

Prevalence and genotype distribution of human papillomavirus among Hakka women in China

Xiang-Xing Zeng¹, Li-Xiang Yan¹, Xiu-Xia Huang², Cai-Hua He¹, Wei-Guo Liu³, Wen-Qing Yuan⁴, Yan-Ping Qiu⁴, Zhi-Xiang Liu²

¹Department of Clinical Laboratory, ²Department of Women's Health, ³Department of Gynecology, ⁴Department of Surgery, Heyuan Women and Children's Hospital, Heyuan 517000, China

Contributions: (I) Conception and design: XX Zeng; (II) Administrative support: XX Huang, ZX Liu; (III) Provision of study materials or patients: All authors; (IV) Collection and assembly of data: All authors; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Zhi-Xiang Liu. Department of Women's Health, Heyuan Women and Children's Hospital, Heyuan 517000, China. Email: 468766614@qq.com.

Background: Human papillomavirus (HPV) infection is the primary risk factor for cervical cancer. HPV genotypes are associated with varying degrees of pathogenicity. To better formulate strategies for cervical cancer prevention, we investigated the population-specific distribution of HPV genotypes, including those with high carcinogenicity.

Methods: From January to December 2012, a cervical cancer-screening program for HPV infection in Hakka women of Heyuan City Guangdong province was conducted. Of 736,000 women residents, 8,284 volunteers were recruited. The cytology specimens of 107 women were not adequate and excluded. Thus, 8,177 women submitted to polymerase chain reaction (PCR) sequencing of 16 HPV genotypes via MassARRAY spectrometry.

Results: Risk stratification based on genotypes indicated that the prevalence of overall, high-risk, and low-risk HPV infections was 12.27%, 14.20%, and 0.79%, respectively. Of the 1,003 women positively infected, 82.75% were infected with a single HPV type; 17.25% were infected with ≥ 2 types. Analysis revealed a U-shaped curve in HPV prevalence that correlated with age group, with peaks at ages 18–24 y (22.03%) and 60–65 y (25%). The most frequently detected HPV genotype was HPV-52 (26.81%), and then HPV-16 (17.54%), HPV-58 (14.25%), HPV-18 (10.16%), HPV-68 (8.27%), HPV-39 (5.68%), and HPV-51 (5.38%).

Conclusions: HPV-52 is the most prevalent genotype infecting Hakka women. Therefore, vaccination against HPV-52 is imperative. The prevalence of HPV infection is highest in the younger (18–24 y) and older (60–65 y) age groups, indicating that screening for HPV in Hakka women should be performed early and maintained in the elderly.

Keywords: Human papillomavirus (HPV); cervical cancer; genotypes; Hakka

Submitted Nov 11, 2015. Accepted for publication May 15, 2016.

doi: 10.21037/atm.2016.05.49

View this article at: <http://dx.doi.org/10.21037/atm.2016.05.49>

Introduction

Cervical cancer is the second most common cancer in women worldwide. In the year 2008, ~530,000 new cases of cervical cancer were diagnosed, and 275,000 women died of the disease (1). It is now widely accepted that cervical cancers highly correlate with infection by the human papillomavirus (HPV).

More than 150 genotypes of HPV have been identified by DNA sequencing, and can be stratified according to oncogenic risk. Fifteen genotypes are clearly linked to the development of adenocarcinoma and squamous cell carcinoma of the cervix and therefore rank as high-risk: HPV-16, -18, -31, -33, -35, -39, -45, -51, -52, -56, -58, -59, -68, -73, and -82. Twelve genotypes are either not

potentially oncogenic, or not involved in the development of cervical cancer, and are considered low-risk: HPV-6, -11, -40, -42, -43, -44, -54, -61, -70, -72, -81, and CP6108 (1,2)

Increasing evidence suggests that multiple factors contribute to the development of cervical cancer, including genetic susceptibility or host genome, HPV variants, co-infection of HPV and other agents, and life-style factors (3). Although all genotypes of HPV correlate with cervical cancers, HPV-16 and HPV-18 infection accounts for about 70% of all cases.

