HPV16 genetic variation and the development of cervical cancer worldwide

I Cornet¹, T Gheit¹, MR Iannacone¹, J Vignat¹, B S Sylla¹, A Del Mistro², S Franceschi¹, M Tommasino¹, G M Clifford *¹ and the IARC HPV Variant Study Group ³

¹International Agency for Research on Cancer, 150 cours Albert Thomas, 69372 Lyon Cedex 08, France and ²Istituto Oncologico Veneto, Via Gattamelata 64, 35128 Padova, Italy

Background: Factors that favour a small proportion of HPV16 infections to progress to cancer are still poorly understood, but several studies have implicated a role of HPV16 genetic variation.

Methods: To evaluate the association between HPV16 genetic variants and cervical cancer risk, we designed a multicentre case–control study based on HPV16-positive cervical samples (1121 cervical cancer cases and 400 controls) from the International Agency for Research on Cancer biobank. By sequencing the E6 gene, HPV16 isolates were classified into variant lineages and the European (EUR)-lineage isolates were subclassified by the common polymorphism T350G.

Results: Incidence of variant lineages differed between cases and controls in Europe/Central Asia (P = 0.006, driven by an underrepresentation of African lineages in cases), and South/Central America (P = 0.056, driven by an overrepresentation of Asian American/North American lineages in cases). EUR-350G isolates were significantly underrepresented in cervical cancer in East Asia (odds ratio (OR) = 0.02 vs EUR-350T; 95% confidence interval (CI) = 0.00–0.37) and Europe/Central Asia (OR = 0.42; 95% CI = 0.27–0.64), whereas the opposite was true in South/Central America (OR = 4.69; 95% CI = 2.07–10.66).

Conclusion: We observed that the distribution of HPV16 variants worldwide, and their relative risks for cervical cancer appear to be population-dependent.

High-risk human papillomavirus (HPV) types are the aetiological agents of cervical cancer (zur Hausen, 2002), of which HPV16 is the most prevalent type worldwide, both in women without cervical abnormalities and in cervical cancer (Guan et al, 2012). Factors that favour a small proportion of HPV16 infections to progress to cancer are still poorly understood, but several studies have implicated a role of HPV16 genetic variation (Villa et al, 2000; Berumen et al, 2001; Sathish et al, 2005; Zuna et al, 2009; Schiffman et al, 2010; Gheit et al, 2011).

HPV16 variants have previously been classified into four major lineages based upon common phylogenetic patterns of single-nucleotide polymorphisms: (1) European Asian, including the sublineages European (EUR), and Asian (As), (2) African 1 (AFR1), (3) African 2 (AFR2) and (4) Asian American/North American (AA/NA), including the sublineages Asian American 1, Asian American 2 and North American (Yamada et al, 1995, 1997; Cornet et al, 2012). Other positions are frequently polymorphic within one or more lineages, but do not define phylogenetic sublineages (Chen et al, 2005). A common such polymorphism within the EUR lineage is T350G that is localised in the E6 oncogene, and leads to an amino acid change from leucine to valine (L83V). Thus, the EUR lineage can be divided into isolates containing 350T (EUR-350T, which includes the prototype HPV16 sequence) or 350G (EUR-350G). This polymorphism has been suggested to influence the persistence and risk of progression to precancerous cervical lesions of EUR variant lineages (Zehbe et al, 2001; Grodzki et al, 2006; Gheit et al, 2011), and also occurs in non-EUR lineages.
Valid E6 sequence data was obtained for a total of 1121 HPV16-positive cases and 400 HPV16-positive controls. Distribution of cases and controls by variant lineage and country/region is shown in Table 1. The EUR lineages accounted for the majority of isolates and were common in all regions except Sub-Saharan Africa and East Asia. The AA/NA lineages were detected in all regions, except Sub-Saharan Africa. The AFR lineages predominated in North and Sub-Saharan Africa, whereas the As lineage was specific for East Asia (Table 1).

In most regions, patterns of the HPV16 variant lineages did not differ significantly between cases and controls. However, variant lineages differed significantly between cases and controls in Europe/Central Asia (Fisher’s exact test \( P = 0.006 \)), driven by an overrepresentation of AFR lineages in controls from Italy. There was also a difference of borderline statistical significance in South/Central America (Fisher’s exact test, \( P = 0.056 \)), apparently driven by an overrepresentation of AA/NA isolates in cases.

