

Decoding the Genetic and Structural Features of HPV16 E5 Oncogene in Cervical Cancer Isolates from Pakistan: A Pilot Study

Naureen Ehsan Ilahi^{1,2*}, Nayyer Siddique³, Muhammad Ibrahim Rashid², Mamoon Noreen⁴, Sheeba Murad¹

¹Atta-Ur-Rehman School of Applied Biosciences, National University of Science and Technology, H-12 Campus, Hucknall Road, Islamabad, Pakistan; ²Department of Biotechnology, Virtual University of Pakistan, Rawalpindi, Pakistan; ³Department of Molecular Biology and Genetics, Khyber Medical University, Peshawar, Pakistan; ⁴Department of Zoology, The Women University Multan, Multan, Pakistan

OPEN ACCESS

Received: 10 January 2023
Accepted: 12 August 2023
Published online: 23 August 2023

ABSTRACT

Background: Many anogenital cancers are caused by high-risk HPV. The most common subtype is HPV16, which is prevalent in the world, including Pakistan. Various amino acid residues in HPV16 E5 are associated with high cell cycle progression and proliferation. Lack of studies on HPV16E5 in Pakistan prompted the current study. This is the first report on the occurrence of pathogenic E5 variant of HPV16 in tissue sections obtained from invasive cervical cancerous patients in Pakistan.

Methods: A subset of 11 samples from HPV-positive biopsies were subjected to E5 gene amplification using PCR and analyzed using bioinformatics programs. The bioinformatics analysis detected mutations causing structural variations, which potentially contribute to the oncogenic properties of proteins.

Results: The two-point mutations, C3979A and G4042A, observed in isolate 11 caused the substitution of isoleucine for leucine and valine at positions 44 and 65 in E5 protein. The rest of the isolates had Leu44Val65 amino acids. Intratypic variations and phylogenetic analysis revealed that the majority of the isolates were closely clustered with European-Asian lineage. Moreover, C3979A and G4042A contributed to higher degree of interactions with host receptors, i.e. EGFR.

Conclusion: This is the first study reporting HPV16 variants in a Pakistani population based on variations in the E5 region. Our findings indicate that isolate 11 has a strong interaction with the intracellular domain of EGFR, which may enhance the generation of downstream signals. Since this was a pilot study to explore E5 gene mutation, future studies with large samples are absolutely needed. **DOI: 10.61186/ibj.3884**

Citation:

Ilahi NE, Siddique N, Rashid MI, Noreen M, Murad S. Decoding the Genetic and Structural Features of HPV16 E5 Oncogene in Cervical Cancer Isolates from Pakistan: A Pilot Study. *Iranian biomedical journal* 2023; 27(6): 388-396.

Keywords: Human papillomavirus 16, Oncogenes, Uterine cervical neoplasms

Corresponding Author: Naureen Ehsan Ilahi

Atta-Ur-Rehman School of Applied Biosciences, National University of Science and Technology, H-12 Campus, Hucknall Road, Islamabad, Pakistan; Tel.: (+042-111) 880880 (Ext. 8713); E-mail: naureen.ehsan@vu.edu.pk

INTRODUCTION

High-risk types of HPVs play an important role in the etiology of cervical cancer [1-5]. HPV16 is the predominant high-risk subtype responsible

for the majority of HPV-caused malignancies. As a DNA virus, HPV16 genome is relatively stable, though it is believed to have adapted to its host. To date, three oncogenic proteins, E5, E6, and E7, have been identified. Genetic mutation in E5 gene may impact

List of Abbreviations:

EGFR: epidermal growth factor receptor; **HLA-1:** human leukocyte antigens-I; **HPV:** human papillomavirus; **MHC-I:** major histocompatibility complex-1

oncogenicity, antigenicity, and immunogenicity of the encoded proteins. Therefore, reports on the sequence analysis of HPV16 E5 from different parts of the world would be of importance.

Intratype variants of HPV16 have formerly been categorized into four major lineages (A, B, C, and D) centered in E6, L1, L2, and LCR regions^[6,7] and 16 sub-lineages, including A-A4 (European-Asian), B1-B4 (African 1), C1-C4 (African 2), and D1-D4 (North American/Asian-American) lineages^[8]. However, limited data on HPV16 lineages have been reported for Pakistan.

HPV16 E5 is a transmembrane and hydrophobic protein consisting of 83 amino acids^[9]. E5 is known to associate with the membranes of Golgi apparatus and endoplasmic reticulum^[10] and participate in the early stages of the oncogenic process^[11-14]. Recent studies have identified an important role for E5 in the later oncogenic stages. E5 has also been found to link with various host proteins such as MHC-I, Bap31, 16 kDa pore sub-unit of vacuolar-ATPase, and EGFR^[15-18]. Previously, it has been shown that HPV16 E5 binds to EGFR^[18], and cells expressing 16 E5 increases mitogen-activated protein kinase activity, as well as improves responsiveness to EGF^[19,20]. In recent years, there has been a growing interest in early E5 oncoprotein of HPV16 due to its sequence and functional diversity across phylogeny and especially due to its role as a viroporin; controlling vesicle trafficking, virion formation, ion homeostasis, and viral genome entrance 22357280^[21]. E5 mRNA and protein have been found in low-grade cervical intraepithelial neoplasia, indicating that E5 actively participates in the early phases of neoplastic transformation^[11,12]. In addition to E6 and E7, E5 is the third oncoprotein with the potential therapeutic target. E5 gene is divergent between different HPV types. Four different protein subgroups α to δ have been proposed for E5. HPV 16, 18, and 31, belong to the subgroups α ^[22]. The comparative small size of E5 makes it an attractive target for retrospective intratype variant analysis^[23]. Thus, intratype amino acid variations may be relevant to the generation of specific immune

