

SHORT REPORT

Open Access



# Genomic organization of a Gamma-6 papillomavirus metagenomic discovered from vaginal swab samples of Chinese pregnant women

Yu Ling<sup>1,2†</sup>, Jianqiang Wang<sup>3†</sup>, Jun Yin<sup>4</sup>, Jianpu Xu<sup>4</sup>, Yufan Wu<sup>4</sup>, Rui Zhou<sup>1</sup>, Juan Lu<sup>1</sup>, Shixing Yang<sup>1</sup>, Xiaochun Wang<sup>1</sup>, Quan Shen<sup>1\*</sup> and Wen Zhang<sup>1\*</sup>

## Abstract

A complete genome sequence of human papillomaviruses (HPV) named as HPV-ujs-21015 was determined by viral metagenomic and PCR methods. The complete genome is 7354 bp in length with GC content of 41.7%, of which the genome was predicted to contain six ORFs (Open Reading Frame, ORF) coding for four early proteins (E7, E1, E4, and E2) and two late proteins (L1 and L2). Phylogenetic analysis based on the complete genome and the L1 protein showed that HPV-ujs-21015 belongs to a type 214 member within genus Gamma-6 papillomavirus. It is the first complete genome of Gamma-6 papillomavirus discovered from pregnant women in China.

**Keywords:** Human papillomaviruses, Gamma-6 papillomavirus, Virus metagenomics, Complete genome, Genomic organization

## Main text

Human papillomavirus (HPVs), a member of the *Papillomaviridae* family, are nonenveloped, double-strand circular DNA viruses with an approximately 8 kb genome in length. In the circular genome of HPV, eight genes are typically encoded. L1 and L2 code capsid proteins of virus, which can help virus entry into the basal layer keratinocytes [1, 2]. E2 protein is required for the transcription of viral genes and replication, and also recruits the viral DNA helicase E1 to keeping viral genomes in host cells [3]. E6 and E7 are believed to drive cellular immortalization and maintain the transformed phenotype during tumor progression, to exert functions by binding with many cellular protein to activate cancer

hallmarks [4]. HPVs are classified into genera (alpha, beta, gamma, mu, and nu), species, types and even variants based on the nucleotide similarity, with the different types having different life-cycle characteristics and disease associations [5, 6]. HPV persistent infection is the main risk factor for the development of many tumors especially cervical tumor [7]. Although there were numerous ways to prevent the infection of HPV, such as vaccination, over 600,000 cases per year of cervical cancer were recorded worldwide [8]. According to the data from International HPV Reference Center at the Karolinska Institute, Stockholm, Sweden, as of May 6, 2016, two hundred and twenty-six reference HPV types, ranging from HPV-1 to HPV-226, were officially recognized ([https://www.hpvcenter.se/human\\_reference\\_clones/](https://www.hpvcenter.se/human_reference_clones/)).

The determination of HPV genome can be helpful to understand the genomic characteristics and the clinical relevance of these new HPV strains. In recent years, in addition to frequently-used methods like PCR, some

\* Correspondence: [shenquan@ujcs.edu.cn](mailto:shenquan@ujcs.edu.cn); [zhangwen@ujcs.edu.cn](mailto:zhangwen@ujcs.edu.cn)

<sup>†</sup>Yu Ling and Jianqiang Wang contributed equally to this work.

<sup>1</sup>School of Medicine, Jiangsu University, 301 Xuefu Road, Zhenjiang, Jiangsu 212013, People's Republic of China

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

new methods including viral metagenomics were used to acquire the genome of HPV more efficiently [9, 10].

In our current study, the viral nucleic acid sequences from vaginal swabs were investigated through viral metagenomics. A total of 100 vaginal swabs were collected from the health pregnant women who visited hospital for antenatal follow-up of pregnancy in Shanghai City, China, in 2017. The total viral nucleic acid was isolated using QiaAmp Mini Viral RNA kit (Qiagen, USA) according to the manufacturer's protocol after centrifugation, filtration and DNase and RNase digestion, as we described previously, and pooled into 10 libraries [9]. The produced nucleic acids (both DNA and RNA) were subjected to reverse transcript with N8 random primers (Sangon, Shanghai, China), and the second strand was generated using Klenow enzyme (NEB, Ipswich, USA). The libraries were then constructed by the Nextera XT DNA sample Preparation Kit (Illumina, CA, USA) following the protocol, and the prepared libraries were sequenced by Illumina Miseq platform with 250 bases paired ends with dual barcoding for each pool.

