GMR
The Original by FUNPEC-RP

E6 and E7 gene sequences of human papillomavirus type 13 isolated from a young native girl from the Brazilian Amazon

C.M. Borborema-Santos[1], N.S. de Lima Freitas[1], P.J.B. Santos[2], R.A. Barbosa-Filho[1], E. Costa-Lira[1], S. Talhari[3] and S. Astolfi-Filho[1]

[1] Programa de Pós-Graduação em Biotecnologia, Universidade Federal do Amazonas, Manaus, AM, Brasil
[2] Fundação Centro de Controle de Oncologia do Estado do Amazonas, Manaus, AM, Brasil
[3] Fundação de Medicina Tropical Doutor Heitor Vieira Dourado, Manaus, AM, Brasil

Corresponding author: C.M. Borborema-Santos
E-mail: crisantos@ufam.edu.br

Genet. Mol. Res. 19 (1): gmr18393
Received June 19, 2019
Accepted January 21, 2020
Published March 30, 2020
DOI http://dx.doi.org/10.4238/gmr18393

ABSTRACT. Human papillomavirus (HPV) type 13 is associated with focal epithelial hyperplasia, a benign and self-limited disease of oral mucosa. It is a rare finding, most often seen among young American natives. We carried out a molecular characterization case study of HPV type 13 isolated from a 14-year-old native girl living in Amazonia, in the northern region of Brazil, by describing the E6 and E7 open reading frame (ORF) sequences and comparing them to the only reference sequence in GenBank, which came from a Turkish isolate. The E6 and E7 ORFs were amplified by PCR, and the amplicons were purified and cloned in *E. coli* using the pCR®-4 TOPO® vector. The purified recombinant plasmids were automatically sequenced, showing 98 and 94% similarity to the E6 and E7 ORFs from the reference sequence, respectively. Six mutations were identified in the E6 ORF, but these changed only three amino acids in the final protein. In the E7 ORF, we found three transversions and one deletion of nine base pairs, which led to three amino acid mutations and three amino acid deletions in the final protein. The mutations in E6 ORF gave no predicted changes in the final protein, while the mutations in E7 were predicted to cause changes in the

protein. We conclude that the E6 and E7 ORFs in HPV-13 from the Amazonia and Turkey isolates have a high degree of genetic similarity.

Key words: Human papillomavirus; HPV-13; E6/E7 genes; Focal epithelial hyperplasia; Polymerase Chain Reaction; Brazil

INTRODUCTION

The group of human papillomaviruses (HPV) is quite extensive and heterogeneous. These non-enveloped, icosahedra viruses possess a genome of double stranded DNA, approximately 8,000 bp in length. The genome is organized into early "E" and late "L" genes. Some early genes possess proliferation and stimulating activity (E5, E6 and E7), while others are regulatory (E1, E2 and E4); late genes, on the other hand, encode structural proteins (L1 and L2) of the viral capsid zur Hausen (2002). HPV have traditionally been classified into types and further classified into high or low risk groups, depending on their association with malignant versus benign lesions, or cutaneous versus mucosal lesions (Gross and Barrasso, 1999). Taxonomically, HPV types have been classified to genus and species. According to de Villiers et al. (2004), a newly isolated HPV must have its complete genome cloned, with the L1 gene more than 10% different relative to the closest known HPV, in order to be classified as a new type. Differences between 2% and 10% define a subtype, and less than 2% defines a variant.

Human papillomavirus type 13 is associated with focal epithelial hyperplasia (FEH), a benign and self-limited disease of oral mucosa. Focal epithelial hyperplasia or Heck's disease is a rare contagious disease caused by human papillomavirus that was first described in 1965, most commonly found among young American native individuals, North-American Eskimos and South Africans but has also been found in other ethnic groups (Archard et al., 1965; Terezhalmy et al., 2001; González et al., 2005; Borborema-Santos et al., 2006). According to Carlos and Sedano (1994), HPV-13 is frequently found exclusively in the mouth, but there are some reports of it in other mucosal sites (Rolighed et al., 1992; Santos et al., 2005; Castro et al., 2011). The complete genome of HPV-13 was sequenced and deposited in GenBank by Van Ranst et al. (1992) based on a sample from a 13-year-old Turkish girl (GenBank accession no. X62843). It is 7,880 bp in length and contains all of the early and late open reading frames (ORF) described for HPV.