responses, especially in the context of rational vaccine designing^[24].

The current study was designed to identify sequence variation on E5 gene in HPV16-positive cervical carcinoma tissues and to determine whether specific genetic variations can aid in characterizing the HPV16 E5 variants present in our population.

MATERIALS AND METHODS

Sample collection and HPV genotyping

This survey was conducted as an extension of our previously reported work^[25]. A total of 49 paraffin-embedded biopsies were collected for this study. Among 37 tumors positive for HPV 16 and 18 DNA, 25 (51%) and 8 (16%) tumors were positive for HPV 16 DNA and HPV 18, respectively, and coinfection was detected in 4 (8%) tumors. Also, 12 tumors (24%) were negative for both HPV 16 and 18. Among the 37 positive samples, 11 samples were randomly selected to study E5 gene. Tissue samples were obtained from the Armed Forces Institute of Pathology, Rawalpindi (Pakistan) after the formal approval of the study. Histological types of epithelial cancers consisted of squamous cell carcinoma (n = 7) and nonkeratinizing squamous cell carcinoma (n = 4). DNA extraction from samples was performed manually by using the reported protocol^[25]. Extracted DNA was screened for the presence of HPV by GP5⁺/GP6⁺ primers and HPV16 by new TS16 primer (Table 1) as described previously^[26].

E5 gene amplification and sequencing

E5 gene amplification was carried out by a primer set reported earlier by our group^[25] (Table 1). HPV16-positive samples were PCR amplified for E5 (252 bp) product, which was purified using the Gene JET PCR Purification Kit (Thermo Fisher Scientific, USA). The eluted and purified DNA samples of E5 gene were sent to Eurofins Genomics Europe (Germany) for sequence analysis. Sequencing was performed twice by both forward and reverse primers to ensure the nucleotide sequences.

Table 1. Primers used for HPV genotyping

Primers	Sequence (5'-3')	Annealing temperature (°C)	Target gene	Product size (bp)	Reference
GP5 ⁺ /GP6 ⁺	F: TTTGTTACTGTGGTAGATAC R: GAAAAATAAACTGTAAATCA	48	L1	150	[27]
New TS16	F: GGTCCGGTGGACCGGTTCGATG R: GCAATGTAGGTGTATCTCCA	58	HPV16 E6	96	[28]
16-E5	F: GAATTCATGACAAATCTTGATACTG R: GGATCCTTATGTAATTA AAAAAGCGT	54.6	HPV16 E5	252	[25]

Molecular variant detection in HPV16 E5 gene

The obtained E5 sequences were aligned with each other using Multalin software to observe the nucleotide variation^[29]. Sequences were also translated into protein using the translate tool of ExpASY server (<http://web.expasy.org/translate/>) and aligned to determine the amino acid alteration caused due to changes in the nucleotide. E5 sequences from Pakistani isolates (numbered 1-11) were then aligned with the representative sequences of different HPV16 lineages mentioned before^[8], i.e. lineage A: K02718.1^[30], AF536179.1^[31], HQ644236.1^[32], and AF534061.1^[31], lineage B (African 1 [AFR1]): AF536180.1^[31] and HQ644298.1^[32], lineage C (African 2 [AFR2]): AF472509.1^[31,33], and lineage D: AF402678.1^[33,34], AY686579.1^[33] and HQ644257.1^[32]. Phylogenetic analysis was conducted using the freeware Phylogeny (<http://www.phylogeny.fr/>)^[35,36].

Bioinformatics analysis

Identification of epitopes

After analyzing sequences and identifying mutations, we assessed the effects of these mutations on the structure of E5 protein. For this purpose, we first detected the epitopes of E5 protein and then distinguished both MHC I and II epitopes for the reference E5 sequence, isolates 10 and 11 using TepiTool (<http://tools.iedb.org/tepitool/>). This tool predicted Class I and II MHC epitopes using IC₅₀ value <500 nM.

Structural analysis of the mutated E5 proteins

Mutations at the DNA level can impact the protein structure by altering the amino acids, since some alterations are crucial in protein-protein interaction. To examine the effect of mutations on protein-protein interaction, we studied the structures of the most common isolates, i.e. 11 and 10, using the FALCON2 server (<http://falcon.ictda.cn:89/#home>). To identify differences between these structures and the reference sequence, we made a comparison with the YASARA MUSTANG method using YASARA visualization software (version 23.9.29)^[37]. To further investigate the impact of these mutated proteins on the EGFR downstream signaling pathway, we analyzed the interaction of these structures with the EGFR protein.

