Distinct geographic clustering of oncogenic human papillomaviruses multiple infections in cervical cancers: results from a worldwide cross-sectional study

Ville N. Pimenoff 1,2*, Sara Tous 3,4*, Yolanda Benavente 3,4, Laia Alemany 3,4, Wim Quint 5, Francesc Xavier Bosch 3,4, Ignacio G. Bravo 6, Silvia de Sanjose 3,4

1 Unit of Biomarkers and Susceptibility, Bellvitge Institute of Biomedical Research (IDIBELL), Catalan Institute of Oncology (ICO), Barcelona (Spain)
2 Department of Epidemiology, University of Tampere, Finland.
3 Unit of Infections and Cancer (UNIC), Cancer Epidemiology Research Program, Catalan Institute of Oncology, Institut Català d’Oncologia (ICO), L’Hospitalet de Llobregat, Barcelona (Spain)
4 CIBER Epidemiología y Salud Pública, CIBERONC, Barcelona, (Spain)
5 DDL Diagnostic Laboratory, Rijswijk (The Netherlands)
6 National Center for Scientific Research (CNRS), Laboratory MIVEGEC (UMR CNRS, IRD, UM), Montpellier (France).

* Authors contributed equally to this work.

To whom correspondence should be addressed:

vnpimenoff@iconcologia.net (V.N.P)
Bellvitge Institute of Biomedical Research (IDIBELL)
Catalan Institute of Oncology (ICO)
Av Gran Via 199, L’Hospitalet de Llobregat, 08908 Barcelona
Tel: +(34) 93 2607186 / Fax: +(34) 93 2607188

Keywords: Multiple HPVs infections, invasive cervical cancer, geographic clustering

Running title: Multiple HPVs infections in invasive cervical cancers

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/ijc.31964

This article is protected by copyright. All rights reserved.
While it isn’t common for an invasive cervical cancer (ICC) to carry more than one strain of HPV, it does occur. Might these co-infections affect tumor progression or other characteristics of ICC? In this HPV-genotyping study of ICC biopsies worldwide, the authors found several strains that were more likely to occur together. This supports the occasional role of Alpha-10 LR-HPVs in cervical cancers. The study also found a distinct co-infection pattern in African women. These results may be valuable in estimating type competition, and indicate that monitoring even after HPV vaccination is warranted.

ABSTRACT

Co-infections by multiple Human Papillomaviruses (HPVs) are observed in approximately 6-8% of invasive cervical cancer (ICC) cases worldwide. But neither the presence of persistent HPVs co-infections nor their etiological role in the development of ICC is well understood. Cervical HPVs co-infections have been observed randomly, mostly in women with pre-neoplastic lesions, and only few studies have globally analyzed ICC cases. Here we explored the HPVs multiple infection patterns in a large worldwide sample of cross-sectional ICC cases. Paraffin-embedded ICC biopsy samples were tested using stringent HPV genotyping. Logistic regression models were used to identify the most likely pairwise HPV types in multiple infections. Multivariate analysis was applied to detect significant HPV co-infection patterns beyond pairwise HPVs comparison. Among 8,780 HPV DNA-positive ICC cases worldwide, 6.7% (N=587) contained multiple HPVs. Pairwise analysis revealed that HPV16|74, HPV31|33, HPV31|44, HPV33|44 and HPV45|70 pairs were significantly more frequently found together in multiple infections compared to any other HPV type combination, which supports the occasional role of Alpha-10 LR-HPVs in cervical cancers. In contrast, HPV16|31, HPV16|45, HPV16|51 and HPV18|HPV45 pairs were significantly less frequently found together than with any other HPV pair combination. Multivariate analysis sustained the results and revealed for the first time a distinct co-infection pattern in African women.
pattern in African ICCs stemming from the clustering of oncogenic HPV51/35/18/52 co-infections in African women. We suggest that the differential geographic HPVs co-infections clustering observed might be compatible with a specific modulation of the natural history/oncogenic potential of particular HPVs multiple infections and warrant monitoring for post-vaccinated.