The total numbers of sequence reads generated for the 10 libraries were 73,264 (swab01), 45,462 (swab02), 100,518 (swab03), 111,398 (swab04), 82,612 (swab05), 436,560 (swab06), 903,618 (swab07), 71,544 (swab08), 273,046 (swab09), and 51,590 (swab10). Raw data were processed according to the standard procedure which included debarcoding, trimming and assembling [11]. Contigs and singlet reads were then matched against a customized viral proteome database using BLASTx with an E value cutoff of  $< 10^{-5}$ . Bioinformatics analysis was performed according to a previous study [9]. PCR and sanger sequencing were carried out to bridge the gaps between sequences as well as assess the prevalence of HPV strain identified in this study. Putative ORFs (Open Reading Frame, ORF) in the genome of HPV-uj-21015 were predicted by Geneious Prime software (version 2020.0.4). The closest viral strains based on best BLASTx hits and the representative members of species and genera were selected to perform the phylogenetic analyses (Table 1). In order to construct the phylogenetic tree, sequence alignment was performed using Clustal W with the default settings. Phylogenetic tree was generated using the maximum likelihood method based on Jones-Taylor-Thornton (JTT) model by MEGA 7.0 with 1000 bootstrap. Bootstrap values for each node are given in the trees.

## Results and discussion

A strain of HPV named as HPV-uj-21015 (GenBank accession no. MN400665, see Additional file 1) was determined in the vaginal swab (containing 1654 reads in library swab02), of which the complete genome is 7354 bp in length with GC content of 41.7%. The genome of

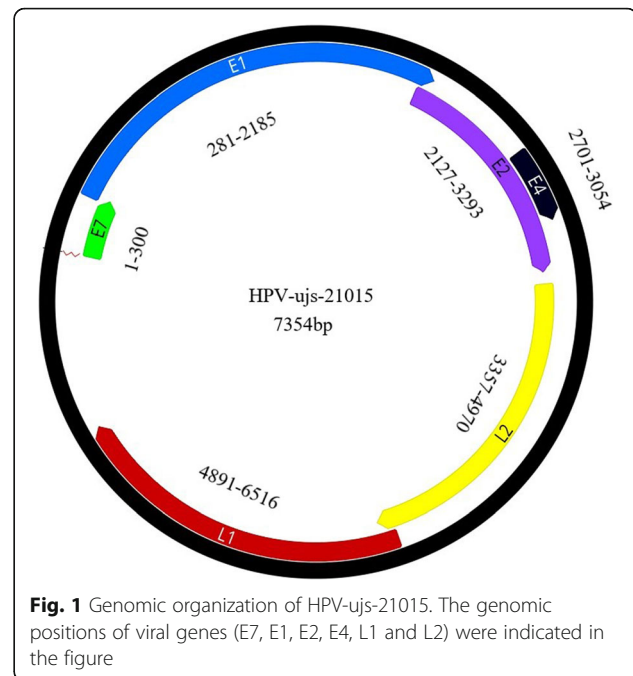
**Table 1** The reference HPV strains and their genera and species. Classification was based on International Committee on Taxonomy of Viruses ([https://talk.ictvonline.org/ictv-reports/ictv\\_online\\_report/dsdna-viruses/w/papillomaviridae](https://talk.ictvonline.org/ictv-reports/ictv_online_report/dsdna-viruses/w/papillomaviridae)), International HPV Reference Center at the Karolinska Institute, Stockholm, Sweden (<http://www.hpvcntr.se>), and Bernard et al., *Virology*. 2010 May 25; 401(1): 70–79