The EUR isolates were further stratified into EUR-350T and EUR-350G, and their distribution compared between cases and controls for all regions, except sub-Saharan Africa (Table 2). The relative proportion of EUR-350T to EUR-350G among controls varied considerably by country/region (EUR-350T was more common than EUR-350G among cases in all regions, except South/Central America), as did their relationship with cervical cancer (Table 2). In South/Central America, the EUR-350G isolates were significantly overrepresented in cervical cancer compared with controls (\( OR = 4.69 \) vs EUR-350T; 95% CI = 2.07–10.66), an effect observed both in Argentina and Chile. In contrast, EUR-350G isolates were significantly underrepresented in cervical cancer in East Asia (\( OR = 0.02 \) vs EUR-350T; 95% CI = 0.00–0.37) and Europe/Central Asia (\( OR = 0.42 \) vs EUR-350T; 95% CI = 0.27–0.64), an effect that was consistent in all countries within those regions. The crude OR estimate for EUR-350G vs EUR-350T for South/Central America was significantly heterogeneous (\( P < 0.05 \)) from that of each of the other regions (data not shown).

In Sub-Saharan Africa, where the AFR lineages predominated, no differences were observed between cases and controls in the pattern of AFR sublineages [as defined by Cornet et al (2012)], namely AFR1a (23 cases and 20 controls), AFR1b (0 and 1), AFR2a (18 and 18) and AFR2b (3 and 3); 5 other AFR isolates were not classifiable by AFR sub-lineage based upon E6 alone; data not shown].

### DISCUSSION

Using a multicentre case–control comparison based on HPV16-positive samples selected from the IARC biobank, we were able to identify significant associations between HPV16 variants and cervical cancer risk. Furthermore, we observed that the distribution of HPV variant lineages and EUR-350T/G worldwide, and the corresponding relative risks for cervical cancer, were population-dependent.

The distribution of major variant lineages around the world was confirmed to be highly geographically/ethnically specific (Yamada et al, 1997; Tornesello et al, 2004; Kang et al, 2005; Chimeddorj et al, 2008; Pande et al, 2008), limiting the possibility to compare their carcinogenic potential in a standardised way across all regions. However, our data did suggest an underrepresentation of AFR variants in cervical cancer cases in Europe/Central Asia, and possibly an overrepresentation of AA/NA variants in cervical cancer cases in South/Central America. Although these findings require further clarification, they do fit with the results of a study suggesting that HPV16 variants tend to persist and progress to
cervical intra-epithelial neoplasia grade 3 (CIN3) better in a host whose ethnicity shares an ancestral origin (Xi et al, 2007), even if this was not observed among greater racial admixed populations of Latin America (Hildesheim et al, 2001; Sichero et al, 2007). The AA/NA variants have been previously associated with higher CIN3 and cervical cancer risk in studies from Costa Rica (Smith et al, 1998), and Mexico (Berumen et al, 2001), respectively, and with higher capacity for in vitro transformation of human keratinocytes (Sichero et al, 2012).

The EUR lineages accounted for a large proportion of HPV16 isolates in all regions, except Sub-Saharan Africa. Hence, analyses were well powered to study cervical cancer risks associated with the common EUR-350T/G polymorphism, revealing significant heterogeneity by world region; in Europe/Central Asia and East Asia, cervical cancer risk was significantly associated with the EUR-350T variants in comparison with EUR-350G. However, the opposite was true in South/Central America. Although puzzling, hints of this regional heterogeneity can be found in previous country-specific studies; EUR-350T variants were overrepresented in cervical cancers in comparison with EUR-350G in studies from both the Netherlands (Bontkes et al, 1998) and China (Chan et al, 2002), although no difference was seen in some other small series from Europe (Nindl et al, 1999; Hu et al, 2001a,b; Tornesello et al, 2004). Another study suggested that the carcinogenicity of EUR-350T vs EUR-350G might be population-dependent also within Europe (Zehbe et al, 2001). Furthermore, EUR-350T infections were more likely to persist and progress to CIN3 in comparison with EUR-350G in Denmark (Gheit et al, 2011). With respect to the opposite finding in South/Central America, a similarly strong association of EUR-350G with cervical cancer has been observed in a previous study from Argentina (Picconi et al, 2003), and a Brazilian study has recently reported a higher capacity for EUR-350G than EUR-350T variants to transform human keratinocytes in vitro (Sichero et al, 2012).

