, N., et al., *Epidemiologic classification of human papillomavirus types associated with cervical cancer.* N Engl J Med, 2003. **348**(6): p. 518-27.
5. Giuliano, A.R., G. Anic, and A.G. Nyitray, *Epidemiology and pathology of HPV disease in males.* Gynecol Oncol, 2010. **117**(2 Suppl): p. S15-9.
6. Silva, L.L.D., et al., *Malignancy Associated with Low-Risk HPV6 and HPV11: A Systematic Review and Implications for Cancer Prevention.* Cancers (Basel), 2023. **15**(16).
7. Clifford, G., et al., *Chapter 3: HPV type-distribution in women with and without cervical neoplastic diseases.* Vaccine, 2006. **24 Suppl 3**: p. S3/26-34.
8. Olesen, T.B., et al., *Human papillomavirus prevalence among men in sub-Saharan Africa: a systematic review and meta-analysis.* Sex Transm Infect, 2014. **90**(6): p. 455-62.
9. Zheng, L., et al., *Human papillomavirus prevalence and genotype distribution in Liaocheng men between 2016 and 2022.* J Med Virol, 2024. **96**(1): p. e29360.
10. Tekkesin, N., et al., *Male Human Papillomavirus Infection and Genotyping in Turkey.* Asian Pac J Cancer Prev, 2023. **24**(12): p. 4187-4193.
11. Ngan, L.T.K., et al., *Prevalence and distribution of human papillomavirus in genital warts in Can Tho City, Vietnam: a cross-sectional study.* Infect Agent Cancer, 2025. **20**(1): p. 56.
12. Hoai, B.N., et al., *Human papillomavirus prevalence and genotype distribution in Vietnamese male patients between 2016 and 2020.* J Med Virol, 2022. **94**(6): p. 2892-2896.
13. Burd, E.M., *Human papillomavirus and cervical cancer.* Clin Microbiol Rev, 2003. **16**(1): p. 1-17.
14. Yu, L., V. Majerciak, and Z.M. Zheng, *HPV16 and HPV18 Genome Structure, Expression, and Post-Transcriptional Regulation.* Int J Mol Sci, 2022. **23**(9).
15. Doorbar, J., et al., *Human papillomavirus molecular biology and disease association.* Rev Med Virol, 2015. **25 Suppl 1**(Suppl Suppl 1): p. 2-23.
16. Wilson, V.G., et al., *Papillomavirus E1 proteins: form, function, and features.* Virus Genes, 2002. **24**(3): p. 275-90.
17. Frattini, M.G. and L.A. Laimins, *Binding of the human papillomavirus E1 origin-recognition protein is regulated through complex formation with the*

E2 enhancer-binding protein. Proc Natl Acad Sci U S A, 1994. **91**(26): p. 12398-402.
18. Jenkins, O., et al., *Characterization of the helicase and ATPase activity of human papillomavirus type 6b E1 protein.* J Gen Virol, 1996. **77 (Pt 8)**: p. 1805-9.
19. Chojnacki, M. and T. Melendy, *The human papillomavirus DNA helicase E1 binds, stimulates, and confers processivity to cellular DNA polymerase epsilon.* Nucleic Acids Res, 2018. **46**(1): p. 229-241.
20. Castro-Muñoz, L.J., et al., *The Human Papillomavirus (HPV) E1 protein regulates the expression of cellular genes involved in immune response.* Sci Rep, 2019. **9**(1): p. 13620.
21. Rana, A., et al., *Mechanisms of Viral DNA Replication of Human Papillomavirus: E2 Protein-Dependent Recruitment of E1 DNA Helicase to the Origin of DNA Replication.* Int J Mol Sci, 2025. **26**(9).
22. Doorbar, J., *The E4 protein; structure, function and patterns of expression.* Virology, 2013. **445**(1-2): p. 80-98.
23. Roberts, S., et al., *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): p. 6432-45.
24. Davy, C.E., et al., *Human papillomavirus type 16 E1 E4-induced G2 arrest is associated with cytoplasmic retention of active Cdk1/cyclin B1 complexes.* J Virol, 2005. **79**(7): p. 3998-4011.
25. Raj, K., et al., *E1^E4 Protein of Human Papillomavirus Type 16 Associates with Mitochondria.* Journal of Virology, 2004. **78**(13): p. 7199-7207.
26. Chen, J., *Signaling pathways in HPV-associated cancers and therapeutic implications.* Rev Med Virol, 2015. **25 Suppl 1**: p. 24-53.
27. Howie, H.L., R.A. Katzenellenbogen, and D.A. Galloway, *Papillomavirus E6 proteins.* Virology, 2009. **384**(2): p. 324-34.
28. Kombe Kombe, A.J., et al., *Epidemiology and Burden of Human Papillomavirus and Related Diseases, Molecular Pathogenesis, and Vaccine Evaluation.* Front Public Health, 2020. **8**: p. 552028.
29. Johnson, K.M., et al., *Role of heparan sulfate in attachment to and infection of the murine female genital tract by human papillomavirus.* J Virol, 2009. **83**(5): p. 2067-74.
30. Richards, R.M., et al., *Cleavage of the papillomavirus minor capsid protein, L2, at a furin consensus site is necessary for infection.* Proc Natl Acad Sci U S A, 2006. **103**(5): p. 1522-7.
31. Day, P.M., D.R. Lowy, and J.T. Schiller, *Heparan sulfate-independent cell binding and infection with furin-precleaved papillomavirus capsids.* J Virol, 2008. **82**(24): p. 12565-8.
32. de Villiers, E.M., et al., *Classification of papillomaviruses.* Virology, 2004. **324**(1): p. 17-27.
33. Papillomaviruses, H.J.L., France: IARC, *IARC monographs on the evaluation of carcinogenic risks to humans.* 2011.