| genus               | species      | abbreviation | GenBank accession number |
|---------------------|--------------|--------------|--------------------------|
| alphapapillomavirus | alpha-1      | HPV32        | NC_001586                |
|                     | alpha-2      | HPV3         | X74462                   |
|                     | alpha-3      | HPV61        | NC_001694                |
|                     | alpha-4      | HPV2         | NC_001352                |
|                     | alpha-5      | HPV26        | NC_001583                |
|                     | alpha-6      | HPV30        | NC_038889                |
|                     | alpha-7      | HPV18        | NC_001357                |
|                     | alpha-7      | HPV85        | AF131950                 |
|                     | alpha-8      | HPV7         | NC_001595                |
|                     | alpha-9      | HPV16        | NC_001526                |
|                     | alpha-9      | HPV58        | D90400                   |
|                     | alpha-10     | HPV6         | NC_001355                |
|                     | alpha-11     | HPV34        | NC_001587                |
|                     | alpha-13     | HPV54        | NC_001676                |
| alpha-14            | HPV71        | NC_039089    |                          |
| betapapillomavirus  | beta-1       | HPV5         | NC_001531                |
|                     | beta-2       | HPV9         | NC_001596                |
|                     | beta-2       | HPV23        | U31781                   |
|                     | beta-3       | HPV49        | NC_001591                |
|                     | beta-4       | HPV92        | NC_004500                |
|                     | beta-5       | HPV96        | NC_005134                |
|                     | unclassified | isol Ki88    | EU410347ACC78262         |
| gammapapillomavirus | gamma-1      | HPV4         | NC_001457                |
|                     | gamma-1      | HPV95        | AJ620210                 |
|                     | gamma-2      | HPV48        | NC_001690                |
|                     | gamma-3      | HPV50        | NC_001691                |
|                     | gamma-4      | HPV60        | NC_001693                |
|                     | gamma-5      | HPV88        | NC_010329                |
|                     | gamma-6      | HPV101       | NC_008189                |
|                     | gamma-6      | mw03c65      | MF588697                 |
|                     | gamma-7      | HPV109       | NC_012485                |
|                     | gamma-8      | HPV112       | NC_012486                |
|                     | gamma-9      | HPV116       | NC_013035                |
|                     | gamma-10     | HPV121       | NC_014185                |
| gamma-11            | HPV126       | NC_016157    |                          |
| gamma-12            | HPV127       | NC_014469    |                          |
| gamma-12            | HPV132       | GU117632     |                          |
| gamma-12            | HPV148       | GU129016     |                          |
| gamma-12            | HPV199       | KJ913662     |                          |

**Table 1** The reference HPV strains and their genera and species. Classification was based on International Committee on Taxonomy of Viruses ([https://talk.ictvonline.org/ictv-reports/ictv\\_online\\_report/dsdna-viruses/w/papillomaviridae](https://talk.ictvonline.org/ictv-reports/ictv_online_report/dsdna-viruses/w/papillomaviridae)), International HPV Reference Center at the Karolinska Institute, Stockholm, Sweden (<http://www.hpvcenter.se>), and Bernard et al., *Virology*. 2010 May 25; 401(1): 70–79 (Continued)

| genus            | species      | abbreviation | GenBank accession number |
|------------------|--------------|--------------|--------------------------|
|                  | gamma-13     | HPV128       | NC_014952                |
|                  | gamma-14     | HPV131       | NC_014954                |
|                  | gamma-15     | HPV135       | NC_017993                |
|                  | gamma-16     | HPV137       | NC_017995                |
|                  | gamma-17     | HPV144       | NC_017997                |
|                  | gamma-18     | HPV156       | NC_033781                |
|                  | gamma-19     | HPV161       | NC_038522                |
|                  | gamma-19     | HPV162       | JX413108                 |
|                  | gamma-19     | HPV166       | JX413104                 |
|                  | gamma-20     | HPV163       | NC_028125                |
|                  | gamma-21     | HPV167       | NC_022892                |
|                  | gamma-22     | HPV172       | NC_038523                |
|                  | gamma-23     | HPV175       | NC_038524                |
|                  | gamma-24     | HPV178       | NC_023891                |
|                  | gamma-24     | HPV197       | KM085343                 |
|                  | gamma-25     | HPV184       | NC_038914                |
|                  | gamma-27     | HPV201       | NC_027528                |
|                  | unclassified | HPV-ZJ01     | KX082661                 |
|                  | unclassified | isol CH2     | KF791917                 |
|                  | unclassified | isol Fi864   | KC311731                 |
|                  | unclassified | isol KC5     | JX444073                 |
| mupapillomavirus | mu-1         | HPV1         | NC_001356                |
|                  | mu-2         | HPV63        | NC_001458                |
|                  | mu-3         | HPV204       | NC_038525                |
| nupapillomavirus | nu-1         | HPV41        | NC_001354                |