These differences by population might be explained by residual genetic heterogeneity within HPV16 genomes classified solely upon position 350, which, although highly polymorphic, does not seem to robustly define phylogenetic sublineages (Chen et al, 2005). Alternatively, host genetic factors, such as HLA haplotypes or TP53 polymorphisms, which differ by population, could have a role in the association between a particular HPV16 variant and cervical cancer development (Bontkes et al, 1998; van Duin et al, 2000; Zehbe et al, 2001). Whatever the underlying cause, such apparent geographical differences reveal an inherent complexity in studies of HPV16 variants and cervical cancer risk, and give a warning about the extent to which data can be pooled across countries/regions.

The availability of HPV16-positive controls, rather than HPV16-positive cases, is the limiting factor in the statistical power of this study and other studies of HPV16 variants and cervical cancer, given that they need to be derived from large population-based samples. This is particularly the case in low-resource settings with no cervical screening programmes. Nevertheless, thanks to the reliance on 20 years of IARC studies on HPV and cancer from around the world, the number of HPV16-positive cases, is the limiting factor in the statistical power of this study and other studies of HPV16 variants and cervical cancer, given that they need to be derived from large population-based samples. This is particularly the case in low-resource settings with no cervical screening programmes. Nevertheless, thanks to the reliance on 20 years of IARC studies on HPV and cancer from around the world, the number of HPV16-positive controls in the current study is the largest studied to date.

In conclusion, although our findings suggest that HPV16 variant analysis has no clinical application, understanding the genetic basis of differences in the carcinogenicity of HPV16 variants (which may be linked to genetic changes in non-E6 parts of the genome) may help us unravel important biological and/or immunological interactions between virus and host that could lead to better tools to control HPV infection and its malignant consequences.

### Table 1. Distribution of HPV16 variant (sub)lineages in cases and controls, by country and region

| Country/region | N cases/N controls | EUR | As | AFR | AA/NA | P-valuea |
|----------------|--------------------|-----|----|-----|-------|----------|
| Africa, North  | 234/17             | 89/7| 1/0| 116/6| 28/4  | 0.359    |
| Algeria         | 109/12             | 57/6| 1/0| 35/4 | 16/2  | 1.000    |
| Morocco         | 125/5              | 32/1| —  | 81/2 | 12/2  | 0.126    |
| Africa, Sub-Saharan | 42/50            | 1/1 | 0/0| 42/49| 0/0   | 1.000    |
| Guinea          | 40/25              | 1/0 | —  | 39/5 | —     | 1.000    |
| Nigeria         | 3/25               | 0/1 | —  | 3/24 | —     | 1.000    |
| America, South/Central | 102/44     | 86/43| 0/0| 2/0  | 14/1  | 0.056    |
| Chile           | 67/15              | 56/15| —  | 1/0  | 10/0  | 0.342    |
| Argentina       | 35/29              | 30/28| —  | 1/0  | 4/1   | 0.366    |
| Asia, East      | 258/28             | 37/5| 181/20| 11/0 | 29/3  | 0.839    |
| Korea           | 47/4               | 7/2 | 35/2| —    | 5/0   | 0.181    |
| Thailand        | 211/24             | 30/3| 146/18| 11/0 | 24/3  | 0.889    |
| Asia, West      | 151/75             | 136/67| 2/1| 1/0  | 12/7  | 0.940    |
| India           | 101/59             | 95/53| 2/0| —    | 4/6   | 0.215    |
| Nepal           | 24/12v             | 22/10| 0/1| —    | 2/1   | 0.510    |
| Pakistan        | 26/4v              | 19/4| —  | 1/0  | 6/0   | 0.612    |
| Europe/Central Asia | 332/186       | 324/177| 0/1| 0/5  | 9/3   | 0.006    |
| Georgia         | 52/3               | 49/3| —  | —    | 3/0   | 1.000    |
| Italy           | 146/110            | 141/102| 0/5| 5/3  | 0.039 |
| Poland          | 83/26              | 82/26| —  | —    | 1/0   | 1.000    |
| Mongolia        | 52/47              | 52/46| 0/1| —    | —     | 0.475    |
| Total           | 1121/400           | 673/300| 184/22| 172/60| 92/18 |

Abbreviations: AA/NA = Asian American/North American; AFR = African; As = Asian; EUR = European.

*P-value of Fisher’s exact test.
The authors declare no conflict of interest.

REFERENCES

Berumen J, Ordonez RM, Laxcano E, Salmeron J, Galvan SC, Estrada RA, Yunes E, Garcia-Carranca A, Gonzalez-Lira G, Madrigal-de la Campa A (2001) Asian-American variants of human papillomavirus 16 and risk for cervical cancer: a case-control study. J Natl Cancer Inst 93: 1325–1330.