34. Chen, Z., et al., *Evolution and taxonomic classification of human papillomavirus 16 (HPV16)-related variant genomes: HPV31, HPV33, HPV35, HPV52, HPV58 and HPV67.* PLoS One, 2011. **6**(5): p. e20183.
35. Bzhalava, D., C. Eklund, and J. Dillner, *International standardization and classification of human papillomavirus types.* Virology, 2015. **476**: p. 341-344.
36. Basiletti, J.A., et al., *Human papillomavirus genotyping using next generation sequencing (NGS) in cervical lesions: Genotypes by histologic grade and their relative proportion in multiple infections.* PLoS One, 2022. **17**(11): p. e0278117.
37. Manga, M.M., A. Fowotade, and M. Yahaya, *Human papillomavirus (HPV) infection in males: a need for more awareness*, in *Current perspectives in human papillomavirus*. 2019, IntechOpen.
38. Manga, M., A. Fowotade, and Y. Mohammed, *Human Papillomavirus (HPV) Infection in Males: A Need for More Awareness*, in *Current Perspectives in Human Papillomavirus*, S.K. Saxena, Editor. 2019, IntechOpen: London.
39. Zhao, K.N., W.J. Liu, and I.H. Frazer, *Codon usage bias and A+T content variation in human papillomavirus genomes.* Virus Res, 2003. **98**(2): p. 95-104.
40. Hebner, C.M. and L.A. Laimins, *Human papillomaviruses: basic mechanisms of pathogenesis and oncogenicity.* Rev Med Virol, 2006. **16**(2): p. 83-97.
41. Rampias, T., C. Sasaki, and A. Psyrri, *Molecular mechanisms of HPV induced carcinogenesis in head and neck.* Oral Oncol, 2014. **50**(5): p. 356-63.
42. Molijn, A., et al., *Molecular diagnosis of human papillomavirus (HPV) infections.* J Clin Virol, 2005. **32 Suppl 1**: p. S43-51.
43. Moody, C.A. and L.A. Laimins, *Human papillomavirus oncoproteins: pathways to transformation.* Nat Rev Cancer, 2010. **10**(8): p. 550-60.
44. Castellsagué, X., F.X. Bosch, and N. Muñoz, *The male role in cervical cancer.* Salud Publica Mex, 2003. **45 Suppl 3**: p. S345-53.
45. Lu, B., et al., *Factors associated with acquisition and clearance of human papillomavirus infection in a cohort of US men: a prospective study.* J Infect Dis, 2009. **199**(3): p. 362-71.
46. Arias-Pulido, H., et al., *Human papillomavirus type 18 variant lineages in United States populations characterized by sequence analysis of LCR-E6, E2, and L1 regions.* Virology, 2005. **338**(1): p. 22-34.
47. Burk, R.D., A. Harari, and Z. Chen, *Human papillomavirus genome variants.* Virology, 2013. **445**(1-2): p. 232-43.
48. Burk, R.D., et al., *Classification and nomenclature system for human Alphapapillomavirus variants: general features, nucleotide landmarks and assignment of HPV6 and HPV11 isolates to variant lineages.* Acta Dermatovenerol Alp Pannonica Adriat, 2011. **20**(3): p. 113-23.
49. Jelen, M.M., et al., *Global Genomic Diversity of Human Papillomavirus 11 Based on 433 Isolates and 78 Complete Genome Sequences.* J Virol, 2016. **90**(11): p. 5503-5513.

50. Barbosa, M.S., D.R. Lowy, and J.T. Schiller, *Papillomavirus polypeptides E6 and E7 are zinc-binding proteins.* J Virol, 1989. **63**(3): p. 1404-7.
51. Donne, A.J., et al., *The role of HPV type in Recurrent Respiratory Papillomatosis.* Int J Pediatr Otorhinolaryngol, 2010. **74**(1): p. 7-14.
52. Hummel, M., J.B. Hudson, and L.A. Laimins, *Differentiation-induced and constitutive transcription of human papillomavirus type 31b in cell lines containing viral episomes.* J Virol, 1992. **66**(10): p. 6070-80.
53. Dartmann, K., et al., *The nucleotide sequence and genome organization of human papilloma virus type 11.* Virology, 1986. **151**(1): p. 124-30.