This is a case study of a young native girl from the Central Amazon region who sought treatment at the Amazonas State Foundation of Tropical Medicine (FMT-AM) in Manaus, the capital city of the state of Amazonas (AM), Brazil. A HPV-13 sample was isolated from this 14-year-old girl and was sequenced in the Biotechnology Division / Molecular Diagnostic Laboratory of Federal University of Amazonas. Afterwards, the HPV-13 findings were deposited in GenBank under accession no. DQ344807.

Since E6 and E7 are important oncogenes and given that there are some reports of HPV-13 in malignant and pre-malignant diseases (Rolighed et al., 1992; Castro et al., 2011), we sequenced and analyzed the E6 and E7 ORFs from the virus that we isolated the Amazon region girl and compared them to the reference sequences of the Turkish isolate in GenBank.

MATERIAL AND METHODS

All the benefits and risks involved in participating in the study were made clear to the patient, who sought out the services at the Amazonas State Foundation of Tropical Medicine (FMT-AM) and signed an informed consent form. This study was approved by the Human

Research Ethics Committee at the Federal University of Amazonas and performed in accordance with the Declaration of Helsinki.

Clinical samples

The sample was collected in 2002 with a cervical brush by scraping the oral mucosa of the source patient, a 14-year-old native American girl from the Central Brazilian Amazon, and subsequently re-suspended in 400 µL of TE (10 mM Tris-HCl pH 8.0 and 1 mM EDTA) and stored at -20°C until further processing.

Preparation of the samples for PCR assay

DNA extraction and purification procedures were performed. The sample was placed in a 400 µL TPK solution [TE, 20% Tween, 10 mg/mL proteinase K], followed by incubation at 56°C for 1 h and then boiling for 10 min. Afterwards, the phenol-chloroform procedure was followed, as described by Sambrook et al. (1989).

PCR amplification and cloning of the HPV E6 and E7 sequences

The HPV-positive DNA sample was used to amplify a 1242 bp fragment of the complete E6 and E7 genes by PCR using primer sets of forward (5'ATAggTgggACCgAAAACgg3') HPV13 27-46 bp (20 mer) and reverse (5'CCACTTCAgAATAgCCATATCC3') HPV13 1269 - 1248 bp (22 mer). The PCR reaction consisted of 5.0 µL of DNA, 5.0 µL of 10X PCR buffer, 2.0 µL of 50 mM MgSO4, 1.0 µL of 10mM dNTP, 5.0 µL of each primer (5.0 pmol/µL), 0.2 µL of 1U/µL T*aq* DNA polymerase high fidelity, and 26.8 µL Milli-Q water to complete the 50 µL volume. The amplification was performed in a PXE 0.2 thermal cycler (Thermo Electron Corp) thermocycler using the following program with 40 amplification cycles: pre-denaturing at 94°C for 30 s, denaturing at 94°C for 30 s, primer annealing at 58°C for 30 s, and primer extension at 68°C for 5 min, followed by a final extension at 68°C for 5 min. Following amplification, fragments were purified and cloned in *E. coli* using the pCR®-4 TOPO® vector (Invitrogen Life Technologies). The purified recombinant plasmids were automatically sequenced using the Dye-terminator method (MegaBace 1000 System – Amersham Biosciences). Sequences were aligned using the CLUSTAL W multiple sequence alignment program (Thompson et al., 1998). The aligned sequences were translated into amino acids, and mutations were defined as divergence from the HPV 13 described by Van Ranst et al. (1992), which was considered to be the reference sequence. The similarity of the generated nucleotide and amino acids sequences was determined using the Basic Local Alignment Search (BLAST), available at http://www.ncbi.nlm.nih.gov (Altschul et al., 1997). At least three independent plasmid sequences were generated for each forward and reverse primer. Sequence changes found three times were considered to be variants.