The structures PDB id: 1M14 and 3NJP were obtained from the Protein Data Bank server (<https://www.rcsb.org/>)^[38]. The interaction between the two structures was then analyzed using the docking web server ClusPro 2.0 (<https://cluspro.bu.edu/login.php>)^[39]. The energy profiles were evaluated based on the electrostatic and van der Waals interactions using the following equation: $E = 0.40E_{rep} + -0.10E_{att} + 600E_{elec} + 1.00E_{DARS}$. Models with lower energy were selected and visualized using Biovia Discovery Studio 2021 (<https://www.3ds.com/products/biovia/discovery-studio/ligand-and-pharmacophore-based-design>).

RESULTS

In the current study, HPV16 E5 gene was amplified from cervical cancer patients. Sequences obtained after PCR and DNA sequencing were deposited in the GenBank (accession nos. KX758540 and KX822075-84). Variations in nucleotide sequences of E5 open reading frame in the isolates are shown in Table 2 and Figure 1. Sequence variant from isolate 11 contained two point mutations, C3979A and G4042A, in comparison with the nucleotide sequence of HPV16 reference sequence (KF880690.1)^[40]. In isolate 10, the following substitutions were observed: C3991T and G4017A. Both isolates 10 and isolate 11 belonged to the poorly and moderately differentiated squamous cell carcinoma.

A consensus sequence of the E5 was generated from all the isolates. Analysis of E5 nucleotide sequences showed 96-99% similarity between the study samples and the reference sequence (data not shown). Upon alignment with molecular variants of E5 originating from distinct lineages, the analysis unveiled the presence of five significant variations (Table 3). These variations exhibited a high degree of conservation within the majority of samples included in the current study, predominantly among lineage A (European-Asian lineage) variants. The majority of the African and American lineages had a different nucleotide at these positions (Table 3). On the basis of intratypic variations and phylogenetic analysis, it was observed that majority of the isolates were closely clustered with European-Asian lineage (Fig. 2).

Table 2. Minor nucleotide variation observed in HPV16 E5 gene in the isolates and other HPV16 variant lineages

Sequence no.	Isolate no.	Nucleotide position in HPV16 E5 gene (3850-4101)	Type of variation and nucleotide change	HPV16 lineage with similar change
1	11	3979	Transversion C → A	A1
2	10	3991	Transition/transversion C/G → T	B1, B2, D1
3	10	4017	Transition G → A	D1-D3
4	11	4042	Transition G → A	A1, A2

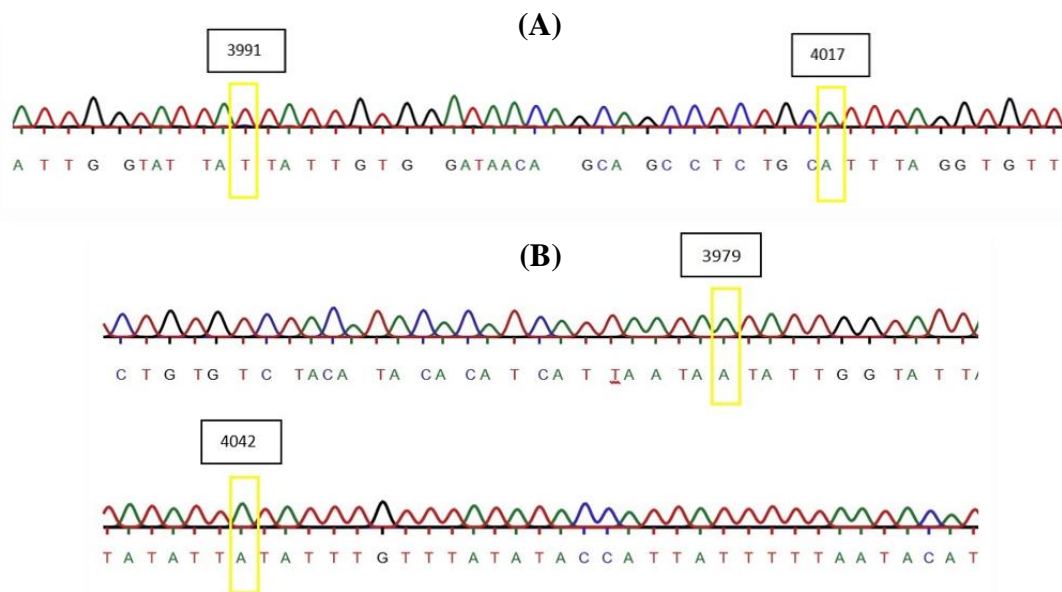


Fig. 1. Sequencing chromatogram of mutations observed in (A) isolate 10 and (B) isolate 11.

The TepiTool software identified approximately 100 epitopes of the E5 proteins for each reference, as well as for the isolates 10 and 11 (Table 4 and Fig. 3). Epitopes containing mutations have different IC₅₀ values compared to the reference sequence. Epitopes with an IC₅₀ value higher than 250 nmol can evade the immune system, since a higher viral load is required to trigger an immune response.