APPENDIX

>IS.HPV11.24_81 7932bp

ATGGAAAGTAAAGATGCCTCCACGTCTGCAACATCCATAGACCAGTTGTGCAAGACGTTTAATC
TTTCTTTGCACACTCTGCAAATTCAGTGCGTGTTTTGCAGGAATGCACTGACCACCGCAGAGAT
ATATGCATATGCCTATAAGAACCTAAAGGTTGTGTGGCGAGACAACTTTCCCTTTGCAGCGTGT
GCCTGTTGCTTAGAACTGCAAGGGAAAATTAACCAATATAGACACTTTAATTATGCTGCATATG
CACCTACAGTAGAAGAAGAAACTAATGAAGATATTTTAAAAGTGTTAATTCGTTGTTACCTGTG
TCACAAGCCGTTGTGTGAAATAGAAAAACTAAAGCACATATTGGGAAAGGCACGCTTCATAAA
ACTAAATAACCAGTGGAAGGGTCGTTGCTTACACTGCTGGACAACATGCATGGAAGACTTGTTA
CCCTAAAGGATATAGTACTAGACCTGCAGCCTCCTGACCCTGTAGGGTTACATTGCTATGAGCA
ATTAGAAGACAGCTCAGAAGATGAGGTGGACAAGGTGGACAAACAAGACTCACAACCTTTAAC
ACAACATTACCAAATACTGACCTGTTGCTGTGGATGTGACAGCAACGTCCGACTGGTTGTGGAG
TGCACAGACGGAGACATCAGACAACTACAAGACCTTTTGCTGGGCACACTAAATATTGTGTGTC
CCATCTGCGCACCAAAACCATAACAAGGATGGCGGACGATTCAGGTACAGAAAATGAGGGGTC
GGGGTGTACAGGATGGTTTATGGTAGAAGCCATAGTAGAGCACACTACAGGTACACAAATATC
AGAAGATGAGGAAGAGGAGGTGGAGGACAGTGGGTATGACATGGTGGACTTTATTGATGACAG
GCATATTACACAAAATTCTGTGGAAGCACAGGCATTGTTTAATAGGCAGGAGGCGGATGCTCAT
TATGCGACTGTGCAGGACCTAAAACGAAAGTATTTAGGCAGTCCATATGTAAGTCCTATAAGCA
ATGTAGCTAATGCAGTAGAAAGTGAGATAAGTCCACGGTTAGACGCCATTAAACTTACAACAC
AGCCAAAAAAGGTAAAGCGACGGCTGTTTGAAACACGGGAATTAACGGACAGTGGATATGGCT
ATTCTGAAGTGGAAGCTGCAACGCAGGTAGAGAAACATGGCGACCCGGAAAATGGGGGAGAT
GGTCAGGAAAGGGACACAGGGAGGGACATAGAGGGTGAGGGGGTGGAACATAGAGAGGCGGA
AGCAGTAGACGACAGCACCCGAGAGCATGCAGACACATCAGGAATATTAGAATTACTAAAATG
TAAGGATATACGATCTACATTACATGGTAAGTTTAAAGACTGCTTTGGGCTGTCATTTGTTGATT
TAATTAGGCCATTTAAAAGTGATAGAACCACATGTGCCGATTGGGTGGTTGCAGGATTTGGTAT
ACATCATAGCATAGCAGATGCATTTCAAAAGTTAATTGAGCCATTAAGTTTATATGCACATATA
CAATGGCTTACAAATGCATGGGGAATGGTACTATTAGTATTAATAAGGTTTAAAGTAAATAAGA
GCAGATGTACCGTGGCACGTACATTAGGTACGTTATTAAATATACCTGAAAATCACATGTTAAT
TGAGCCTCCTAAAATACAAAGTGGCGTAGCAGCCCTGTATTGGTTTAGGACAGGCATTTCAAAT
GCAAGTACAGTTATAGGGGAGGCGCCGGAATGGATAACGCGCCAGACCGTTATTGAACATAGT
TTGGCTGACAGTCAATTTAAATTAACTGAAATGGTGCAGTGGGCATATGATAATGATATTTGTG
AAGAAAGTGAGATAGCATTTGAATATGCACAGCGTGGAGACTTTGACTCCAATGCAAGGGCCT
TTTTAAATAGTAATATGCAGGCTAAATATGTAAAAGATTGTGCAATTATGTGCAGACATTATAA
ACATGCAGAAATGAAAAAGATGTCTATTAAACAATGGATTAAGTATAGGGGTACTAAAGTTGA
CAGTGTAGGTAACTGGAAGCCAATTGTGCAGTTTCTAAGACATCAAAACATAGAATTTATTCCA
TTTTTAAGCAAACTAAAATTATGGCTGCACGGAACGCCCAAAAAAAATTGTATAGCCATTGTAG
GGCCACCTGACACTGGGAAGTCGTGCTTTTGCATGAGTTTAATTAAGTTTTTGGGGGGAACAGT