RESULTS

Sequencing of the E6 ORF revealed close similarity (98%) between HPV-13 from the Amazonian and Turkish sample. In this 453 bp ORF, neither insertions nor deletions were found, but six mutations were identified. Five mutations were transversions (nt 139: t → g; nt 236: g → t; nt 276: a → c; nt 541: a → c; nt 547: c → a), while only one was a transition (nt 275: c → t). Although six mutations were identified, only three amino acids were changed in the protein: S

(serine) > A (alanine), S (serine) > Q (glutamine) and D (aspartic acid) >E (glutamic acid), as shown in Figure 1.

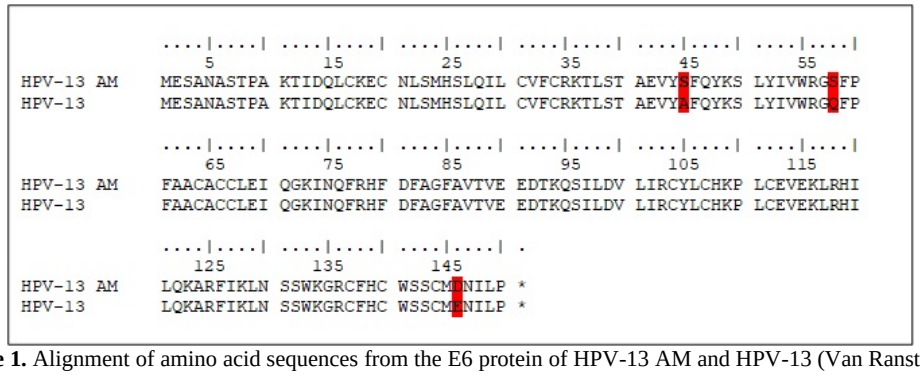

Figure 1. Alignment of amino acid sequences from the E6 protein of HPV-13 AM and HPV-13 (Van Ranst et al., 1992) shows three mutations. Three amino acids are changed in the protein: S (serine) > A (alanine), S (serine) > Q (glutamine) and D (aspartic acid) >E (glutamic acid). Red amino acids.

The E7 ORF from the Amazonian HPV-13 isolated showed 94% similarity to the Turkish sample and was 297 bp in length. In this ORF, we identified three transversions (nt 541: a -> c; nt 547: c -> a; nt 573: g-> c) and one deletion of nine base pairs (nt 647 – 655: CACGCAAGC). Neither insertions nor transitions were found in the E7 ORF. The amino acids sequence showed three amino acid mutations: K (lysine) > Q (glutamine), E (glutamic acid) > D (aspartic acid), P (proline) > T (threonine) and a deletion of three amino acids: alanine, threonine and glutamine. The E7 protein of the Amazonian HPV-13 isolate was 98 amino acids in length rather than the 101 amino acids reported for the Turkish sample described by Van Ranst et al. (Santos et al., 2005). A comparison between the two proteins was made (Figure 2).

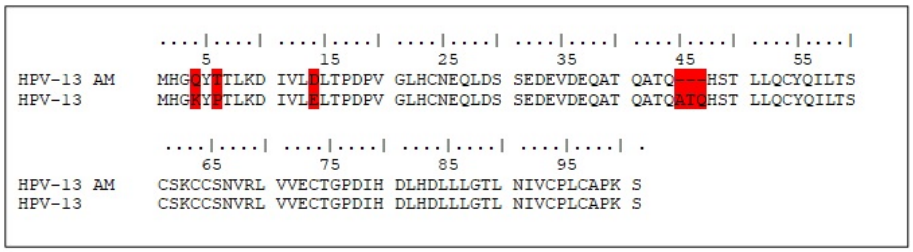

Figure 2. Alignment of amino acid sequences from the E7 protein of HPV-13 AM and HPV-13 (Van Ranst et al., 1992) shows three mutations: K (lysine) > Q (glutamine), E (glutamic acid) > D (aspartic acid), P (proline) > T (threonine) and a deletion of three amino acids: alanine, threonine and glutamine.