Bontkes HJ, van Duin M, de Gruijl TD, Duggan-Keen ML, Walboomers JM, Stukart MJ, Verheijen RH, Helmerhorst TJ, Meijer CJ, Schepers RJ, Stevens FR, Dyer PA, Sinnott P, Stern PL (1998) HPV 16 infection and progression of cervical intra-epithelial neoplasia: analysis of HLA polymorphism and HPV 16 E6 sequence variants. Int J Cancer 78: 166–171.

Bosch FX, Manos MM, Muñoz N, Sherman M, Jansen AM, Peto J, Schiffman MH, Moreno V, Kurman R, Shah KV (1995) Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. International biological study on cervical cancer (IBSSC) Study Group. J Natl Cancer Inst 87: 796–802.

Chan PK, Lam CW, Cheung TH, Li WW, Lo KW, Chan MY, Cheung JL, Xu LY, Cheng AF (2002) Human papillomavirus type 16 intratypic variant infection and risk for cervical neoplasia in southern China. J Infect Dis 186: 696–700.

Chen Z, Terai M, Fu L, Herrera R, Desalle R, Burk RD (2005) Diversifying selection in human papillomavirus type 16 lineages based on complete genome analyses. J Virol 79: 7014–7023.

Chimeddoorij B, Pak CY, Damdin A, Okamoto N, Miaygi Y (2008) Distribution of HPV-16 intratypic variants among women with cervical intraepithelial neoplasia and invasive cervical cancer in Mongolia. Asian Pac J Cancer Prev 9: 563–568.

Clifford GM, Gallus S, Herrera R, Muñoz N, Snijders PJ, Vaccarella S, Anh PT, Ferreccio C, Hieu NT, Matos E, Molano M, Rajkumar R, Ronco G, de Sanjose S, Shin HR, Sukvichar S, Thomas JO, Tunsakul S, Meijer CJ, Franceschi S. and the IARC HPV Prevalence Surveys Study Group (2005) Worldwide distribution of human papillomavirus types in cytologically normal women in the International Agency for Research on Cancer HPV prevalence surveys: a pooled analysis. Lancet 366: 991–998.

Corneil G, Gheit T, Franceschi S, Vignat J, Burk RD, Sylla BS, Tommasino M, Clifford GM (2012) Human papillomavirus type 16 genetic variants: phylogeny and classification based on E6 and LCR. J Virol 86: 6855–6861.

Del Mistro A, Frayle-Salamanca H, Trevisan R, Matteucci M, Pinarello A, Zambenedetti P, Buoso R, Fantin GP, Zorzi M, Minucci D (2010) Triage of women with atypical squamous cells of undetermined significance (ASC-US): results of an Italian multicentric study. Gynecol Oncol 117: 77–81.

Ferreccio C, Van De Weygaard V, Oclay F, Dominguez MA, Puschel K, Corvalan AH, Franceschi S, Snijders PJ (2011) High-risk HPV infection after five years in a population-based cohort of Chilean women. Infect Agent Cancer 6: 21.

Gheit T, Corneil I, Clifford GM, Iftner T, Munk C, Tommasino M, Kjaer SK (2011) Risks for persistence and progression by human papillomavirus type 16 variant lineages among a population-based sample of Danish women. Cancer Epidemiol Biomarkers Prev 20: 1315–1321.

Grodzki M, Besson G, Clavel C, Arslan A, Franceschi S, Birembaut P, Tommasino M, Zehbe I (2006) Increased risk for cervical disease progression of French women infected with the human papillomavirus type 16 E6-350G variant. Cancer Epidemiol Biomarkers Prev 15: 821–822.

Guan P, Howell-Jones R, Li N, Bruni L, de Sanjose S, Franceschi S, Clifford GM (2012) Human papillomavirus types in 115,789 HPV-positive women: a meta-analysis from cervical infection to cancer. Int J Cancer 131: 2349–2359.

Hildesheim A, Schiffman M, Bromley C, Wacholder S, Herrero R, Rodriguez A, Bratti MC, Sherman ME, Scarpidis U, Lin QQ, Terai M, Bromley RL, Buetow K, Apple RJ, Burk RD (2001) Human papillomavirus type 16 variants and risk of cervical cancer. J Natl Cancer Inst 93: 315–318.

Hu X, Pang T, Guo Z, Mazurenko N, Kissievich F, Ponten J, Nistér M (2001a) HPV16 E6 gene variations in invasive cervical squamous cell carcinoma and cancer in situ from Russian patients. Br J Cancer 84: 791–798.