TATTAGTTATGTTAATTCCTGCAGCCATTTCTGGCTACAGCCACTAACGGATGCAAAAGTGGCA
TTATTGGATGATGCCACACAACCATGTTGGACATATATGGATACATATATGAGAAACCTATTAG
ATGGTAATCCTATGAGCATAGATAGAAAACATAGAGCATTAACATTAATTAAGTGTCCACCGCT
ACTGGTTACATCAAATATAGACATTAGCAAAGAGGAGAAATACAAATATTTACATAGTAGGGT
TACCACATTTACATTTCCAAATCCATTCCCCTTTGACAGAAATGGGAATGCAGTATATGAACTAT
CAGATGCAAACTGGAAATGTTTCTTTGAAAGACTGTCGTCCAGCCTAGACATTGAGGATTCAGA
GGACGAGGAAGATGGAAGCAATAGCCAAGCGTTTAGATGCGTGCCAGGATCAGTTGTTAGAAC
TTTATGAAGAAAACAGTATTGATATACACAAACACATTATGCATTGGAAATGCATACGATTGGA
AAGTGTATTACTACACAAAGCAAAACAAATGGGCCTGAGCCACATTGGGCTACAAGTAGTACC
ACCATTAACTGTGTCAGAGACTAAAGGACATAATGCTATTGAAATGCAAATGCATTTAGAATCC
TTAGCAAAAACTCAGTATGGTGTGGAACCTTGGACATTACAGGACACCAGTTATGAAATGTGGC
TAACACCACCCAAACGGTGCTTTAAAAAACAGGGAAATACTGTGGAGGTAAAATTTGATGGCT
GTGAAGACAATGTAATGGAGTATGTGGTATGGACACATATATACCTGCAGGACAACGACTCAT
GGGTAAAAGTAACTAGTTCCGTAGATGCCAAGGGCATATATTATACATGTGGACAGTTTAAAAC
ATATTATGTAAATTTTAATAAAGAGGCACAAAAGTATGGTAGTACCAATCATTGGGAAGTATGT
TATGGCAGCACAGTTATATGTTCTCCTGCATCTGTATCTAGCACTGTACGAGAAGTATCCATTGC
TGAACCTACTACATACACCCCCGCACAGACCACCGCCCCTACAGTGTCCGCCTGCACCACGGAA
GACGGCGTGTCGGCGCCGCCTAGAAAGCGAGCACGTGGACCGTCCACTAACAACACCCTGTGT
GTGGCCAACATTAGATCCGTGGACAGTACAATCAACAACATCGTCACTGACAATTACAACAAG
CACCAAAGAAGGAACAACTGTCACAGTGCAGCTACGCCTATAGTGCAACTGCAAGGTGATTCC
AATTGTTTAAAATGTTTTAGATATCGACTGAATGACAAATATAGACATTTGTTTGAATTAGCATC
TTCAACGTGGCATTGGGCCTCACCTGAGGCACCACATAAAAATGCAATTGTAACATTAACCTAT
AGCAGTGAGGAACAACGTCAGCAATTTTTAAACAGTGTAAAAATACCACCCACCATTAGGCAT
AAGGTGGGGTTTATGTCATTACATTTATTGTAACCATTGCACCTGTATATATGTATATGTGTACA
TAACATACGTGTATGGAGGTAGTGCCTGTACAAATTGCTGCAGCAACAACTACAACATTGATAT
TGCCTGTTGTTATTGCATTTGCAGTATGTTTTCTTAGTATTGTACTTATAATATTAATATCTGATT
TTCTAGTATATACATCTGTGCTGGTACTAACACTTCTTTTATATTTGCTTTTGTGGCTTTTATTAA
CAACCCCTTTGCAATTCTTTTTACTAACACTGTGTGTGTGCTATTTTCCTGCCTTTTATATACACA
TATACATTGTGCAAACGCAACAATAATGGTGATGTTAACCTGTCACTTAAATGATGGTGATACA
TGGTTGTTTCTGTGGTTGTTTACTGCATTTGTTGTAGCTGTACTTGGATTGTTGTTACTACATTAC
AGGGCTGTACATGGTACTGAAAAAACTAAATGTGCTAAGTGTAAATCAAACCGCAATACTACT
GTGGATTATGTGTATATGTCACATGGTGATAATGGAGATTATGTGTACATGAACTAGAGTAAAC
CTTTTTTATACAGTGTGTGGTGTACGTTAGTTATATATAATGAAACCTAGGGCACGCAGACGTA
AACGTGCGTCAGCCACACAACTATATCAAACATGCAAGGCCACTGGTACATGTCCCCCAGATGT
AATTCCTAAAGTTGAACATACTACTATTGCAGATCAAATATTAAAATGGGGAAGCTTAGGGGTT
TTTTTTGGTGGGTTAGGTATTGGTACAGGGGCTGGTAGTGGCGGTCGTGCAGGGTATATACCCT
TGGGAAGCTCTCCTAAGCCTGCTATTACTGGGGGGCCAGCAGCACGTCCGCCAGTGCTTGTGGA
GCCTGTTGCCCCTTCCGATCCCTCCATTGTGTCCTTAATTGAGGAGTCTGCTATTATTAATGCTG
GTGCACCTGAGGTGGTACCCCCTACACAGGGTGGCTTTACTATAACATCATCTGAATCGACTAC