DISCUSSION

The mutations in the E6 ORF led to no major predicted changes in the final protein. Although a deletion of nine nucleotides was found in the E7 ORF, the reading frame was not shifted, and there were no other major protein modifications predicted. In the E7 protein, the lysine, glutamine, glutamic acid, aspartic acid and threonine residues have a negative hydropathic index, while proline has a positive hydrophobic index. Three

mutations were synonymous, while only one mutation was not [P (proline) > T (threonine)], predicting changes in the final protein (Hoop and Woods, 1983; Ying et al., 2010).

The E6 and E7 ORFs in HPV-13 from our Amazonian isolate are very similar to the same ORFs in HPV-13 from the Turkish isolate reported by Van Ranst et al. (1992). Due to the high genetic similarity between HPV-13 isolated from Amazonia and the HPV-13 isolate (Van Ranst et al., 1992), we could presume that FEH had the same clinical characteristics described by Archard et al. (1965) and confirmed by Pfister et al. (1983) when describing the viral etiology and HPV-13 characterization from a 13-year-old Turkish girl with FEH lesions. As the molecular analysis allowed a conclusive diagnosis, no histological examination was performed for diagnostic purposes in this study. The location of FEH lesions (lip, tongue and buccal mucosa) is very characteristic (Borborema-Santos et al., 2006).

Using PCR, followed by sequencing, Castro et al. (2011) reported the presence of HPV-13 in a case of cervical cancer in Manaus – Amazonas. Rolighed et al. (1992) reported a case of bowenoid papulosis in an HIV-positive male using in situ hybridization, but HPV 16, 6, 11 and 33 were also detected. One of us, PJB Santos, (unpublished results) also detected HPV 13 in three of 36 (3/36) oral cancer patients in Manaus, AM.

The methodologies used by Castro et al. (2011) and PJB Santos (unpublished results) (PCR following the sequencing technique), showed that multiple types of HPV infections (HPV types associated with non-malignant oral lesions) could not be detected. According to de Villiers et al. (2004), these types of HPV are associated with low-risk benign mucosal lesions.

Santillán (2003) found that cases of this disease are rare. Ethnic differences in the prevalence of specific HPV types have not been reported, except for Heck's disease or focal epithelial hyperplasia produced by HPV types 13 and 32 among American Indians (Archard et al., 1965) and North-American Eskimos (Jarvis and Gorlin, 1972). According to Cheah and Looi (1998) and Castro et al. (2011), little is known about the interrelationship between oral and cervical HPV infections. Rarely, papilloma type 13 can undergo malignant transformation. Therefore, further studies of proteins E6 and E7 from HPV 13 will be necessary to understand other malignant lesions that have been reported by some authors (Rolighed et al., 1992; Castro et al., 2011).

The comparison of E6 and E7 ORFs of HPV-13 isolates from Amazonia and Turkey was done in accordance with the principles outlined by Bernard (2005). These suggest that there are a limited number of common variants of each HPV type, and that these variants show maximal divergence when they are sampled from ethnic groups that have evolved for a long time without contact.

We cannot affirm that this is a variant because the L1 gene was not studied. Our findings indicate that E6 and E7 genes of HPV-13 isolates, included in this comparison, may have evolved more recently and despite the great geographic distance, they present high similarity (GenBank accession no. DQ344807).

ACKNOWLEDGMENTS

The authors are grateful to the Federal University of Amazonas (UFAM), Fundação de Amparo à Pesquisa do Estado do Amazonas (FAPEAM), Conselho Nacional de

Desenvolvimento Científico e Tecnológico (CNPq) and Superintendência da Zona Franca de Manaus (SUFRAMA), for the support received during the development of this research.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Altschul SF, Madden TL, Schäffer AA, Zhang J, et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25: 3389-3402.

Archard HO, Heck JW and Stanley HR. (1965). Focal epithelial hyperplasia: an unusual oral mucosal lesion found in Indian children. *Oral Surg Oral Med Oral Pathol.* 20: 201-212.