**Novelty and Impact**

Following clonal origin of HPV-induced lesions co-infections are rare in ICCs, and thus difficult to study at population-level. Here we provide worldwide estimations on multiple HPV-infected ICCs. Our results support an occasional role of Alpha-10 low-risk HPVs in ICCs. Further, we identified a distinct co-infection pattern in African ICCs compared to other world regions. These results, although with likely minor public health importance, may be valuable in estimating type competition, and warrants monitoring for post-vaccinated.
BACKGROUND

Among the more than 350 different Papillomaviruses (PVs) identified (https://pave.niaid.nih.gov), above 200 are Human Papillomaviruses (HPVs). According to their carcinogenic potential, twelve HPVs are acknowledged as carcinogenic to humans (IARC Group I: HPV16/18/31/33/35/39/45/51/52/56/58/59), while 13 additional HPVs are classified as probably or possibly carcinogenic to humans (IARC Groups 2a and 2b: HPV26/30/34/53/66/67/68/69/70/73/82/85/97). All of them are usually referred to as high-risk (HR) HPVs. Two HPVs cannot be classified as to their carcinogenicity to humans (IARC Group 3: HPV6 and 11) and are referred to as low-risk (LR) HPVs. Other viruses such as HPV40, 42, 43, 44, 54, and 74 have not been included in any of the IARC Groups but classified by IARC as LR-HPVs. All of these above mentioned HPV types belong to Alphapapillomaviruses within the Papillomaviridae family.

HPVs are the most common infectious agents in sexually transmitted infections: above 80% of sexually active women are infected by any HPVs during their lifetime. Fortunately, most of these HPVs infections are transient and cleared by the immune system, while persistent infection with HR-HPVs is a necessary cause for invasive cervical cancer (ICC) development. Co-infections by multiple HPVs are rare in women with normal cytology (between 1% and 3%), common in women with pre-cancerous lesions (between 15% and 41%) and display intermediate levels in HPV-induced ICCs with co-infections being observed between 6% and 8% of the cases. The simultaneous infection by different HPVs seems to have an impact on the natural history of the viral infection. Thus, when compared to single HPV infections, cancerous
lesions with multiple HR-HPVs infections contain more oncogenic \textit{E6/E7} HPV mRNA, often from more than one HR-HPV, thus suggesting that HPVs co-infections in pre-cancerous lesions might be more susceptible to carcinogenesis than single infections. Additionally, women co-infected with at least two different HPVs in the cervix have a higher risk of high-grade cervical lesions, of ICC or of acquiring another HPV infection, compared to those infected with a single type.

Current consensus on the etiology of HPVs-related cancers is that individual lesions are associated with individual HPV infections, meaning that lesions, and thus HPVs-induced cancers, are clonal. Accordingly, given the large differences in prevalence of HPVs co-infections in lesions of different grade and grade combinations, the etiological role of multiple HPVs co-infections in the development of virus-induced cancers, if any, still needs to be evaluated. This is especially true given the increasing number of large-scale HPVs vaccination programs currently implemented using bivalent vaccine protective against HPV16/18/31/33/35/45 types, and further inducing herd effects against HPV18/31/33/35 types. In parallel, the 9-valent HPV vaccine is protective against seven HR-HPV types (HPV16/18/31/33/45/52/58) and when these vaccines are widely implemented they will thoroughly change the prevalence pattern of the targeted HR-HPVs infections in humans.

The aim of this study was to explore the prevalence and patterns of HPVs co-infections in a large worldwide cross-sectional study of ICC cases using standardized HPV genotyping methods and applying different statistical models.
METHODS