ACCTGCTATTTTAGATGTGTCTGTTACCAATCACACTACCACTAGTGTGTTTCAAAATCCACTGT
TTACAGAACCGTCTGTAATACAGCCCCAACCACCTGTGGAGGCCAGTGGTCACATACTTATATC
TGCCCCAACAATAACATCCCAACATGTAGAAGACATTCCACTAGACACTTTTGTTGTATCCTCTA
GTGATAGTGGACCTACATCCAGTACTCCTCTTCCTCGTGCTTTTCCTCGGCCTCGGGTGGGTTTG
TATAGTCGTGCCTTACAGCAGGTACAGGTTACGGACCCCGCGTTTTTGTCCACGCCACAGCGAT
TGGTAACTTATGACAACCCTGTCTATGAAGGAGAAGATGTAAGTTTACAATTTACCCATGAGTC
TATCCACAATGCACCTGATGAAGCATTTATGGATATTATTAGACTACATAGACCAGCTATAACG
TCCAGACGGGGTCTTGTGCGTTTTAGTCGCATTGGGCAACGGGGGTCCATGTACACACGCAGTG
GACAACATATAGGTGCCCGCATACATTATTTTCAGGACATTTCACCAGTTACACAAGCTGCAGA
GGAAATAGAACTGCACCCTCTAGTGGCTGCAGAAAATGACACGTTTGATATTTATGCTGAACCA
TTTGACCCTATCCCTGACCCTGTCCAACATTCTGTTACACAGTCTTATCTTACCTCCACACCTAAT
ACCCTTTCACAATCGTGGGGTAATACCACAGTCCCATTGTCAATCCCTAGTGACTGGTTTGTGCA
GTCTGGGCCTGACATAACTTTTCCTACTGCATCTATGGGAACACCCTTTAGTCCTGTAACTCCTG
CTTTACCTACAGGCCCTGTTTTTATTACAGGTTCTGACTTCTATTTGCATCCTACATGGTACTTTG
CACGCAGACGCCGTAAACGTATTCCCTTATTTTTTACAGATGTGGCGGCCTAGCGACAGCACAG
TATATGTGCCTCCTCCCAACCCTGTATCCAAGGTTGTTGCCACGGATGCGTATGTTAAACGCACC
AACATATTTTATCATGCCAGCAGTTCTAGACTCCTTGCTGTGGGACATCCATATTACTCTATCAA
AAAAGTTAACAAAACAGTTGTACCAAAGGTGTCTGGATATCAATATAGAGTGTTTAAGGTAGTG
TTGCCAGATCCTAACAAGTTTGCATTACCTGATTCATCTCTGTTTGACCCCACTACACAGCGTTT
AGTATGGGCGTGCACAGGGTTGGAGGTAGGCAGGGGTCAACCTTTAGGCGTTGGTGTTAGTGG
GCATCCATTGCTAAACAAATATGATGATGTAGAAAATAGTGGTGGGTATGGTGGTAATCCTGGT
CAGGATAATAGGGTTAATGTAGGTATGGATTATAAACAAACCCAGCTATGTATGGTGGGCTGTG
CTCCACCGTTAGGTGAACATTGGGGTAAGGGTACACAATGTTCAAATACCTCTGTACAAAATGG
TGACTGCCCCCCGTTGGAACTTATTACCAGTGTTATACAGGATGGGGACATGGTTGATACAGGC
TTTGGTGCTATGAATTTTGCAGACTTACAAACCAATAAATCGGATGTTCCCCTTGATATTTGTGG
AACTGTCTGCAAATATCCTGATTATTTGCAAATGGCTGCAGACCCTTATGGTGATAGGTTGTTTT
TTTATTTGCGAAAGGAACAAATGTTTGCTAGACACTTTTTTAATAGGGCCGGTACTGTGGGGGA
ACCTGTGCCTGATGACCTGTTGGTAAAAGGGGGTAATAATAGATCATCTGTAGCTAGTAGTATT
TATGTACATACACCTAGTGGCTCATTGGTGTCTTCAGAGGCTCAATTATTTAATAAACCATATTG
GCTTCAAAAGGCTCAGGGACATAACAATGGTATTTGCTGGGGAAACCACTTGTTTGTTACTGTG
GTAGATACCACACGCAGTACAAATATGACACTATGTGCATCTGTGTCTAAATCTGCTACATACA
CTAATTCAGATTATAAGGAATACATGCGCCATGTGGAGGAGTTTGATTTACAGTTTATTTTTCAA
TTGTGTAGCATTACATTATCTGCAGAAGTCATGGCCTATATACACACAATGAATCCTTCTGTTTT
GGAGGACTGGAACTTTGGTTTATCGCCTCCACCAAATGGTACACTGGAGGATACTTATAGATAT
GTACAGTCACAGGCCATTACCTGTCAGAAACCCACACCTGAAAAAGAAAAACAGGATCCCTAT
AAGGATATGAGTTTTTGGGAGGTTAACTTAAAAGAAAAGTTTTCAAGTGAATTAGATCAGTTTC
CCCTTGGACGTAAGTTTTTATTGCAAAGTGGATATCGAGGACGGACGTCTGCTCGTACAGGTAT
AAAGCGCCCAGCTGTGTCTAAGCCCTCTACAGCCCCCAAACGAAAACGTACCAAAACCAAAAA
GTAATATATGTGTGTCAGTGTGTTGTGTTATTTATATGTTGTTGTAGTGTGTATATGTTTCTTGTA

TTGTGTATATGTGTATATGTTTGTGTATATGTGTATGTTATGTATGTTATGTTGTTATGTATGTTT
GTGTGTTTAGTGTGTGTATATATTTGTGGAATGTGTATGTATGTTTTTGTGCAATAAACAATTAT
TATGTGTGTTCTGTTACACCCAGTGACTAAGTTGTGTTTGCACGCGCCGTTTGTGTTGCCTTCAT
ATTATATTATACATATTTGTAATATACCTATACTATGTTACCCCCCCCCCACTTGCAACCGTTTTC
GGTTGCCCTTACATACACTTACCTCAAATTTGTTATAACGTGTTTTGTACTAATCCCATATGTTGT
GTGCCAAGGTACATATTGCCCTGCCAAGTATCTTGCCAACAACACACCTGGCGCCAGGGCGCGG
TATTGCATGACTAATGTACAATAAACCTGTCGGTTTGTACAATGTTGTGGATTGCAGCCAAAGG
TTAAAAGCATTTTTGGCTTCTAGCTGAACATTTTTGTACCCTTAGTATATTATGCACAATACCCA
CAAAATGAGTAACCTAAGGTCACACACCTGCAACCGGTTTCGGTTACCCACACCCTACATATTT
CCTTCTTATACTTAATAACAATCTTAGTTTAAAAAAGAGGAGGGACCGAAAACGGTTCAACCGA
AAACGGTTATATATAAACCAGCCCAAAAAATTAGCAGACGAGGCATT