HPV-ujs-21015 was predicted to contain six ORFs coding for four early proteins (E7, E1, E4, and E2) and two late proteins (L1 and L2) (Fig. 1). The nucleic acid lengths of these proteins were 300, 1905, 354, 1167, 1626 and 1614, respectively, and the positions on the genome were showed in Fig. 1. Notably, the E6 gene that plays a crucial role in the cell transformation through binding of p53 tumor suppressor protein was absent in this strain, which was consistence with other HPV214 strains [12, 13]. E6 as well as E7 is believed to be directly responsible for the development of HPV-induced carcinogenesis. In the high risk HPVs, they do this cooperatively by targeting diverse cellular pathways including the regulation of cell cycle control. Meanwhile, there is a view that the lost function of E6 in HPV214 may be

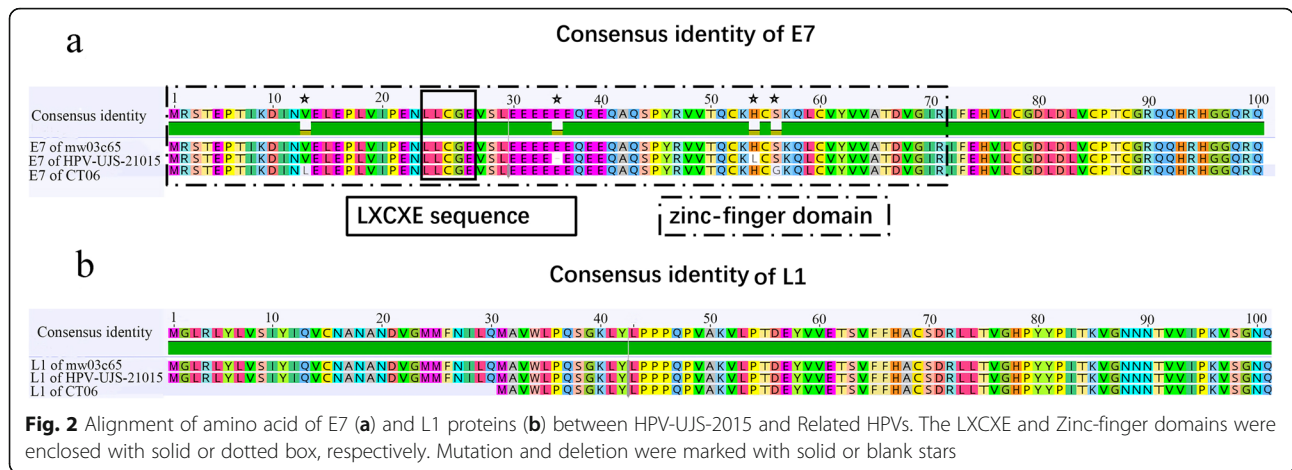


**Fig. 1** Genomic organization of HPV-ujs-21015. The genomic positions of viral genes (E7, E1, E2, E4, L1 and L2) were indicated in the figure

compensated for in its E7 protein which has an LXCXE (Fig. 2a) motif that has been shown to bind pRB in HPV16 and other high risk HPV types.

According to the International Committee on Taxonomy of Viruses (ICTV), a viral type within a species has 71 to 89% identity with other types within the same species based on the comparative homology of the L1 DNA sequence. Additionally, there are several subtypes and variants within a type, which share 90 to 98% and more than 98% identity, respectively. In the current study, sequence analysis indicated that HPV-ujs-21015 shared the highest nucleotide (nt) sequence identity (99%) with a type 214 strain named CT06 isolated from South African strain (GenBank no. MF509819), as well as strain mw03c65 (GenBank accession no. MF588697), which was an unclassified strain detected in patients with immunodeficiency in USA.