Study design

This study is based on an international retrospective cross-sectional study of presence of HPVs in ICC cases as previously described. The original study included information from 10,575 ICCs and HPV DNA was detected in 85% of the cases. Briefly, paraffin-embedded specimens from consecutive cases (aged 16-97 years) diagnosed with ICC were obtained from histological archives from 38 countries worldwide. This study and all protocols were approved by both local and ICO ethics committees. After pathological confirmation of primary ICC, HPV DNA detection was performed by PCR with SPF-10 broad-spectrum primers followed by DNA enzyme immunoassay (DEIA) and genotyping with a reverse hybridization line probe assay (LiPA25 SPF10-PCR) – LiPA25 that detects 25 HPV types (HPV6, 11, 16, 18, 31, 33, 34, 35, 39, 40, 42, 43, 44, 45, 51, 52, 53, 54, 56, 58, 66, 68, 70, and 74) (see details in ). Cases showing DEIA positive but LiPA25 negative results (undetermined HPV) were further subjected to Sanger sequencing for HPV identification, as previously described. Special care was given to maintain stringent conditions for HPV type-level detection using the SPF10-PCR LiPA25, a method particularly suitable for HPV genotyping from FFPE specimens. To further prevent contaminations the paraffin blocks in this study were processed under strict conditions: four paraffin sections were systematically obtained from each block. First and fourth sections were used for histopathological assessment after haematoxylin and eosin staining, and the second and third sections were used for HPV DNA detection and genotyping as described. Importantly, a blank paraffin section was cut and processed in-between specimens to control for any carryover contamination in addition to the standard controls.
Statistical analyses

ICC cases were classified into single or multiple HPV-infected cases based on the above described genotyping results. Appropriate statistical tests were used to compare the proportion of single and multiple HPVs infections by the design variables (age at diagnosis, geographical origin, period of diagnosis and histological type). We estimated prevalence ratio and 95% bootstrapped-confidence interval of single/multiple (S/M) infections for each individual HPV. Unconditional logistic regression models were used to estimate association between demographic and clinical characteristic of the cases, probability of multiple infections, and to identify pairwise co-infections of HPVs that were more likely to be detected together compared to being in combination with any other HPV in multiple infections. To simultaneously estimate the possible associations beyond pairwise HPVs comparison among multiple HPV infected ICCs collected worldwide, we fitted the histological type, period of diagnosis and geographical origin of the sample into our generalised linear model (GLM) of the multivariate HPVs multiple infections dataset\textsuperscript{16}. Further, a univariate tests for each HPV in co-infected ICCs were implemented within multivariate analysis to estimate whether there were type-specific differences in HPV co-infection distribution by histology, period of diagnosis or geographic origin. The observed pairwise and composition level HPVs co-infection pattern differences were visualized using network plots constructed with \textit{igraph} R package (R Core Team, http://www.R-project.org/). Classification and regression trees (CART) analyses \textsuperscript{17} were performed using the \textit{rpart} R package to estimate and visualize the possible HPVs co-infection patterns beyond pairwise HPVs comparison, accounting for age at diagnosis, histology, period of diagnosis and geographic origin of the cases. In practice, CART analyses use recursive partitioning to
structure the data into a series of dichotomous splits, e.g. presence or absence of each HPV type and other demographic variables, to create a decision tree with the goal of correctly classifying clusters of the above mentioned interacting variables, for example, HPV16 or HPV18 co-infected ICCs. Each independent variable is examined and a split is introduced to maximize the sensitivity and specificity, calculated using the predicted probability from the multivariable logistic regression of the classification, and resulting in a decision tree. All regression models were adjusted by the design variables and Bonferroni’s correction was implemented in pairwise analyses to control for multiple comparisons. All statistical tests were two-sided and rejecting the null hypothesis was set at a $P$ value of less than 0.05. Analyses were carried out with Stata software (Release 13, StataCorp, USA) except otherwise specified.