To analyze the effects of mutations on the protein structures, a structure-structure alignment was performed. The findings revealed a difference with an RMSD value of 0.544 Å over 81 aligned residues and a 97.53% sequence identity (Fig. 4). This structural variation may contribute to the oncogenic properties of the different E5 variants. The interaction between E5 and the EGFR protein was examined through docking analysis. The results clearly demonstrated that the E5 protein of isolate 11 had more interactions (eight hydrogen bonds) with EGFR compared to the reference sequence (five hydrogen bonds). Moreover, the mutated isolate showed a stronger interaction with the intracellular domain of EGFR than the reference

sequence, which may enhance the activation of the downstream signaling pathways (Fig. 5).

DISCUSSION

The current study aimed to investigate the presence of sequence variants of HPV16 E5 in cervical cancer biopsies of a Pakistani population. We confirmed the etiological involvement of HPV16 in the same cases of cervical cancer in our previous study using PCR and immunohistochemistry^[25]. At present, we amplified the E5 gene by PCR from 11 HPV16-infected cervical cancer tissues. Heterogeneity in the viral oncogenes of HPV16 was found in a particular geographical region, which would be of functional significance. Studies have indicated that the prevalence of cervical cancer in different countries can be attributed to the distribution of specific viral variants of HPV genome, as HPV variants are differently dispersed among geographic regions^[41,42].

Table 3. Major nucleotide variation observed in HPV16 E5 gene in the isolates and other HPV16 variant lineages

Sequence no.	Isolate no.	Nucleotide position in HPV16 E5 gene (3850-4101)	Type of variation and nucleotide change	HPV16 lineage with similar change
1	1-10	3979	Transversion A → C	All lineages except A1
2	1-9, 11	3991	Transition/transversion T/G → C	A1, A2, A4
3	1-9, 11	4017	Transition A → G	A1-A4, C, B1, B2
4	1-10	4042	Transition A → G	A3, A4, C, D1-D3
5	1-11	4089	Transition C → T	A1-A4

Table 4. Selected MHC I and MHC II epitopes covering the mutated region

MHC I	41-49	42-50	43-51	61-69	63-71	64-72	65-73
Reference E5	416.10	190.53	205.81	495.27	348.18	492.2	152.21
Isolate 10	416.10	190.53	205.81	495.27	348.18	492.2	
Isolate 11	214.67	195.38		465.63	221.09	224.4	143.73
MHC II	33-47	51-65	58-72	63-77			
Reference E5	285.58	488.90	495.49	452.16			
Isolate 10	285.58	488.90	495.49	452.16			
Isolate 11	330.62	472.44	193.50	478.47			

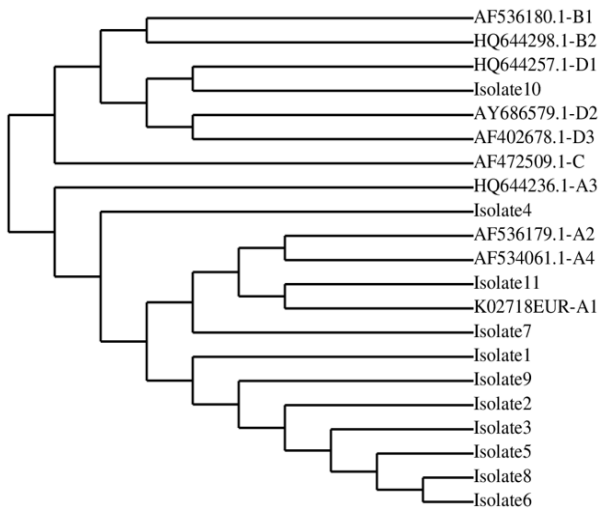


Fig. 2. A phylogenetic tree based on HPV16 E5 analyses of 11 isolates (I-11) and reference sequences of different variants obtained from the GenBank database. A1-A4, B1-B2, C, and D1-D3 show European-Asian, African 1, African 2, and Asian American/North American lineages, respectively.

Molecular analysis of E5 gene in our samples revealed more conservation than variation in the nucleotide sequences. In a limited number of samples, only nucleotide substitutions with no deletion/insertion were identified. Four-point mutations were also observed in the isolates 10 (C3991T and G4017A) and 11 (C3979A and G4042A). In isolate 11, there were two non-synonymous mutations, which led to the replacement of leucine (L) at position 44 and valine at 65 by isoleucine (I) amino acid. Formerly, a functional analysis of this mutation revealed the importance of leucine and valine as α -helix stabilizers at the mentioned positions; however, substitution at amino acid 44 and 65 by isoleucine, a helix destabilizer, affected the structural integrity of E5 protein^[43]. All other amino acids were fully conserved in the isolates. Similar to all E5 proteins of A9 group of papillomaviruses, the consensus sequences from the isolates belonged to E5 α family^[22].