The prevalence of HPV infection has been studied extensively, and epidemiologic characteristics of the HPV genotypes vary by geographic region (4). Forman *et al.* (5) reported that, worldwide, the prevalence of HPV infection without cervical abnormalities is 11% to 12%. de Sanjosé *et al.* (4) found that, overall, in Asia and China the prevalence of HPV in females without cytological abnormalities was 8.0% and 11.4–20.3%, respectively. However, the prevalence of HPV infection is 10% in Guangzhou (6), 6.2% in Hong Kong (6), 24.5% in Chaozhou (7), 18.4% in Shenzhen (8), 12.6% in Shanghai (9), 13.3% in Zhejiang (10), 13.6% in Shenyang (11), and 29.9% in Macau (12).

The Hakka people are a distinct Han Chinese population who speak Hakka Chinese and live in several provincial areas of China. About 60% of Hakka people live in Guangdong province, particularly in the northeast. The city of Heyuan in Guangdong province is the only city on the mainland in which only Hakka people reside. Over centuries, the Hakka have formed distinctive physiological characteristics. Based on genetic studies, the Hakkas are thought to originate from the Han people of northern China, and are distinctive in their unique tradition and culture. The physical characteristics of Hakka are different from that of populations in northern and southern China (13). However, the prevalence and genotype distribution of HPV in Hakka women so far has not been reported.

To estimate the efficacy of a future HPV vaccination program, in the present study we selected Hakka women residents of the city of Heyuan in Guangdong province to investigate the prevalence and distribution of their HPV genotypes.

Methods

Study area and population

All participants were identified through the Cervical Cancer Screening Program of Heyuan and recruited from the

Heyuan Women and Children's Hospital, from January 2012 to December 2012. The program is a comprehensive public health program that helps the women residing in Heyuan gain access to cervical cancer screening services. Since the inception, 8,284 of the 736,000 registered Hakka women in Heyuan participated in the screening program. Participants were divided into nine sub-groups based on their age (18–24, 25–29, 30–34, 35–39, 40–44, 45–49, 50–54, 55–59, and 60–65 years).

The specimens of the enrollees and corresponding clinical information were obtained under protocols approved by the Ethics Committee of the hospital. All patients provided written informed consent. Participants were excluded for the following: history of total hysterectomy, acute genital inflammation, virgin status, sexual intercourse in the previous 3 days, or pregnancy. After exclusions, the genomic DNA for HPV detection and genotyping was successfully extracted from the remaining 8,177 women.

Questionnaire interview

The participants were interviewed by a trained interviewer with a standardized questionnaire, which was designed to record general demographic data, physical and mental status, marital status (yes or no), menstrual status, and childbearing history. The interview was conducted face-to-face using the local language.

Cervical specimen collection

All cervical specimens were collected by trained physicians. The physicians first removed secretions with cotton swabs and then collected exfoliated cervical cells with a cervical brush [Beijing Genomics Institute (BGI), Shenzhen, China]. The brush was inserted 1 to 1.5 cm into the endocervical canal and rotated 3 to 5 full turns. Each brush, containing cellular material, was mixed well with 1.5 mL of specimen transport medium and stored at 4 °C immediately. All specimens were labeled with a coded number.

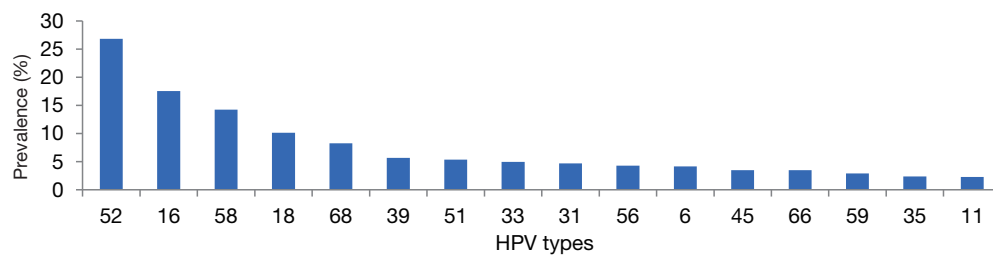
HPV detection and genotyping

HPV detection and genotyping were performed using a MassARRAY platform (Sequenom, San Diego, CA, USA), i.e., matrix-assisted laser desorption/ionization time-of flight mass spectrometer, as previously described (14,15). HPV DNA samples were extracted from the cervical specimens using a commercial magnetic beads kit (Chemagen,

Table 1 Distribution of multiple HPV infections in HPV-positive subjects by numbers of cases

Risk categories of HPV types	Number of HPV genotypes per subject				Total
	Two	Three	Four	Five	
High-risk and low-risk	13	3	0	1	17
Low-risk only	1	0	0	0	1
High-risk only	128	24	3	0	155
Total number of cases	142	27	3	1	173

HPV, human papillomavirus.