RESULTS

Demographic characteristics of the ICCs

From the 10,575 ICCs previously reported, we excluded the ICC cases with a negative HPV DNA test (n=1,598) and cases containing DNA from undetermined HPV further tested with Sanger sequencing (n=197). Thus, the dataset used in this study contained 8,780 ICC HPV-positive cases, most of them (91.8%; N=8,064) histologically diagnosed as squamous cell carcinomas (SCCs, Table 1). The final dataset of HPV-positive ICC cases consisted of women with a mean age of 51 years (range 17-97) at diagnosis and 23.1% recruited from Europe, 39.6% from the Americas, 31.3% from Asia including Oceania, and 6% from Africa (Table 1). Since infection with other pathogens, notably Human Immunodeficiency Virus, could have an impact on age at ICC onset, we verified that the mean age at ICC diagnosis exclusively for African women did not differ from
that for women outside Africa ($P=0.240$, data not shown). Most (n=8,193) HPV-positive ICC cases showed the presence of a single HPV, and in 83.8% (N=492) of the cases with multiple infections only two distinct HPVs were detected (Supplementary Figure 1). Logistic regression analysis among all ICCs showed prevalence differences between single and multiple HPVs infected ICCs by geographic origin and by period of diagnosis (Table 1). By geographic region, prevalence of multiple HPVs infections was consistently about 0.3-fold less frequent ($P<0.001$) among women from Europe (0.36 [95% CI: 0.27-0.49]), Asia (0.31[95% CI: 0.23-0.42]) or the Americas (0.36 [95% CI: 0.27-0.49]) compared with women from Africa (Table 1). Stratification of the African ICC cases revealed that 41.7% were from Mozambique, 30.1% from Nigeria, 19.6% from Uganda and 8.5% from Algeria (data not shown). Cases diagnosed between 1990-2009 showed 7.3% prevalence of multiple infections, 1.32-fold (95%CI: 1.08-1.60) higher than ICCs diagnosed between 1940-1989 (5.8% prevalence) (Table 1). However the difference in prevalence for multiple HPVs infections was not significant when stratifying ICCs into cases collected before and after 2000 (results not shown).

**HPV type-specific distribution**

HPV type-specific distribution in ICCs is presented in Figure 1. With 61.9% (N=5,432) prevalence, HPV16 was the most frequent HPV, followed HPV18 (11.1% prevalence; N=977), and HPV45 (6.7% prevalence; N=585). Their respective contributions in co-infected ICCs were also the highest: HPV16, 58.4% (N=343), HPV18, 21.0% (N=123); and HPV45, 18.9% (N=111). According to viral taxonomy, the Alpha-9 HPVs (which include HPV16) were involved in 86.5% (N=506) of the co-infected ICCs, while Alpha-7...
HPVs (which include HPV18) were involved in 44.6% (N=262) of the co-infected ICC cases (see also Figure 1).

Prevalence of single and multiple HPVs infections in ICC cases along with the patient characteristics, stratified by HR- and LR-HPVs, is given in Supplementary Table 1. Among ICCs with multiple infections (n=587), 76.0% were exclusively infected with HR-HPVs while only two cases were exclusively infected with LR-HPVs (HPV54|74 and HPV6|53|66). The remaining 23.7% of the multiple infected ICC cases showed a combination of HR- and LR-HPVs. The corresponding values for ICCs with single infections were 99.5% (N=8,149) for only HR-HPVs and 0.5% (N=44) for only LR-HPVs (Supplementary Table 1). Altogether, six (HPV16/18/31/33/35/45) and seven (HPV16/18/31/33/45/52/58) HR-HPVs covered/cross-protected by the 2- and 9-valent HPV vaccines, accounted for respectively 87.9% and 90.9% of the single-infected ICC cases and were present in 92.0% and 95.9% of the multiple HPV co-infected ICCs.

Prevalence of HPVs in single and multiple HPVs infections as well as S/M prevalence ratio are shown in Supplementary Table 2. Virtually all HR-HPVs, except HPV66 and HPV70, showed S/M prevalence ratios above one indicating a higher prevalence in single than in multiple infections. In contrast, all LR-HPVs (except HPV42) and particularly Alpha-10 LR-HPVs showed S/M prevalence ratios below one, indicating a higher prevalence in multiple than in single infections.

**HPV co-infection pattern analysis**

Among all HPVs pairwise combinations detected in ICCs with multiple infections, the prevalence of eight HPV pairs was significantly different than expected by chance: HPV16|74, HPV31|33, HPV31|44, HPV33|44 and HPV45|70 pairs were more likely to
be detected together than with any other HPV type, while HPV16|31, HPV16|45 and HPV16|51 pairs were less likely to be detected together compared with other HPVs (Figure 2A). Restricting analysis to ICCs infected only with two HPVs revealed that the HPV45|70 combination was more likely and HPV16|45 and HPV18|45 less likely to be detected together compared with any other combination of HPV types (Figure 2B).