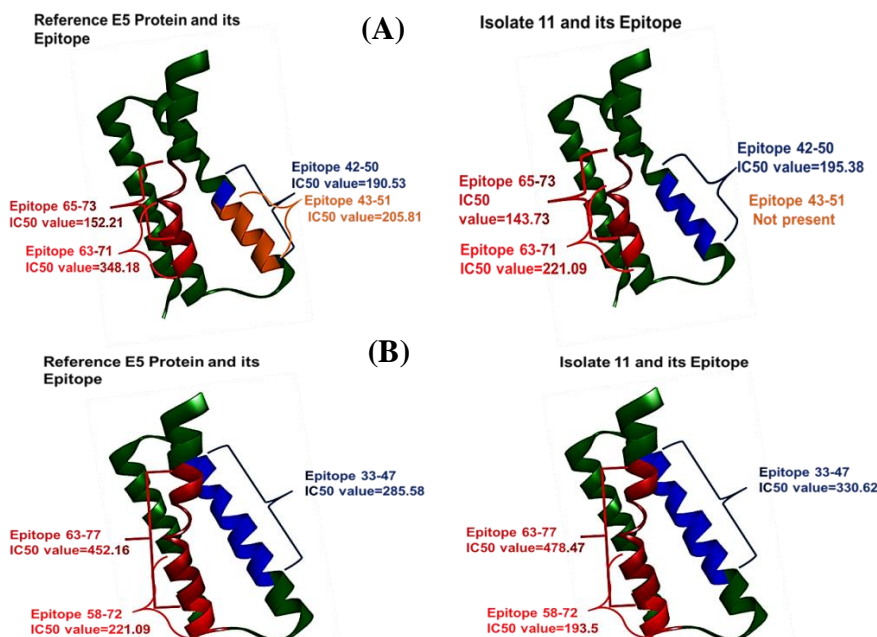


Fig. 3. Different (A) MHC I and (B) MHC II epitopes of the reference E5 protein from isolate 11 exhibiting notable variations in their IC₅₀ values, which allow clear differentiation.

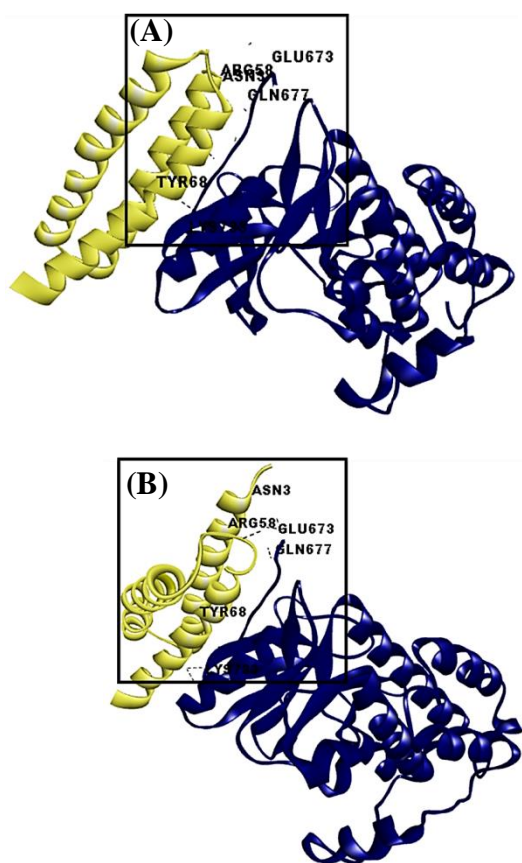


Fig. 5. Interaction between the reference E5 protein (yellow) and EGFR (blue). (A) Reference E5 protein and EGFR and (B) isolate 11 and EGFR. The structural difference observed between the reference and isolate 11 may contribute to more interaction between the two proteins.

In a previous study, similar to our study, no substitutions have been noted within amino acids 11-24^[44]. This region forms a stretch of hydrophobic residues, potentially representing a transmembrane helical region. This transmembrane region has been introduced as the most conserved segment of E5 among HPV types of the A9 HPV group^[45]. The sequence conservation observed in E5 can be attributed to the fact that E5 sequence variants have more than 98% similarity among their nucleotide sequences^[46]. Additionally, the fact that almost identical strains have been found in unrelated people living in geographically different places without any known interaction is evidence of the sequence conservation in HPV16 E5 variations. Our study showed that HPV16 E5 variants in Pakistan are mostly related to one distinct group of HPV16 variants (European-Asian type). Sequence variations in E5 have not formerly been used to characterize the variants. Therefore, we attempted to connect the variations in E5 of different lineages to the ones included in our study.

Interestingly, a high prevalence of European lineage found in the isolates, indicating a possible epidemiological link between Europe and Pakistan regarding the dissemination of HPV16 infections in Pakistan. Nevertheless, due to the small sample size, this finding should be verified by larger studies.