Bernard HU (2005). The clinical importance of the nomenclature, evolution and taxonomy of human papillomaviruses. *J. Clin. Virol.* 32: S1-S6.

Borborema-Santos CM, de Castro MM, Santos PJB, Talhari S, et al. (2006). Oral focal epithelial hyperplasia: report of five cases. *Braz. Dent. J.* 17: 79-82.

Carlos R and Sedano HO (1994). Multifocal papilloma virus epithelial hyperplasia. Oral Surg Oral Med Oral Pathol. 77: 631-635.

Castro MM, Borborema-Santos CM, Farias IP, Correia G, et al. (2011). Prevalence of human papillomavirus (HPV) type 16 variants and rare HPV types in the central Amazon region. *Genet. Mol. Res.* 10: 186-196.

Cheah PL and Looi LM (1998). Biology and pathological associations of the human papillomaviruses: a review. *Malaysion J. Pathol.* 20: 1-10.

de Villiers EM, Fauquet C, Broker TR, Bernard HU, et al. (2004). Classification of papillomaviruses. *Virology*. 324: 17-27.

González LV, Gaviria AM, Sanclemente G, Rady P, et al. (2005). Clinical, histopathological and virological findings in patients with focal epithelial hyperplasia from Colombia. *Int. J. Dermatol.* 44: 274-279.

Gross GE and Barrasso R (1999). Infecção por Papilomavírus Humano – Atlas Clínico de HPV; trad. Maria Alice Guimarães. Porto Alegre. Editora Artes Médicas Sul Ltda.

Hopp TP and Woods KR (1983). A computer program for predicting protein antigenic determinants. *Mol. Immunol.* 20: 483-489.

Jarvis A and Gorlin RJ (1972). Focal epithelial hyperplasia in an Eskimo population. *Oral Surg. Oral Med. Oral Pathol.* 33: 227-228.

Pfister H, Hettich I, Runne U, Gissmann L, et al. (1983). Characterization of human papillomavirus type 13 from focal epithelial hyperplasia Heck lesions. *J. Virol.* 47: 363-366.

Rolighed J, Sorensen IM, Jacobsen NO and Lindeberg H (1991). The presence of HPV types 6/11, 13, 16 and 33 in bowenoid papulosis in an HIV-positive male, demonstrated by DNA in situ hybridization. *APMIS.* 99: 583-585.

Sambrook J, Fritsch EF and Maniatis T (1989). Molecular Cloning: A Laboratory Manual. 2nd edn. Cold Springs Harbor Laboratory Press, Cold Springs Harbor, New York.

Santillán EM (2003). Hiperplasia epitelial focal (enfermidade de Heck). Reporte de un caso. *Rev. Mex. Pediatr.* 70: 132-134.

Santos PJB, Borborema-Santos CM, Mendonça RR, Vieira do Carmo MA, et al. (2005). Human papillomavirus type 13 infecting the conjunctiva. *Diagn. Microbiol Infect Dis.* 53: 71-73.

Syrjänen S (2003). Human papillomavirus infections and oral tumors. *Med. Microbiol Immunol.* 192: 123-128.

Terezhalmy GT, Riley CK and Moore WS (2001). Focal epithelial hyperplasia (Heck's disease). Quintessence Int 32: 664-665.

Thompson J, Gibson TJ, Plewniak F, Jeanmougin F, et al. (1998). The CLUSTAL_ X windows interface: flexible strategies for multiple sequence alignment aided by quality tools. *Nucleic Acids Res.* 25: 4876-4882.

Van Ranst M, Fuse A, Fiten P, Beuken E, et al. (1992). Human papillomavirus type 13 and pygmy chimpanzee papillomavirus type 1: comparison of the genome organizations. *Virology.* 190: 587-596.

Ying X, Dong X and Jie L (2010). Computational Methods for Protein Structure Prediction and Modeling: v.2: Structure Prediction. *Protein Sci.* 10: 455-458.

zur Hausen H (2002). Papillomaviruses and cancer: from basic studies to clinical application. *Nat. Rev. Cancer*. 2: 342-350.