To explore individual viral prevalence differences in ICCs with multiple HPVs infections beyond pairwise analysis, we performed a multivariate analysis. A significant difference in HPVs composition by geographic origin [likelihood ratio test (LRT) 201.24, \( P = 0.001 \)], by histology (LRT 58.604, \( P = 0.001 \)) and by period of diagnosis (LRT 43.35, \( P = 0.008 \)) was observed (Table 2). A univariate test applied to the multivariate analysis showed a significant virus-specific level difference by histology for HPV31 and by period of diagnosis for HPV74 (Table 2). Furthermore, seven HR-HPVs (HPV18, HPV35, HPV51, HPV52, HPV56, HPV58 and HPV74; Table 2) showed significant or nearly significant co-infection prevalence differences by geography. Accordingly, crude prevalence estimates of the above listed seven HR-HPVs in co-infected ICCs from Africa were either about 2-fold higher, or at least 5-fold lower than in any other world region (Table 2). In agreement, stratifying by histology and restricting the multivariate analysis on the SCC cases showed significant HPV co-infection pattern differences by period of diagnosis (LRT 41.60, \( P = 0.012 \)) and by geographic origin (LRT 193.6, \( P = 0.001 \)). Univariate test revealed also significant prevalence differences by period of diagnosis for HPV74 and by geography for five HR-HPVs (Supplementary Table 3). HPV type-specific networks shown in Figure 3 stratified by geographic origin visualize the multiple HPVs co-infection combinations observed and more importantly, the significant higher frequency of HPV18, HPV35, HPV51 and HPV52 present in multiple
infections in samples from Africa compared with samples outside Africa. The HPV pairs identified with significantly different pairwise co-infections by logistic regression are also visualized in Figure 3.

Finally, two different recursive partitioning analyses were performed to explore the appearance of HPVs co-infections in ICCs with either HPV16 (Figure 4) or HPV18 (Supplementary Figure 2) and accounting for age at diagnosis, geographic origin and period of diagnosis of the cases. First CART analysis searched for a homogenous split of the data on the chosen variable (i.e. HPV16) and the subsequent best splits among all variables that partitioned the data in each step into further most homogenous subgroups, accounting also for age at diagnosis, period of diagnosis and geographic origin of the cases. The predicted probabilities (PP) for each split for other HPVs occurring with HPV16 (PP= 58.4% [95%CI: 54.3-62.5], Figure 4) or HPV18 (PP= 21.0%, Supplementary Figure 2) were presented. Second and third most likely splits suggest respectively a major exclusion of HPV31 (PP= 62.6% [95%CI: 58.3-66.9]) and of HPV45 (PP= 67.9% [95% CI: 63.1-72.5]) from HPV16-related co-infections. In parallel, the two major exclusion variables from HPV18 co-infections and their predicted probabilities were non-African cases (PP= 39.6%) and HPV45 co-infection exclusion (PP= 47.4%) (Supplementary Figure 2).

DISCUSSION

In this study, we describe distinct clustering patterns of different HPVs detected in multiple HPVs infections in the largest available global collection of ICCs in unvaccinated women. Our results confirm that the simultaneous detection of different HPVs in ICCs is rare albeit far from negligible (6.7% in our dataset, N=8780). The
association patterns identified suggest further that the clustering of multiple HPVs in ICCs is not random. Our most relevant findings were a significant causal/coincidental infection of LR-HPV types 44 or 74 in ICCs in combination with HR-HPV types 16, 31 or 33, and a marked geographic clustering of particular oncogenic HPV types 51, 35, 18 and 52 (in descending significance) in multiple HPVs infected cervical cancers in African women compared with women outside Africa. A number of studies have previously communicated that clustering patterns observed among HPVs do not always occur at random in multiple infections detected in normal cytology, in cervical pre-cancerous lesions or in cancer lesions\textsuperscript{7,18–22}. Nevertheless, our results disagree with the previous studies, which suggest that the observed non-random HPVs clustering in multiple HPVs infections is mainly due to technical biases originating from the HPV genotyping methods\textsuperscript{23}.