Our research focused on bioinformatics analysis of sequences, with particular emphasis on identification of epitopes for the reference sequence and the two primary isolates (10 and 11), due to their notable mutations. The epitope selection was based on the targeting the ones with significant mutations, guided by the understanding that E5 of HPV16 physically interacts with the heavy chain component of MHC-I antigen. This interaction retains HLA -I in Golgi bodies and endoplasmic reticulum prevents its transport to the cell surface, which consequently interferes with the immune presentation of viral peptides and reduces recognition by cytotoxic T lymphocytes^[47,48].

Our study assessed the structural variations resulting from the mutations C3979A and G4042A, specifically examining the interaction between E5 proteins and human EGFR. E5 protein activates EGFR, which triggers downstream signaling pathways, a phenomenon previously linked to protection against apoptosis^[49]. Altogether, our findings suggest the potential activation of downstream signaling pathway due to a strong interaction of mutant E5 with EGFR. This dual focus on epitope identification and structural variations enhances our understanding of how HPV16 E5 may modulate immune recognition and downstream signaling and offers valuable insights for future research and potential therapeutic interventions.

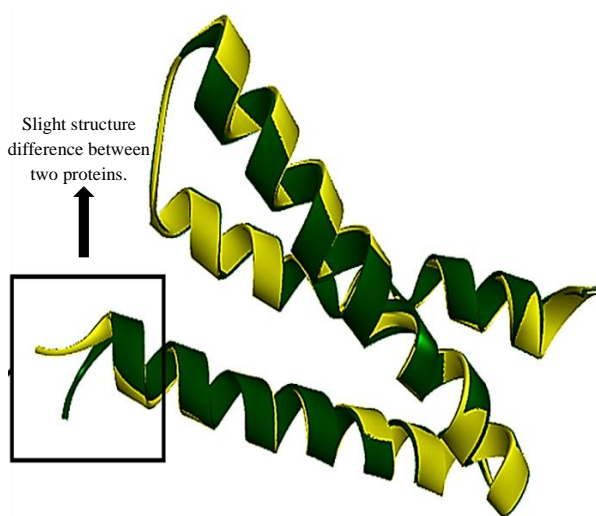


Fig. 4. Superimposed structure of Wild type Reference E5 protein (green) and isolate 11 (yellow) shows a noticeable difference in their structures.

DECLARATIONS

Acknowledgments

We are obliged to the Higher Education Commission of Pakistan for providing a chance to conduct research.

Ethical statement

This article does not contain any studies with human participants or animals performed by any of the authors.

Data availability

The raw data supporting the conclusions of this article are available from the corresponding author upon reasonable request.

Author contributions

NEI: conceptualization, experimentation, formal analysis, original draft, writing-review, and editing; NS and MI: bioinformatics analysis, writing-review, and editing; MN: formal analysis and experimentation; SM: funding acquisition, project administration, supervision, validation, and resources.

Conflict of interest

None declared.

Funding/support

A recurring budget of Attar-Ur-Rahman School of Applied Biosciences was provided by National University of Sciences and Technology (Islamabad, Pakistan) for the financial support of this project.