Component Genes

>L1 1506bp

ATGTGGCGGCCTAGCGACAGCACAGTATATGTGCCTCCTCCCAACCCTGTATCCAAGGTTGTTG
CCACGGATGCGTATGTTAAACGCACCAACATATTTTATCATGCCAGCAGTTCTAGACTCCTTGCT
GTGGGACATCCATATTACTCTATCAAAAAAGTTAACAAAACAGTTGTACCAAAGGTGTCTGGAT
ATCAATATAGAGTGTTTAAGGTAGTGTTGCCAGATCCTAACAAGTTTGCATTACCTGATTCATCT
CTGTTTGACCCCACTACACAGCGTTTAGTATGGGCGTGCACAGGATTGGAGGTAGGCAGGGGTC
AACCTTTAGGCGTTGGTGTTAGTGGGCATCCATTGCTAAACAAATATGATGATGTAGAAAATAG
TGGTGGGTATGGTGGTAATCCTGGTCAGGATAATAGGGTTAATGTAGGTATGGATTATAAACAA
ACCCAGCTATGTATGGTGGGCTGTGCTCCACCGTTAGGTGAACATTGGGGTAAGGGTACACAAT
GTTCAAATACCTCTGTACAAAATGGTGACTGCCCCCCGTTGGAACTTATTACCAGTGTTATACA
GGATGGGGACATGGTTGATACAGGCTTTGGTGCTATGAATTTTGCAGACTTACAAACCAATAAA
TCGGATGTTCCCCTTGATATTTGTGGAACTGTCTGCAAATATCCTGATTATTTGCAAATGGCTGC
AGACCCTTATGGTGATAGGTTGTTTTTTTATTTGCGGAAGGAACAAATGTTTGCTAGACACTTTT
TTAATAGGGCCGGTACTGTGGGGGAACCTGTGCCTGATGACCTGTTGGTAAAAGGGGGTAATA
ATAGATCATCTGTAGCTAGTAGTATTTATGTACATACACCTAGTGGCTCATTGGTGTCTTCAGAG
GCTCAATTATTTAATAAACCATATTGGCTTCAAAAGGCTCAGGGACATAACAATGGTATTTGCT
GGGGAAACCACTTGTTTGTTACTGTGGTAGATACCACACGCAGTACAAATATGACACTATGTGC
ATCTGTGTCTAAATCTGCTACATACACTAATTCAGATTATAAGGAATACATGCGCCATGTGGAG
GAGTTTGATTTACAGTTTATTTTTCAATTGTGTAGCATTACATTATCTGCAGAAGTCATGGCCTA
TATACACACAATGAATCCTTCTGTTTTGGAGGACTGGAACTTTGGTTTATCGCCTCCACCAAATG
GTACACTGGAGGATACTTATAGATATGTACAGTCACAGGCCATTACCTGTCAGAAACCCACACC
TGAAAAAGAAAAACAGGATCCCTATAAGGATATGAGTTTTTGGGAGGTTAACTTAAAAGAAAA
GTTTTCAAGTGAATTAGATCAGTTTCCCCTTGGACGTAAGTTTTTATTGCAAAGTGGATATCGAG
GACGGACGTCTGCTCGTACAGGTATAAAGCGCCCAGCTGTGTCTAAGCCCTCTACAGCCCCCAA
ACGAAAACGTACCAAAACCAAAAAGTAA

>L2 1368bp

ATGAAACCTAGGGCACGCAGACGTAAACGTGCGTCAGCCACACAACTATATCAAACATGCAAG
GCCACTGGTACATGTCCCCCAGATGTAATTCCTAAAGTTGAACATACTACTATTGCAGATCAAA
TATTAAAATGGGGAAGCTTAGGGGTTTTTTTTGGTGGGTTAGGTATTGGTACAGGGGCTGGTAG
TGGCGGTCGTGCAGGGTATATACCCTTGGGAAGCTCTCCTAAGCCTGCTATTACTGGGGGGCCA
GCAGCACGTCCGCCAGTGCTTGTGGAGCCTGTTGCCCCTTCCGATCCCTCCATTGTGTCCTTAAT
TGAGGAGTCTGCTATTATTAATGCTGGTGCACCTGAGGTGGTACCCCCTACACAGGGTGGCTTT
ACTATAACATCATCTGAATCGACTACACCTGCTATTTTAGATGTGTCTGTTACCAATCACACTAC
CACTAGTGTGTTTCAAAATCCACTGTTTACAGAACCGTCTGTAATACAGCCCCAACCACCTGTG
GAGGCCAGTGGTCACATACTTATATCTGCCCCAACAATAACATCCCAACATGTAGAAGACATTC
CACTAGACACTTTTGTTGTATCCTCTAGTGATAGTGGACCTACATCCAGTACTCCTCTTCCTCGT
GCTTTTCCTCGGCCTCGGGTGGGTTTGTATAGTCGTGCCTTACAGCAGGTACAGGTTACGGACCC
CGCGTTTTTGTCCACGCCACAGCGATTGGTAACTTATGACAACCCTGTCTATGAAGGAGAAGAT
GTAAGTTTACAATTTACCCATGAGTCTATCCACAATGCACCTGATGAAGCATTTATGGATATTAT
TAGACTACATAGACCAGCTATAACGTCCAGACGGGGTCTTGTGCGTTTTAGTCGCATTGGGCAA
CGGGGGTCCATGTACACACGCAGTGGACAACATATAGGTGCCCGCATACATTATTTTCAGGACA
TTTCACCAGTTACACAAGCTGCAGAGGAAATAGAACTGCACCCTCTAGTGGCTGCAGAAAATG
ACACGTTTGATATTTATGCTGAACCATTTGACCCTATCCCTGACCCTGTCCAACATTCTGTTACA
CAGTCTTATCTTACCTCCACACCTAATACCCTTTCACAATCGTGGGGTAATACCACAGTCCCATT
GTCAATCCCTAGTGACTGGTTTGTGCAGTCTGGGCCTGACATAACTTTTCCTACTGCATCTATGG
GAACACCCTTTAGTCCTGTAACTCCTGCTTTACCTACAGGCCCTGTTTTTATTACAGGTTCTGAC
TTCTATTTGCATCCTACATGGTACTTTGCACGCAGACGCCGTAAACGTATTCCCTTATTTTTTAC
AGATGTGGCGGCCTAG