Hu X, Pang T, Guo Z, Ponten J, Nister M, Bernard AG (2001b) Oncogene lineages of human papillomavirus type 16 E6, E7 and E5 in preinvasive and invasive cervical squamous cell carcinoma. J Pathol 195: 307–311.

Jacobs MV, Walboomers JM, Snijders PJ, Voorhorst FJ, Verheijen RH, Hildesheim A, Schiffman M, Bromley C, Wacholder S, Herrero R, Rodriguez A, Bratti MC, Sherman ME, Scarpidis U, Lin QQ, Terai M, Bromley RL, Buetow K, Apple RJ, Burk RD (2001) Human papillomavirus type 16 intratypic variants and risk of cervical cancer. J Natl Cancer Inst 93: 315–318.

Kjaer SK, Munoz N, Meijer CJ, Herrero R, Shah KV (1995) Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. International biological study on cervical cancer (IBSSC) Study Group. J Natl Cancer Inst 87: 796–802.

Kang S, Jeon YT, Kim JW, Park NH, Song YS, Kang SB, Lee HP (2005) Polymorphism in the E6 gene of human papillomavirus type 16 in the cervical tissues of Korean women. Int J Gynecol Cancer 15: 107–112.

Kjaer SK, Munoz N, Meijer CJ (2003) Epidemiologic classification of human papillomavirus types associated with cervical cancer. N Engl J Med 348: 518–527.

Lindl R, Rindfleisch K, Lotz B, Schneider A, Durst M (1999) Uniform distribution of HPV 16 E6 and E7 variants in patients with normal...
histology, cervical intra-epithelial neoplasia and cervical cancer. Int J Cancer 82: 203–207.

Pande S, Jain N, Prusty BK, Bhambhani S, Gupta S, Sharma R, Batra S, Das BC (2008) Human papillomavirus type 16 variant analysis of E6, E7, and L1 genes and long control region in biopsy samples from cervical cancer patients in north India. J Clin Microbiol 46: 1060–1066.

Picconi MA, Alonio LV, Sichero L, Mbayed V, Villa LL, Gronda J, Campos R, Teysse A (2003) Human papillomavirus type-16 variants in Quechua aboriginals from Argentina. J Med Virol 69: 546–552.

Sathish N, Abraham P, Peedicayil A, Sridharan G, Chandy G (2005) HPV 16 E6 sequence variations in Indian patients with cervical neoplasia. Cancer Lett 229: 93–99.

Schiffman M, Rodriguez AC, Chen Z, Wacholder S, Herrero R, Hildesheim A, Villalobo C, Viera SS, Buonaguro FM (2004) Analysis of human papillomavirus type-16 and 18 preferentially associated with cervical neoplasia. J med Virol 74: 117–126.

APPENDIX

The members of the IARC HPV variant study group include previous IARC staff (N Munoz, R Herrero, X Bosch) and local study coordinators (in alphabetical order by country): Algeria (D Hammouda); Argentina (D Loria, E Matos); Chile (C Ferreccio, van Duin M, Snijders PJ, Vossen MT, Klaassen E, Voorhorst F, Verheijen RH, Helmerhorst TJ, Meijer CJ, Walboomers JM (2000) Analysis of human papillomavirus type 16 E6 variants in relation to p53 codon 72 polymorphism genotypes in cervical carcinogenesis. J Gen Virol 81: 317–325.

Villa LL, Sichero L, Rahal P, Caballero O, Ferenczy A, Rohan T, Franco EL (2000) Molecular variants of human papillomavirus types 16 and 18 preferentially associated with cervical neoplasia. J Gen Virol 81: 2959–2968.

Xi LF, Koutsky LA, Hildesheim A, Galloway DA, Wheeler CM, Winer RL, Ho J, Kiviat NB (2007) Risk for high-grade cervical intraepithelial neoplasia associated with variants of human papillomavirus types 16 and 18. Cancer Epidemiol Biomarkers Prev 16: 4–10.

This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution-NonCommercial-Share Alike 3.0 Unported License.

F Aguayo); Georgia (T Alibegashvili, D Kordaia); Guinea (N Keita, M Koulibaly); India (T Rajkumar, R Rajkumar); Italy (G Ronco, M Sideri); Korea (D-H Lee, H R Shin); Morocco (N Chaouki); Nepal (A Sherpa); Nigeria (OJ Thomas, C Okolo, I Adewole); Mongolia (B Dondog); The Netherlands (CJLM Meijer, PJF Snijders); Pakistan (A Raza); Poland (W Zatonski); and Thailand (S Chichareon, S Sukvirach, S Tunsakul).