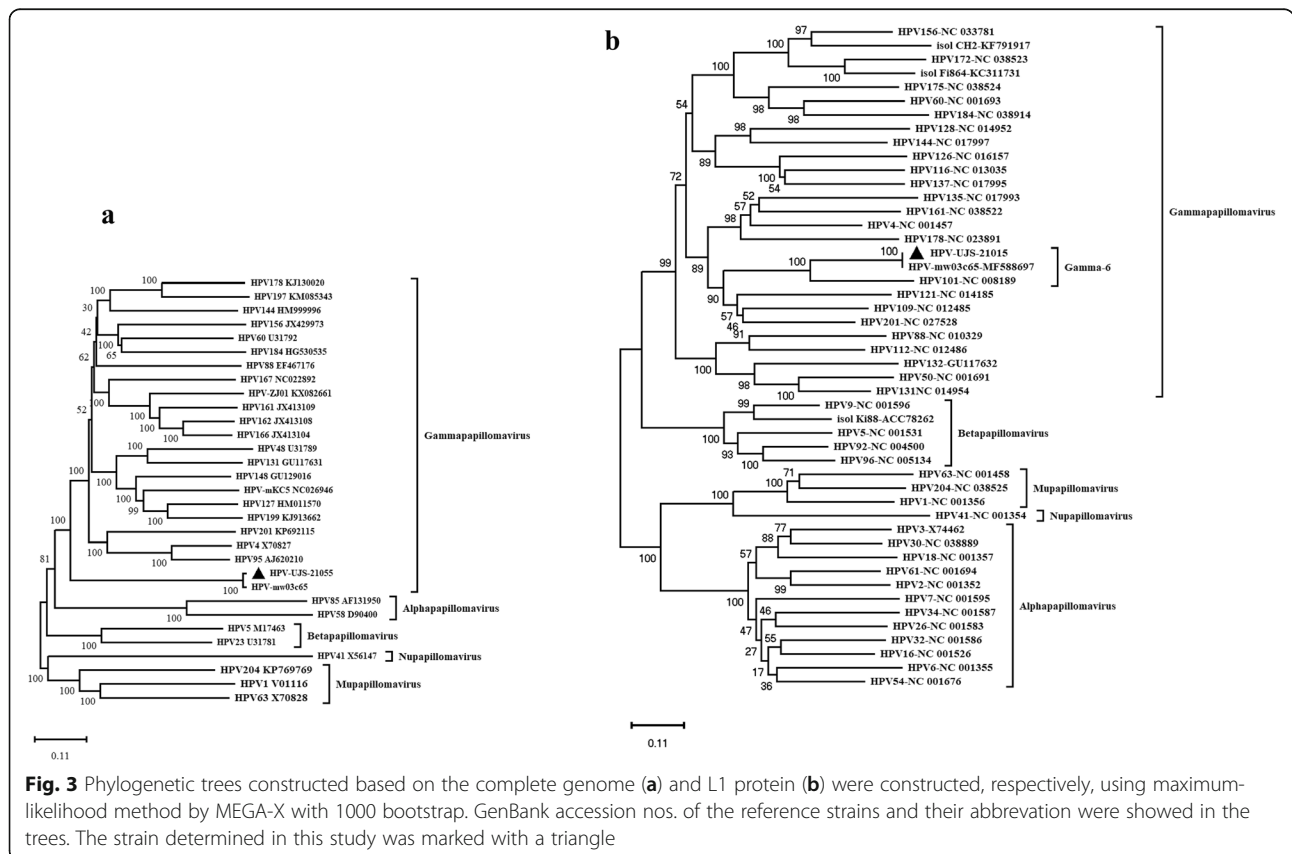
Similar to mw03c65 and CT06 strain, the putative E7 protein of HPV-ujs-21015 strain contained one zinc-finger domain and an LXCXE sequence (Fig. 2a), which is critical for transforming activities by way of binding a number of important cellular regulatory proteins, including tumor suppressor: Retinoblastoma protein (pRb). Compared with these two strains, HPV-ujs-21015 had one amino acid deletion and three mutations (Fig. 2a). Whether the deletion and mutations affect the biological function of E7 will require more research. Intriguingly, another protein with significant diversity was L1, of which HPV-ujs-21015 had the 100% amino acid similarity with mw03c65, but was thirty consensus amino acid longer than that of CT06 strain in the 5'end (Fig. 2b).



To characterize the phylogenetic relationship between HPV-ujs-21015 and related HPV reference strains, two phylogenetic trees based on the complete genome and L1 protein were constructed, respectively, by MEGA 7.0. Both trees revealed that the reference HPVs were clustered well in their genera and types. The phylogenetic tree based on the complete genome showed that HPV-ujs-21015 belonged to Gamma-papillomavirus (Fig. 3a). The other phylogenetic tree based on the L1 protein further assigned HPV-ujs-21015 within the group of type

214 in Gamma-6, being closely related to mw03c65 strain (Fig. 3b). In summary, our results suggest that all of these three strains isolated from different countries were variants with the genotype 214.