>E1 1950bp

ATGGCGGACGATTCAGGTACAGAAAATGAGGGGTCGGGGTGTACAGGATGGTTTATGGTAGAA
GCCATAGTAGAGCACACTACAGGTACACAAATATCAGAAGATGAGGAAGAGGAGGTGGAGGA
CAGTGGGTATGACATGGTGGACTTTATTGATGACAGGCATATTACACAAAATTCTGTGGAAGCA
CAGGCATTGTTTAATAGGCAGGAGGCGGATGCTCATTATGCGACTGTGCAGGACCTAAAACGA
AAGTATTTAGGCAGTCCATATGTAAGTCCTATAAGCAATGTAGCTAATGCAGTAGAAAGTGAGA
TAAGTCCACGGTTAGACGCCATTAAACTTACAACACAGCCAAAAAAGGTAAAGCGACGGCTGT
TTGAAACACGGGAATTAACGGACAGTGGATATGGCTATTCTGAAGTGGAAGCTGCAACGCAGG
TAGAGAAACATGGCGACCCGGAAAATGGGGGAGATGGTCAGGAAAGGGACACAGGGAGGGAC
ATAGAGGGTGAGGGGGTGGAACATAGAGAGGCGGAAGCAGTAGACGACAGCACCCGAGAGCA
TGCAGACACATCAGGAATATTAGAATTACTAAAATGTAAGGATATACGATCTACATTACATGGT
AAGTTTAAAGACTGCTTTGGGCTGTCATTTGTTGATTTAATTAGGCCATTTAAAAGTGATAGAAC
CACATGTGCCGATTGGGTGGTTGCAGGATTTGGTATACATCATAGCATAGCAGATGCATTTCAA
AAGTTAATTGAGCCATTAAGTTTATATGCACATATACAATGGCTTACAAATGCATGGGGAATGG

TACTATTAGTATTAATAAGGTTTAAAGTAAATAAGAGCAGATGTACCGTGGCACGTACATTAGG
TACGTTATTAAATATACCTGAAAATCACATGTTAATTGAGCCTCCTAAAATACAAAGTGGCGTA
GCAGCCCTGTATTGGTTTAGGACAGGCATTTCAAATGCAAGTACAGTTATAGGGGAGGCGCCGG
AATGGATAACGCGCCAGACCGTTATTGAACATAGTTTGGCTGACAGTCAATTTAAATTAACTGA
AATGGTGCAGTGGGCATATGATAATGATATTTGTGAAGAAAGTGAGATAGCATTTGAATATGCA
CAGCGTGGAGACTTTGACTCCAATGCAAGGGCCTTTTTAAATAGTAATATGCAGGCTAAATATG
TAAAAGATTGTGCAATTATGTGCAGACATTATAAACATGCAGAAATGAAAAAGATGTCTATTAA
ACAATGGATTAAGTATAGGGGTACTAAAGTTGACAGTGTAGGTAACTGGAAGCCAATTGTGCA
GTTTCTAAGACATCAAAACATAGAATTTATTCCATTTTTAAGCAAACTAAAATTATGGCTGCAC
GGAACGCCCAAAAAAAATTGTATAGCCATTGTAGGGCCACCTGACACTGGGAAGTCGTGCTTTT
GCATGAGTTTAATTAAGTTTTTGGGGGGAACAGTTATTAGTTATGTTAATTCCTGCAGCCATTTC
TGGCTACAGCCACTAACGGATGCAAAAGTGGCATTATTGGATGATGCCACACAACCATGTTGGA
CATATATGGATACATATATGAGAAACCTATTAGATGGTAATCCTATGAGCATAGATAGAAAACA
TAGAGCATTAACATTAATTAAGTGTCCACCGCTACTGGTTACATCAAATATAGACATTAGCAAA
GAGGAGAAATACAAATATTTACATAGTAGGGTTACCACATTTACATTTCCAAATCCATTCCCCT
TTGACAGAAATGGGAATGCAGTATATGAACTATCAGATGCAAACTGGAAATGTTTCTTTGAAAG
ACTGTCGTCCAGCCTAGACATTGAGGATTCAGAGGACGAGGAAGATGGAAGCAATAGCCAAGC
GTTTAGATGCGTGCCAGGATCAGTTGTTAGAACTTTATGA