The existence of genuine non-random association pattern of HPVs in multiple infections is elusive, and identifying and excluding technology-driven biases is not a simple task. Nevertheless, a public health concern is that HPVs co-infections may have a role in ICCs development not only among unvaccinated\textsuperscript{24} but also among HPV vaccinated women\textsuperscript{25}. The rational is that due to ecological competition between HPV types, the HPV vaccine implementation may further clear the ecological niche for the non-vaccine targeted oncogenic HPVs\textsuperscript{25–27}. Particularly, some HPVs such as HPV39 and HPV51 not included in the vaccine formulations have previously shown increase in prevalence after vaccination suggesting the possibility of type replacement\textsuperscript{25,26}. However, thus far, competition between HPV types, which is the prerequisite for HPV
type replacement, has not been persistently observed among unvaccinated \textsuperscript{7,18,28,29} or vaccinated populations \textsuperscript{25–27}.

**HPVs co-infection prevalence**

Here in this study we describe known as well as novel patterns of non-random HPVs clustering in global ICC cases, identified using a single well established HPV detection and genotyping technique. For clarity, it is important to remind that the current consensus of ICC etiology follows the logic of superinfection exclusion \textsuperscript{30}, with a subsequent clonal origin of HPV-induced lesions indicating that a single HPV virus contributes to the development of a single ICC lesion \textsuperscript{12,13}. This is to say, that the presence of DNA from two or more HPV types in an ICC sample does not necessarily mean that the different HPVs detected, actually contribute to the cancer development. Previously, techniques of laser capture microdissection (LCM) have been successfully used to affirm the likely involvement of causal or coincidental HPVs with particular lesion component in multiple HPVs infected cases, further supporting the hypothesis of clonal origin of HPV-induced cervical cancers \textsuperscript{13,31}.

In ICC cases worldwide, we did not observe significant differences in HPVs co-infection prevalence between the age groups at diagnosis. These results seem to be in contrast to the studies reporting higher prevalence of multiple HPVs infections in sexually active adolescents and in young women \textsuperscript{19,24,32}, and thus, our results might actually indicate that regardless of the sexual risk-taking behavior co-infection by multiple HPVs may be a risk factor for malignant cervical transformations. Interestingly however, a recent study suggests reduced HPV-driven carcinogenesis in cases of co-infection with
oncogenic HPV types from the same papillomavirus species among HIV-infected patients. This was interpreted as an evidence of natural cross-protection against HPV-induced carcinogenesis by multiple HPV infections from the same HPV species group. Nonetheless, we did observe that HPVs co-infection prevalence was consistently about three-fold higher in ICCs from Africa compared to the rest of the world. One explanation could be concomitant HIV infection which may increase the prevalence of multiple HPVs co-infections. Information on the sexual risk-taking behavior or HIV status was, unfortunately, not available for our ICC cases but the corresponding population estimates by Bruni and colleagues showed a relatively high prevalence of HIV infections among women in Mozambique, Uganda and Nigeria (11.3%, 7.2 and 3.7%, respectively). Thus, the larger fraction of women infected by HIV in these African populations compared with other populations worldwide may partly play a role in the higher presence of multiple HPVs infections also in our African ICC cases.

DNA degradation

Despite using a highly sensitive PCR based method for HPV genotyping (based on a 65 base pairs amplicon), we detected a minor 1.3-fold higher prevalence of co-infections in the more recently diagnosed ICC cases (between 1990-2009) than in older cases (between 1940-1989), but the difference was not significant when ICCs were stratified for cases collected before and after 2000 (results not shown), and these results were further supported by our multivariate analysis showing significance only for the 1990 stratification. Essentially, most paraffin embedded samples are initially fixed in formalin-containing solution, which reacts with the lateral amino groups in proteins.
and DNA, and thus, complicates the DNA isolation from FFPE samples. More importantly, the aldehyde in the formalin is unstable if unbuffered (used in most historical FFPE samples at least before the 1990s), and disproportionate rendering methanol and formic acid, causing extra fragmentation of the sample DNA. In short, we cannot completely rule out the possible effect of additional DNA degradation stemming from older tissue-fixation protocols, as described by Gilbert and colleagues. Notwithstanding, we estimate that older tissue-fixation protocol related DNA degradation plays a minor role (e.g. lowest LRT estimated for period of diagnosis) in explaining the higher prevalence of multiple HPVs infections in cases collected since 1990 compared with the older ICC cases.