REFERENCES

- Londesborough P, Ho L, Terry G, Cuzick J, Wheeler C, Singer A. Human papillomavirus genotype as a predictor of persistence and development of high-grade lesions in women with minor cervical abnormalities. *International journal of cancer* 1996; **69**(5): 364-368.
- Zehbe I, Voglino G, Delius H, Wilander E, Tommasino M. Risk of cervical cancer and geographical variations of human papillomavirus 16 E6 polymorphisms. *Lancet* 1998; **352**(9138): 1441-1442.
- Zehbe I, Wilander E, Delius H, Tommasino M. Human papillomavirus 16 E6 variants are more prevalent in invasive cervical carcinoma than the prototype. *Cancer Research* 1998; **58**(4): 829-833.
- Xi LF, Demers GW, Koutsky LA, Kiviat NB, Kuypers J, Watts DH, Holmes KK, Galloway DA. Analysis of human papillomavirus type 16 variants indicates establishment of persistent infection. *The journal of infectious diseases* 1995; **172**(3): 747-755.
- Xi LF, Koutsky LA, Hildesheim A, Galloway DA, Wheeler CM, Winer RL, Ho J, Kiviat NB. Risk for high-grade cervical intraepithelial neoplasia associated with variants of human papillomavirus types 16 and 18. *Cancer epidemiology, biomarkers and prevention* 2007; **16**(1): 4-10.
- Ho L, Chan SY, Burk RD, Das BC, Fujinaga K, Icenogle JP, Kahn T, Kiviat N, Lancaster W, Mavromara-Nazos P. The genetic drift of human papillomavirus type 16 is a means of reconstructing prehistoric viral spread and the movement of ancient human populations. *Journal of virology* 1993; **67**(11): 6413-6423.
- Yamada T, Wheeler CM, Halpern AL, Stewart AC, Hildesheim A, Jenison SA. Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, L2, and L1 coding segments. *Journal of virology* 1995; **69**(12): 7743-7753.
- Burk RD, Harari A, Chen Z. Human papillomavirus genome variants. *Virology* 2013; **445**(1-2): 232-243.
- Borzacchiello G, Roperto F, Campo MS, Venuti A. 1st International workshop on Papillomavirus E5 oncogene-a report. *Virology* 2010; **408**(2): 135-137.
- Conrad M, BubbVJ, Schlegel R. The human papillomavirus type 6 and 16 E5 proteins are membrane-associated proteins which associate with the 16-kilodalton pore-forming protein. *Journal of virology* 1993; **67**(10): 6170-6178.
- Stoler MH, Rhodes CR, Whitbeck A, Wolinsky SM, Chow LT, Broker TR. Human papillomavirus type 16 and 18 gene expression in cervical neoplasias. *Human pathology* 1992; **23**(2): 117-128.
- Kell B, Jewers RJ, Cason J, Pakarian F, Kaye JN, Best JM. Detection of E5 oncoprotein in human papillomavirus type 16-positive cervical scrapes using antibodies raised to synthetic peptides. *The journal of general virology* 1994; **75**(Pt 9): 2451-2456.
- Biswas C, Kell B, Mant C, Jewers RJ, Cason J, Muir P, Raju KS, Best JM. Detection of human papillomavirus type 16 early-gene transcription by reverse transcription-PCR is associated with abnormal cervical cytology. *Journal of clinical microbiology* 1997; **35**(6): 1560-1564.
- Hsieh CH, Tsao YP, Wang CH, Han CP, Chang JL, Lee JY, Chen SL. Sequence variants and functional analysis of human papillomavirus type 16 E5 gene in clinical specimens. *Archives of virology* 2000; **145**(11): 2273-2284.
- Regan JA, Laimins LA. Bap31 is a novel target of the human papillomavirus E5 protein. *Journal of virology* 2008; **82**(20): 10042-10051.
- Cortese MS, Ashrafi GH, Campo MS. All 4 di-leucine motifs in the first hydrophobic domain of the E5 oncoprotein of human papillomavirus type 16 are essential for surface MHC class I downregulation activity and E5 endomembrane localization. *International journal of cancer* 2010; **126**(7): 1675-1682.
- Rodriguez MI, Finbow ME, Alonso A. Binding of human papillomavirus 16 E5 to the 16 kDa subunit c (proteolipid) of the vacuolar H⁺-ATPase can be dissociated from the E5-mediated epidermal growth factor receptor overactivation. *Oncogene* 2000; **19**(33): 3727-3732.
- Hwang ES, Nottoli T, Dimaio D. The HPV16 E5 protein: expression, detection, and stable complex formation with