HPVs comprise five evolutionary groups with different epithelial tropisms and disease associations. Traditionally, based on the location of the certain virus genome was found, HPVs have also been classified as mucosal or cutaneous types [1]. Increasing evidences revealed that Gamma-PVs showed broad tissue tropism, with the



detection locations ranging from health skin and cutaneous lesions to genital lesions [10, 14, 15]. DNA of some Gamma-PVs types were detected in skin cancer raised concerns of some Gamma-PVs associations with cancers, especially in patients with immunodeficiency or immunosuppression [16, 17]. In the current study, HPV-ujs-21015 strain was identified from a health pregnant woman who visited hospital for antenatal follow-up. Vaginitis or other vaginal disease were not found by the attending gynecologist. Generally, both mucosal or cutaneous disease relied on the persistent infection of HPVs. Therefore, whether the infection of HPV-ujs-21015 can cause disease or not is still unknown. A total of one hundred of vaginal swab samples from health pregnant women who visited hospital for antenatal follow-up were screened by PCR method with a set of nested primers (data not showed) designed on HPV-ujs-21015 L1 gene. Result showed that two samples were positive (2/100). The prevalence and disease association of HPV-ujs-21015 need to be clarified through larger sample size, biological and histological experiments.

In conclusion, we determined and characterized the complete genome sequence of a genotype 214 Gamma-6 papillomavirus, which was isolated from a health pregnant woman of China. To the best of our knowledge, it is the first complete genome of Gamma-6 papillomavirus detected in Pregnant Women of China.

### Supplementary information

**Supplementary information** accompanies this paper at <https://doi.org/10.1186/s12985-020-01319-9>.

**Additional file 1.** The complete genome of HPV-ujs-21015.

### Abbreviations

aa: Amino acid; HPV: Human papillomaviruses; ORF: Open Reading Frame; ICTV: International Committee on Taxonomy of Viruses; nt: Nucleotide; pRb: Retinoblastoma protein

### Acknowledgements

Not applicable.

### Authors' contributions

QS and WZ conceived the study and designed the experiments. YL, JQW, JY, JPX and YFW collected the samples, YL and RZ and JL performed the laboratory assays. SXY and XCW contributed to analysis of sequencing data. LY wrote the initial draft, and QS edited the manuscript. All authors read and approved the final manuscript.

### Funding

This study was financially supported by Changzhou Social Development Project (Grant No. CE20185002); Jiangsu social development project (No. BE2016663), Zhangjiagang science and technology planned project (ZKS1703), Research Funds of Jiangsu Entry-Exit Inspection and Quarantine Bureau (No. 2018KJ07).

### Availability of data and materials

The sequences of full-length envelope gene generated in this study have been deposited in GenBank under the accession numbers MN400665.

### Ethics approval and consent to participate

This study did not include experiments with human participants or animals performed by any of the authors.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

### Author details

<sup>1</sup>School of Medicine, Jiangsu University, 301 Xuefu Road, Zhenjiang, Jiangsu 212013, People's Republic of China. <sup>2</sup>Department of Laboratory Medicine, Nanjing Maternity and Child Health Care Hospital, Nanjing, Jiangsu 210004, People's Republic of China. <sup>3</sup>Intensive Care Unit, Jintan District Hospital of Traditional Chinese Medicine, 1006 Huixian Road, Changzhou, Jiangsu 213200, People's Republic of China. <sup>4</sup>Zhangjiagang Customs, 28 Renmin Road, Zhangjiagang, Jiangsu 215699, People's Republic of China.