**Figure 1** Distribution of HPV types among 1,003 HPV-positive Hakka women. HPV, human papillomavirus.

Pekinemer, Waltham, MA, USA) in accordance with the manufacturer's instructions, and amplified through polymerase chain reaction (PCR) technology. Sixteen HPV genotypes were successfully sequenced and identified, including 14 high-risk (HPV-16, -18, -31, -33, -35, -39, -45, -51, -52, -56, -58, -59, -66, and -68) and 2 low-risk (HPV-6, and -11) types. All DNA extraction, PCR amplification, HPV recombination plasmids, mass spectrum detection, and data analysis were performed by BGI in Shenzhen.

Statistical analyses

All statistical analyses were performed using SPSS 17.0 (IBM, Armonk, NY, USA) statistical package for Windows.

Results

Prevalence of specific HPV genotypes

The average age of the participants was 41.34 years (range, 18–65 years). Among the 8,177 eligible participants, the genotype screen revealed that 1,003 women were HPV-positive. The overall prevalence of HPV was 12.27% (1,003/8,177).

Among the 1,003 HPV-positive women, 82.75%

(830/1,003) were infected with only one genotype of HPV, while 17.25% (173/1,003) were infected with two or more (Table 1). Of those with more than one, 82.08% (142/173) were infected with two genotypes, 15.61% (27/173) with three, 1.73% (3/173) with four, and only one subject with five. Patients with multiple HPV genotype infections were categorized as all high-risk, all low-risk, or a combination of high-risk and low-risk.

HPV-52 was the most frequently detected genotype, accounting for 26.81% (269/1,003) of the positive samples, followed by HPV-16 (17.54%), HPV-58 (14.25%), HPV-18 (10.16%), HPV-68 (8.27%), HPV-39 (5.68%), and HPV-51 (5.38%). Other HPV genotypes (-33, -31, -56, -6, -45, -66, -59, -35, and -11) were observed in fewer than 5% of the positive samples (Figure 1).

Age-specific prevalence of overall HPV infection

Interestingly, the age-specific prevalence pattern of overall HPV infections appeared as a U-shaped curve (Table 2). Participants older than 60 years had the highest HPV prevalence (25%). Participants younger than 24 years had the second higher prevalence (22.3%). The prevalence of HPV of these two groups was markedly higher than that of

Table 2 Age-specific rate of prevalence of positive HPV infection among the 8,177 women

Age, y	N	Overall, n (%)	Single, n (%)	Multiple, n (%)	High-risk, n (%)	Low-risk, n (%)
18–24	118	26 (22.03) ^a	20 (16.95)	6 (5.08)	31 (26.27) ^o	3 (2.54) ^c
25–29	886	93 (10.50)	71 (8.01)	22 (2.48)	106 (11.96)	13 (1.47)
30–34	1,681	178 (10.59)	146 (8.69)	32 (1.90)	200 (11.90)	14 (0.83)
35–39	1,629	200 (12.28)	173 (10.62)	27 (1.66)	233 (14.30)	7 (0.43)
40–44	1,604	189 (11.78)	159 (9.91)	30 (1.87)	214 (13.34)	11 (0.69)
45–49	1,171	139 (11.87)	120 (10.25)	19 (1.62)	159 (13.58)	5 (0.43)
50–54	697	102 (14.63)	84 (12.05)	18 (2.58)	121 (17.36)	7 (1.00)
55–59	351	66 (18.80)	50 (14.25)	16 (4.56)	82 (23.36)	4 (1.14)
60–65	40	10 (25.00) ^a	7 (17.50)	3 (7.50)	12 (30.00) ^o	1 (2.50) ^c
Total	8,177	1,003 (12.27)	830 (82.75)	173 (17.25)	1,158 (14.16)	65 (0.79)

^a, HPV prevalence of the 18–24 y and 60–65 y groups were similar, but for each the HPV prevalence was significantly higher than that of the middle-aged groups ($P < 0.05$, all); ^b, HPV prevalence of the 18–24 y and 60–65 y groups were similar, but for group the HPV prevalence was significantly higher than that of the middle-aged groups ($P < 0.05$, all); ^c, HPV prevalence of the 18–24 y and 60–65 y groups was not significantly different from that of any other age group ($P > 0.05$). HPV, human papillomavirus.