**HPV type-specific clustering**

At pairwise association level, we found a significant causal or coincidental infection of LR-HPV types 44 or 74 in combination with HR-HPV types 16, 31 or 33 in global ICCs. Moreover, all Alpha-10 LR-HPVs (HPV6/11/44/74) showed a markedly higher prevalence in multiple HPV type infections compared to single HPV type infected ICCs. Altogether, we identified five positive and four negative significant associations between HPV types among the multiple infected ICCs. Three of the five positive associations were novel (i.e. HPV16|74, HPV31|44 and HPV33|44) while association between Alpha-9 species HPV31 and HPV33 or between Alpha-7 species HPV45 and HPV70 had also been previously estimated. Current epidemiological research supports the double etiological origin for HPV-induced cervical lesions: LR-HPVs, mainly HPV6 and HPV11, mostly leading to benign low-grade squamous intraepithelial lesions (LSILs), and oncogenic HR-HPVs, if persistent, mostly leading to high-grade SILs,
and considered cervical cancer precursors lesions. Nonetheless, low-risk HPVs such as HPV6, 11 and 44 from the Alpha-10 species have previously been occasionally associated with causal and coincidental infections in anogenital cancers. In addition, previous studies have also shown that a fraction of low-grade cervical SILs may also harbor components of HSIL or carcinoma often associated with independent HPVs infections, thus following the logic of “one-virus one-lesion”. In agreement, our results match well the logic of an essentially clonal origin of HPV-related lesions and support the occasional causal and/or coincidental role of low-risk Alpha-10 HPVs in cervical cancer development.

Theoretical models have predicted natural competition between HPV types, such as between types from the same HPV species group, and here we observed four significant negative pairwise HPV type associations: HPV16|HPV31 and HPV18|HPV45 pairs within the same species, and HPV16|HPV45 and HPV16|HPV51 pairs with longer pairwise phylogenetic distance between the types. Furthermore, our multivariate CART analysis beyond pairwise HPVs estimation revealed the same patterns of negative association between three of the HPV pairs, and thus, sustained the idea of possible competition or antagonistic interaction between these types. Novel negative association between HPV16|HPV45 types and two also previously reported negative associations, namely HPV16|HPV31 and HPV18|HPV45 pairs within the same species group. Notably, antagonistic interactions between HPV6/11 and HPV16, and between HPV6 and HPV31 types, earlier infections with the LR-HPVs protecting against cervical cancer risk associated with the latter HR-HPV infections have been documented, and indirectly fit with our observations on less likely co-infections.
with HPV16 and 31, and HPV18 and 45. Taken together, the assumption that HPV types within the same species group, such as Alpha-7 (HPV18, 39, 45, 59, 68, 70 and 85) or Alpha-9 (HPV16, 31, 33, 35, 52, 58 and 67), compete with one another is not clear as we also observed positive pairwise HPV type associations for HPV31|33 and HPV45|70 pairs.