- transmembrane proteins in COS cells. *Virology* 1995; **211**(1): 227-233.
19. Gu Z, Matlashewski G. Effect of human papillomavirus type 16 oncogenes on MAP kinase activity. *Journal of virology* 1995; **69**(12): 8051-8056.
 20. Zyzak LL, MacDonald L M, Batova A, Forand R, Creek KE, Pirisi L. Increased levels and constitutive tyrosine phosphorylation of the epidermal growth factor receptor contribute to autonomous growth of human papillomavirus type 16 immortalized human keratinocytes. *Cell growth and differentiation: the molecular biology journal of the american association for cancer research* 1994; **5**(5): 537-547.
 21. Wetherill LF, Holmes KK, Verow M, Müller M, Howell G, Harris M, Fishwick C, Stonehouse N, Foster R, Eric Blair G, Griffin S. High-risk human papillomavirus E5 oncoprotein displays channel-forming activity sensitive to small-molecule inhibitors. *Journal of virology* 2012; **86**(9): 5341-5351.
 22. Bravo IG, Alonso A. Mucosal human papillomaviruses encode four different E5 proteins whose chemistry and phylogeny correlate with malignant or benign growth. *Journal of virology* 2004; **78**(24): 13613-13626.
 23. Hao M, Wu XZ, XL Wang, Bian ML, Liu HT. [Human papillomavirus type 16 E5 sequence evolution analysis]. *Chinese journal of experimental and clinical virology* 2009; **23**(6): 452-454.
 24. Yamada T, Manos MM, Peto J, Greer CE, Munoz N, Bosch F X, Wheeler CM. Human papillomavirus type 16 sequence variation in cervical cancers: a worldwide perspective. *Journal of virology* 1997; **71**(3): 2463-2472.
 25. Ilahi NE, Naiyar Hashmi S, Anwar S, Murad S. Retrospective analysis of HPV 16/18-related disease burden using archival clinical samples. *Journal of cancer research and clinical oncology* 2016; **142**(11): 2367-2373.
 26. Siddiq A, Zainab M, Qadri I, Bhatti MF, Parish JL. Prevalence and genotyping of high risk human papillomavirus in cervical cancer samples from Punjab, Pakistan. *Viruses* 2014; **6**(7): 2762-2777.
 27. de Roda Husman AM, Walboomers JM, van den Brule AJ, Meijer CJ, Snijders PJ. The use of general primers GP5 and GP6 elongated at their 3' ends with adjacent highly conserved sequences improves human papillomavirus detection by PCR. *The journal of general virology* 1995; **76**(Pt 4): 1057-1062.
 28. Baay MF, Quint WG, Koudstaal J, Hollema H, Duk JM, Burger MP, Stolz E, Herbrink P. Comprehensive study of several general and type-specific primer pairs for detection of human papillomavirus DNA by PCR in paraffin-embedded cervical carcinomas. *Journal of clinical microbiology* 1996; **34**(3): 745-747.
 29. Corpet F. Multiple sequence alignment with hierarchical clustering. *Nucleic acids research* 1988; **16**(22): 10881-10890.
 30. Kennedy IM, Haddow JK, Clements JB. A negative regulatory element in the human papillomavirus type 16 genome acts at the level of late mRNA stability. *Journal of virology* 1991; **65**(4): 2093-2097.
 31. National Library of Medicine. Human papillomavirus type 16 isolate European German 131 type, complete genome. Available at: <https://www.ncbi.nlm.nih.gov/nuccore/AF536179.1>.
 32. Smith B, Chen Z, Reimers L, van Doorslaer K, Schiffman M, Desalle R, Herrero R, Yu K, Wacholder S, Wang T, Burk RD. Sequence imputation of HPV16 genomes for genetic association studies. *PLoS one* 2011. **6**(6): e21375.
 33. Chen Z, Terai M, Fu L, Herrero R, DeSalle R, Burk RD. Diversifying selection in human papillomavirus type 16 lineages based on complete genome analyses. *Journal of virology* 2005; **79**(11): 7014-7023.
 34. National Library of Medicine. Human papillomavirus type 16 Asian-American variant, complete genome Available at: <https://www.ncbi.nlm.nih.gov/nuccore/AF402678>.
 35. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM, Gascuel O. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic acids research* 2008; **36**(Web Server issue): W465-469.
 36. Dereeper A, Audic S, Claverie JM, Blanc G. BLAST-EXPLORER helps you building datasets for phylogenetic analysis. *BMC evolutionary biology* 2010; **10**: 8.
 37. Konagurthu AS, Whisstock JC, Stuckey PJ, Lesk AM. MUSTANG: a multiple structural alignment algorithm. *Proteins* 2006; **64**(3): 559-574.
 38. Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, Feng Z, Gilliland GL, Iype L, Jain S, Fagan P, Marvin J, Padilla D, Ravichandran V, Schneider B, Thanki N, Weissig H, Westbrook JD, Zardecki C. The protein data bank. *Acta crystallographica. Section D, Biological crystallography* 2002; **58**(6): 899-907.
 39. Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, Beglov D, Vajda S. The ClusPro web server for protein-protein docking. *Nature protocols* 2017; **12**(2): 255-278.
 40. Letsolo BT, Faust H, Ekblad L, Wennerberg, Ola Forslund J. Establishment and characterization of a human papillomavirus type 16-positive tonsillar carcinoma xenograft in BALB/c nude mice. *Head and neck* 2016; **38**(3): 417-425.
 41. Tornesello ML, Duraturo ML, Buonaguro L, Vallefucio G, Piccoli R, Palmieri S, Buonaguro FM. Prevalence of human papillomavirus genotypes and their variants in high risk West Africa women immigrants in South Italy. *Infectious agents and cancer* 2007; **2**: 1.
 42. Stoppler MC, Ching K, Stöppler H, Clancy K, Schlegel R, Icenogle J. Natural variants of the human papillomavirus type 16 E6 protein differ in their abilities to alter keratinocyte differentiation and to induce p53 degradation. *Journal of virology* 1996; **70**(10): 6987-6993.
 43. Nath R, Mant CA, Kell B, Cason J, Bible JM. Analyses of variant human papillomavirus type-16 E5 proteins for their ability to induce mitogenesis of murine fibroblasts. *Cancer cell international* 2006; **6**(1): 19.
 44. Eriksson A, Herron JR, Yamada T, Wheeler CM. Human papillomavirus type 16 variant lineages characterized by nucleotide sequence analysis of the E5 coding segment

- and the E2 hinge region. *The journal of general virology* 1999; **80**(Pt 3): 595-600.
45. Halpern AL, McCance DJ. Papillomavirus E5 Proteins. Transformation, 1996. **9**: 52. Available at: <https://pave.niaid.nih.gov/lanl->
46. de Villiers EM, Fauquet C, Broker TR, Bernard HU, zur Hausen H. Classification of papillomaviruses. *Virology* 2004; **324**(1): 17-27.
47. Ashrafi GH, Haghshenas M, Marchetti B, Campo MS. E5 protein of human papillomavirus 16 downregulates HLA class I and interacts with the heavy chain via its first hydrophobic domain. *International journal of cancer* 2006; **119**(9): 2105-2112.
48. Campo MS, Graham SV, Cortese MS, Ashrafi GH, Araibi EH, Dornan ES, Miners K, Nunes C, Man S. HPV-16 E5 down-regulates expression of surface HLA class I and reduces recognition by CD8 T cells. *Virology* 2010; **407**(1): 137-142.
49. Zhang B, Spandau DF, Roman A. E5 protein of human papillomavirus type 16 protects human foreskin keratinocytes from UV B-irradiation-induced apoptosis. *Journal of virology* 2002; **76**(1): 220-231.