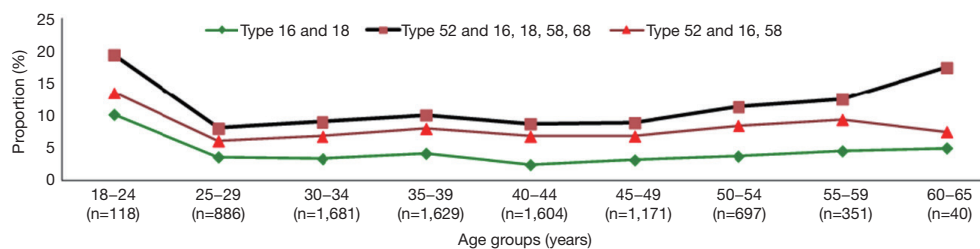


Figure 2 Age-specific prevalence of combined HPV genotypes HPV-16 + HPV-18; HPV-52 + HPV-16 + HPV-18 + HPV-58 + HPV-68; and HPV-52 + HPV-16 + HPV-58. HPV, human papillomavirus.

the middle-aged groups (ages 25 to 59 years).

Similar to the overall distribution pattern of HPV prevalence, a 2-peak (bimodal) pattern of high-risk HPV prevalence was also observed, with the first peak at ages 18–24 years (26.27%) and the second peak at ages 60–65 years (23.36%; *Table 2*). The prevalence of both of these peaks was significantly higher than that of the other age groups.

Age-specific prevalence of specific HPV types

In total, we identified 16 genotypes of HPV. HPV-52 was the most common type identified in 6 of the 9 age groups, that is, all except the age groups 18–24, 35–39, and 60–65 years (*Figure 2*). Infections of combined HPV-52, HPV-16, and HPV-58 constituted 7.19% of all the cases, with no obvious association to any age group. Combined infection of HPV-16

and HPV-18 accounted for 3.40% of all cases, and no obvious association with an age group was observed. The combined HPV genotypes -52, -16, and -18, and HPV genotypes -58 and -68, were similarly prevalent in parallel as the most prevalent genotypes in the age 18–24 y group (19.48%) and in the age 60–65 y group (17.5%). Whereas HPV-16 was the most prevalent genotype in those aged 35–39 years.

HPV-52, HPV-16, and HPV-18 were the three most prevalent types in subjects aged 18–34 years (*Figure 3*). HPV-52, HPV-16, HPV-58 were the three most prevalent types in the five age groups encompassing ages 35–59 years. Some genotypes of HPV were only detected in certain age groups.

Discussion

To the best of our knowledge, this is the first large-scale

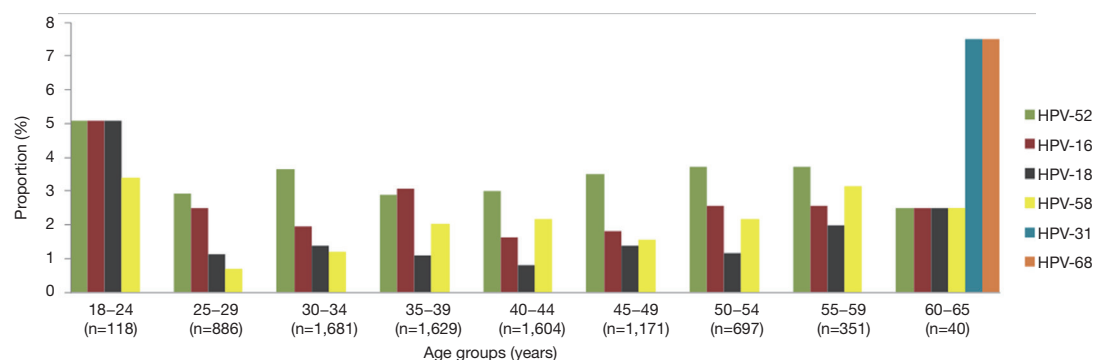


Figure 3 The most common HR-HPV genotypes in each age group. HPV, human papillomavirus.

study to investigate the genotypes of HPV infection in Hakka people in Heyuan. We found that the overall prevalence of HPV was 12.27%, which differs from that of other cities of Guangdong province (6-8). This may be because, firstly, samples in the present study were collected from the general population recruited from the urban or surrounding area of Heyuan, while other investigators collected samples from clinic-based populations. Secondly, the previous studies used PCR and the GenoArray test for the detection and genotyping of HPV, while we used MassARRAY (Sequenom, San Diego, CA, USA). In addition, in our study 16 HPV genotypes were identified, including 14 high-risk and 2 low-risk genotypes. However, several other genotypes of HPV have been detected in previous studies (16).