>E2 1104bp

ATGGAAGCAATAGCCAAGCGTTTAGATGCGTGCCAGGATCAGTTGTTAGAACTTTATGAAGAA
AACAGTATTGATATACACAAACACATTATGCATTGGAAATGCATACGATTGGAAAGTGTATTAC
TACACAAAGCAAAACAAATGGGCCTGAGCCACATTGGGCTACAAGTAGTACCACCATTAACTG
TGTCAGAGACTAAAGGACATAATGCTATTGAAATGCAAATGCATTTAGAATCCTTAGCAAAAAC
TCAGTATGGTGTGGAACCTTGGACATTACAGGACACCAGTTATGAAATGTGGCTAACACCACCC
AAACGGTGCTTTAAAAAACAGGGAAATACTGTGGAGGTAAAATTTGATGGCTGTGAAGACAAT
GTAATGGAGTATGTGGTATGGACACATATATACCTGCAGGACAACGACTCATGGGTAAAAGTA
ACTAGTTCCGTAGATGCCAAGGGCATATATTATACATGTGGACAGTTTAAAACATATTATGTAA
ATTTTAATAAAGAGGCACAAAAGTATGGTAGTACCAATCATTGGGAAGTATGTTATGGCAGCAC
AGTTATATGTTCTCCTGCATCTGTATCTAGCACTGTACGAGAAGTATCCATTGCTGAACCTACTA
CATACACCCCCGCACAGACCACCGCCCCTACAGTGTCCGCCTGCACCACGGAAGACGGCGTGTC
GGCGCCGCCTAGAAAGCGAGCACGTGGACCGTCCACTAACAACACCCTGTGTGTGGCCAACAT
TAGATCCGTGGACAGTACAATCAACAACATCGTCACTGACAATTACAACAAGCACCAAAGAAG
GAACAACTGTCACAGTGCAGCTACGCCTATAGTGCAACTGCAAGGTGATTCCAATTGTTTAAAA
TGTTTTAGATATCGACTGAATGACAAATATAGACATTTGTTTGAATTAGCATCTTCAACGTGGCA
TTGGGCCTCACCTGAGGCACCACATAAAAATGCAATTGTAACATTAACCTATAGCAGTGAGGAA
CAACGTCAGCAATTTTTAAACAGTGTAAAAATACCACCCACCATTAGGCATAAGGTGGGGTTTA
TGTCATTACATTTATTGTAA

>E4 327bp

ATGGTAGTACCAATCATTGGGAAGTATGTTATGGCAGCACAGTTATATGTTCTCCTGCATCTGTA
TCTAGCACTGTACGAGAAGTATCCATTGCTGAACCTACTACATACACCCCCGCACAGACCACCG
CCCCTACAGTGTCCGCCTGCACCACGGAAGACGGCGTGTCGGCGCCGCCTAGAAAGCGAGCAC
GTGGACCGTCCACTAACAACACCCTGTGTGTGGCCAACATTAGATCCGTGGACAGTACAATCAA
CAACATCGTCACTGACAATTACAACAAGCACCAAAGAAGGAACAACTGTCACAGTGCAGCTAC
GCCTATAG

>E5α 276bp

ATGGAGGTAGTGCCTGTACAAATTGCTGCAGCAACAACTACAACATTGATATTGCCTGTTGTTA
TTGCATTTGCAGTATGTTTTCTTAGTATTGTACTTATAATATTAATATCTGATTTTCTAGTATATA
CATCTGTGCTGGTACTAACACTTCTTTTATATTTGCTTTTGTGGCTTTTATTAACAACCCCTTTGC
AATTCTTTTTACTAACACTGTGTGTGTGCTATTTTCCTGCCTTTTATATACACATATACATTGTGC
AAACGCAACAATAA

>E5β 225bp

ATGGTGATGTTAACCTGTCACTTAAATGATGGTGATACATGGTTGTTTCTGTGGTTGTTTACTGC
ATTTGTTGTAGCTGTACTTGGATTGTTGTTACTACATTACAGGGCTGTACATGGTACTGAAAAAA
CTAAATGTGCTAAGTGTAAATCAAACCGCAATACTACTGTGGATTATGTGTATATGTCACATGG
TGATAATGGAGATTATGTGTACATGAACTAG

>E6 453bp

ATGGAAAGTAAAGATGCCTCCACGTCTGCAACATCCATAGACCAGTTGTGCAAGACGTTTAATC
TTTCTTTGCACACTCTGCAAATTCAGTGCGTGTTTTGCAGGAATGCACTGACCACCGCAGAGAT
ATATGCATATGCCTATAAGAACCTAAAGGTTGTGTGGCGAGACAACTTTCCCTTTGCAGCGTGT
GCCTGTTGCTTAGAACTGCAAGGGAAAATTAACCAATATAGACACTTTAATTATGCTGCATATG
CACCTACAGTAGAAGAAGAAACTAATGAAGATATTTTAAAAGTGTTAATTCGTTGTTACCTGTG
TCACAAGCCGTTGTGTGAAATAGAAAAACTAAAGCACATATTGGGAAAGGCACGCTTCATAAA
ACTAAATAACCAGTGGAAGGGTCGTTGCTTACACTGCTGGACAACATGCATGGAAGACTTGTTA
CCCTAA

>E7 297bp

ATGCATGGAAGACTTGTTACCCTAAAGGATATAGTACTAGACCTGCAGCCTCCTGACCCTGTAG
GGTTACATTGCTATGAGCAATTAGAAGACAGCTCAGAAGATGAGGTGGACAAGGTGGACAAAC
AAGACTCACAACCTTTAACACAACATTACCAAATACTGACCTGTTGCTGTGGATGTGACAGCAA
CGTCCGACTGGTTGTGGAGTGCACAGACGGAGACATCAGACAACTACAAGACCTTTTGCTGGGC
ACACTAAATATTGTGTGTCCCATCTGCGCACCAAAACCATAA