Received: 19 October 2019 Accepted: 19 March 2020

Published online: 31 March 2020

### References

1. Doorbar J, Quint W, Banks L, Bravo IG, Stoler M, Broker TR, Stanley MA. The biology and life-cycle of human papillomaviruses. *Vaccine*. 2012;30(Suppl 5):F55–70.
2. Finnen RL, Erickson KD, Chen XJS, Garcea RL. Interactions between papillomavirus L1 and L2 capsid proteins. *J Virol*. 2003;77(8):4818–26.
3. Thomas Y, Androphy EJ. Acetylation of E2 by P300 Mediates Topoisomerase Entry at the Papillomavirus Replicon. *J Virol*. 2019;93(7):e02224–18.
4. Hoppe-Seyler K, Bossler F, Braun JA, Herrmann AL, Hoppe-Seyler F. The HPV E6/E7 oncogenes: key factors for viral carcinogenesis and therapeutic targets. *Trends Microbiol*. 2018;26(2):158–68.
5. Bernard HU, Burk RD, Chen ZG, van Doorslaer K, zur Hausen H, de Villiers EM. Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology*. 2010;401(1):70–9.
6. Bosch FX, Burchell AN, Schiffman M, Giuliano AR, de Sanjose S, Bruni L, Tortolero-Luna G, Kjaer SK, Munoz N. Epidemiology and natural history of human papillomavirus infections and type-specific implications in cervical Neoplasia. *Vaccine*. 2008;26:K1–K16.
7. Schwarz TF. AS04-adjuvanted human papillomavirus-16/18 vaccination: recent advances in cervical cancer prevention. *Expert Rev Vaccines*. 2008;7(10):1465–73.
8. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer*. 2010;127(12):2893–917.
9. Liu ZJ, Yang SX, Wang Y, Shen Q, Yang Y, Deng XT, Zhang W, Delwart E. Identification of a novel human papillomavirus by metagenomic analysis of vaginal swab samples from pregnant women. *Viol J*. 2016;13:122.
10. Pastrana DV, Peretti A, Welch NL, Borgogna C, Olivero C, Badolato R, Notarangelo LD, Gariglio M, FitzGerald PC, McIntosh CE, et al. Metagenomic Discovery of 83 New Human Papillomavirus Types in Patients with Immunodeficiency. *Mosphere*. 2018;3(6).
11. Deng X, Naccache SN, Ng T, Federman S, Li L, Chiu CY, Delwart EL. An ensemble strategy that significantly improves de novo assembly of microbial genomes from metagenomic next-generation sequencing data. *Nucleic Acids Res*. 2015;43(7):e46.
12. Murahwa AT, Meiring TL, Mbulawa ZZA, Williamson AL. Discovery, characterisation and genomic variation of six novel Gammapapillomavirus types from penile swabs in South Africa. *Papillomavirus Res*. 2019;7:102–11.
13. Nobre RJ, Herraes-Hernandez E, Fei JW, Langbein L, Kaden S, Grone HJ, de Villiers EM. E7 oncoprotein of novel human papillomavirus type 108 lacking the E6 gene induces dysplasia in organotypic keratinocyte cultures. *J Virol*. 2009;83(7):2907–16.
14. Bolatti EM, Hosnjak L, Chouhy D, Re-Louhau MF, Casal PE, Bottai H, Kocjan BJ, Stella EJ, Gorosito MD, Sanchez A, et al. High prevalence of Gammapapillomaviruses (gamma-PVs) in pre-malignant cutaneous lesions of immunocompetent individuals using a new broad-spectrum primer system, and identification of HPV210, a novel gamma-PV type. *Virology*. 2018;525:182–91.

15. Bolatti EM, Chouhy D, Casal PE, Perez GR, Stella EJ, Sanchez A, Gorosito M, Bussy RF, Giri AA. Characterization of novel human papillomavirus types 157, 158 and 205 from healthy skin and recombination analysis in genus gamma-papillomavirus. *Infect Genet Evol.* 2016;42:20–9.
16. Botalico D, Chen ZG, Dunne A, Ostoloza J, McKinney S, Sun C, Schlecht NF, Fatahzadeh M, Herrero R, Schiffman M, et al. The Oral cavity contains abundant known and novel human papillomaviruses from the Betapapillomavirus and Gammapapillomavirus genera. *J Infect Dis.* 2011; 204(5):787–92.
17. Dutta S, Robitaille A, Aubin F, Fouere S, Galicier L, Boutboul D, Luzi F, Di Bonito P, Tommasino M, Gheit T. Identification and characterization of two novel Gammapapillomavirus genomes in skin of an immunosuppressed Epidermodysplasia Verruciformis patient. *Virus Res.* 2018;249:66–8.

### **Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