A previous meta-analysis showed that, worldwide, HPV prevalence was highest in the younger age categories (<34 years), with a second peak in the older age categories (≥ 45 years) (4). We observed a similar pattern, with peaks of prevalence in subjects aged 18–24 or 60–65 years. Liu *et al.* (6) also reported that in Hong Kong the prevalence of HPV infection showed peaks at ages 20–29 years and those older than 60 years. Although the mechanism of the association between HPV infection and age is not clear, we speculate there are a few possible reasons. First, this may be partly due to the lower hormone levels and weaker immune function in premenopausal and post-menopausal women, who therefore have a lower ability to withstand HPV infection. Second, other unknown infection(s) may increase the chance of HPV infection. These multiple infections may affect a woman's susceptibility to HPV, or its persistence (17). Lee *et al.* showed that changes in vaginal microbiota in post-menopausal women made them more susceptible to HPV infection (18). Other research found

that bacterial vaginosis was associated with susceptibility to HPV infection (19). Moreover, the social activities of the younger age groups can lower immunity to HPV and increase chance of exposure.

It has been reported that, worldwide, HPV-16 is the genotype with the highest prevalence, followed by HPV-18 and HPV-31. Therefore, most HPV vaccines are against both HPV-16 and HPV-18. However, in the present study we found that HPV-52 was the most prevalent HPV genotype in Hakka, which is consistent with the results reported in Taiwan and Japan (4), as well as the neighboring cities of Guangzhou (6), Chaozhou (7), and Macau (12). However, the genotype with highest prevalence in Hong Kong and Shen Zhen is HPV-16 (6,8).

Furthermore, in the present study HPV-52 was the most common HPV genotype in 6 of the 9 age groups, and the overall prevalence of HPV-52 was 3.29% (269/8177). This is much higher than that of North America, Africa, and Eastern Asia, with prevalence of HPV-52 of 2.1%, 2.4%, and 0.7%, respectively (4). Our study indicates that developing vaccines for HPV-52 is essential, especially in China.

In this study, multiple HPV infections were identified in 17.25% (173/1003) of the positive specimens, and the age-specific prevalence of multiple HPV infections also showed peaks at ages 18–24 and 60–65 years. We speculate that the decline of immunosurveillance in the younger and older age groups may increase the risk of multiple HPV infections. Piris *et al.* (20) also showed that multiple HPV infections were more likely to induce high-grade histological lesions in adolescent girls and women younger than 24 years. Also to be noted, HPV has the ability to avoid host defenses. According to previous studies, people infected with >2 HPV genotypes might have an increased risk of

developing cervical cancer (17). However, the interaction of various genotypes in co-infections remains unclear, and future studies are needed to verify whether coordinated mechanisms in co-infections exist.

Although we achieved a few novel findings in the present study, there remain other relevant factors that should be considered. Firstly, we collected the samples from the general female population, and not specifically cervical cancer patients. Therefore the findings and conclusions drawn from this study may not be applicable for estimating HPV genotype-specific prevalence in woman affected by cervical cancer. Secondly, HPV also causes diseases in men, including cancer of the penis; Smith *et al.* (2) reported that the prevalence of overall HPV was 16% in men. However, the objectives of our study did not include an evaluation of gender-specific prevalence of HPV infection. It would be interesting in future studies to investigate whether the prevalence HPV infection in men matches that of women, since vaccinating males is also considered important in the United States. Finally, previous studies have shown that sexual behavior and the number of sexual partners is associated with HPV infection (1). We did not collect such lifestyle information and therefore could not assess the role of lifestyle on HPV infections. The lack of lifestyle information hampers a practical guide for the population in preventing HPV infection.

Conclusions

In conclusion, our results suggest that a screening program for HPV infection as well as cervical cancer is necessary for both older and younger women. Vaccination and educational campaigns concerning HPV infection should start in younger females. The current findings should be useful and beneficial toward vaccine development and the implementation of vaccination programs.

Acknowledgements

We thank Beijing Genomics Institute (BGI) in Shenzhen for their support performing HPV detection and genotyping, and Medjaden Bioscience for editing the manuscript. We also thank Dr. Zhongren Du for critically reviewing the manuscript. We are grateful to all the participants of the Cervical Cancer Screening Program in Heyuan.

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

Ethical Statement: The study was approved by Ethics Committee of the hospital and written informed consent was obtained from all patients.

References

1. Vinodhini K, Shanmughapriya S, Das BC, et al. Prevalence and risk factors of HPV infection among women from various provinces of the world. *Arch Gynecol Obstet* 2012;285:771-7.
2. Smith JS, Gilbert PA, Melendy A, et al. Age-specific prevalence of human papillomavirus infection in males: a global review. *J Adolesc Health* 2011;48:540-52.
3. de Freitas AC, Gurgel AP, Chagas BS, et al. Susceptibility to cervical cancer: an overview. *Gynecol Oncol* 2012;126:304-11.
4. de Sanjosé S, Diaz M, Castellsagué X, et al. Worldwide prevalence and genotype distribution of cervical human papillomavirus DNA in women with normal cytology: a meta-analysis. *Lancet Infect Dis* 2007;7:453-9.
5. Forman D, de Martel C, Lacey CJ, et al. Global burden of human papillomavirus and related diseases. *Vaccine* 2012;30 Suppl 5:F12-23.
6. Liu SS, Chan KY, Leung RC, et al. Prevalence and risk factors of Human Papillomavirus (HPV) infection in southern Chinese women - a population-based study. *PLoS One* 2011;6:e19244.
7. Chen Q, Luo ZY, Lin M, et al. Prevalence and genotype distribution of human papillomavirus infections in women attending hospitals in Chaozhou of Guangdong province. *Asian Pac J Cancer Prev* 2012;13:1519-24.
8. Wu RF, Dai M, Qiao YL, et al. Human papillomavirus infection in women in Shenzhen City, People's Republic of China, a population typical of recent Chinese urbanisation. *Int J Cancer* 2007;121:1306-11.
9. Zhang R, Shi TY, Ren Y, et al. Risk factors for human papillomavirus infection in Shanghai suburbs: a population-based study with 10,000 women. *J Clin Virol* 2013;58:144-8.
10. Ye J, Cheng X, Chen X, et al. Prevalence and risk profile of cervical Human papillomavirus infection in Zhejiang Province, southeast China: a population-based study. *Virol*

- J 2010;7:66.
11. Li LK, Dai M, Clifford GM, et al. Human papillomavirus infection in Shenyang City, People's Republic of China: A population-based study. *Br J Cancer* 2006;95:1593-7.
 12. Chui SH, Szeto YT, Lam CW. Human papillomavirus infection in Macau women. *Public Health* 2012;126:600-4.
 13. Zheng L, Li Y, Lu S, et al. Physical characteristics of Chinese Hakka. *Sci China Life Sci* 2013;56:541-51.
 14. Jing L, Zhong X, Zhong Z, et al. Prevalence of human papillomavirus infection in Guangdong Province, China: a population-based survey of 78,355 women. *Sex Transm Dis* 2014;41:732-8.
 15. Yi X, Li J, Yu S, et al. A new PCR-based mass spectrometry system for high-risk HPV, part I: methods. *Am J Clin Pathol* 2011;136:913-9.
 16. Chen Z, Wang Q, Ding X, et al. Characteristics of HPV prevalence in Sichuan Province, China. *Int J Gynaecol Obstet* 2015;131:277-80.
 17. Vedham V, Divi RL, Starks VL, et al. *Technol Cancer Res Treat* 2014;13:177-94.
 18. Lee JE, Lee S, Lee H, et al. Association of the vaginal microbiota with human papillomavirus infection in a Korean twin cohort. *PLoS One* 2013;8:e63514.
 19. Allsworth JE, Lewis VA, Peipert JE. Viral sexually transmitted infections and bacterial vaginosis: 2001-2004 National Health and Nutrition Examination Survey data. *Sex Transm Dis* 2008;35:791-6.
 20. Piris S, Bravo V, Alvarez C, et al. Natural history of histologically moderate cervical dysplasia in adolescent and young women. *Onco Targets Ther* 2014;7:2101-6.

Cite this article as: Zeng XX, Yan LX, Huang XX, He CH, Liu WG, Yuan WQ, Qiu YP, Liu ZX. Prevalence and genotype distribution of human papillomavirus among Hakka women in China. *Ann Transl Med* 2016;4(15):276. doi: 10.21037/atm.2016.05.49