Our most relevant finding is the distinct geographic clustering of HPVs multiple infections in ICCs among women from Africa compared to women outside Africa. Particularly, a significant geographic clustering of oncogenic HPV types 51, 35, 18 and 52 in multiple infections was observed in African women. That is, beyond pairwise level association and exclusively among the HPVs co-infected ICCs, we identify for the first time seven (i.e. HPV51, HPV35, HPV18, HPV74, HPV58, HPV52, HPV56) geographic clustering patterns for specific HPVs, in descending significance. Earlier studies regarding worldwide HPV genotype attribution in ICCs have described particular type-specific geographic structure, namely higher prevalence for HPV16, HPV18 and HPV45 in African populations, and a higher prevalence for HPV52 and HPV58 types in Asian populations compared with other world regions. In agreement, our multivariate analysis among the co-infected ICCs or only among the co-infected SCC cases demonstrated the significantly higher prevalence clustering of HPV18 in Africa and HPV58 in Asia. In contrast, HPV52 showed marginally significant higher geographic clustering among African women with intermediate prevalence in Asia while HPV16 or HPV45 co-infection prevalence values did not reveal any geographical structure. Most interestingly the observed patterns of geographic clustering of HPVs among the co-infected ICCs practically originate from differences between Africa and other world regions.
regions. As such, the prevalence of HPV51, HPV35, HPV18 and HPV52 types were about two-fold higher in African multiple infected ICCs, while prevalence of HPV56, HPV58 and HPV74 was either absent or at least two-fold lower in African ICCs compared with any other world region. Most importantly, the highest LRT values observed in our multivariate analysis were both the geographic clustering of multiple infections in general and at type level the geographic clustering of oncogenic HPV51, HPV35, HPV18 and HPV52 types in African multiple infected ICCs. Simultaneously, a recent community-randomized clinical trial data from Finnish population showed an increase of HPV51 and HPV52 in multiple HPVs infected women post-vaccinated47. However, within a public health perspective earlier studies have shown that although HPV51 is classified as possibly/probably oncogenic by the IARC, it presents a low oncogenic potential to cause ICCs and is commonly more prevalent in low-grade cervical lesions48.

In addition, a minor clustering of HPVs by histological type was also observed among the HPV co-infected ICCs. A significantly higher contribution of HPV31 was observed in SCCs compared with other type of ICCs, mostly adenocarcinomas (ADC) or mixed adenosquamous carcinomas (ADSC). Previous studies have already reported these HPV type distribution differences between single infected cervical SCCs and ADCs/ADSCs 5,49, and our results among multiple HPV co-infected ICCs further sustain the results.

Nevertheless, and despite the large worldwide cross-sectional sampling of ICC cases and the standardized HPV genotyping methodology used, our study suffers from certain limitations. Firstly, detection of more than one HPV type or variant in a sample
is not enough to assign causality of the multiple HPVs with the particular lesion, in this case, the ICC. Secondly, further cohort studies and use of other biomarkers providing biological activity of particular HPVs such as mRNA or p16INK4a expression, and with new statistical models are needed to better disentangle the role of multiple HPVs infections in the etiology of cervical cancer, especially in the post-HPV-vaccinated environment. Thirdly, a better characterization of the patients’ genetic background, possible HIV infections, heterogeneous ICCs biopsy material and detection method cross-validation is needed to confirm the potential role of multiple HPVs infections in ICCs. Lastly, the detection and ultimately the likely role of multiple HPVs infections at HPV variant level has only recently been verified thanks to the next-generation sequencing applied to HPV genotyping, but within host HPV type-level variation is beyond the scope of this study.

Our genotyping strategy was based on the differential hybridization of very short PCR products, and as such, we cannot completely rule out the possibility that some of the observed clustering of multiple HPV type co-infections in our study is not partly due to PCR and hybridization biases and directional cross-detection of multiple HPVs. However, several lines of evidence argue against that the possible technology-based cross-detection biases are the main explanation for the HPVs clustering patterns observed. First, it has been previously shown that the SPF10-PCR LiPA25 methodology is more sensitive and less prone to overemphasize multiple HPVs infections than other PCR-based methods. Second, it is unlikely that the observed differential HPVs clustering after histological or geographical stratification exclusively originates from unspecific genotyping, as it would imply that such hybridization biases are both histology and geography dependent. Third, we found the worldwide non-random
clustering of different HR-HPVs in co-infected ICCs at two different analysis depths, at pairwise and at multivariate level as described above.

Taken together, HPVs co-infection patterns are rare in ICCs, and thus difficult to study at population-level. In this study, however, we provide robust and worldwide estimations on the proportion of single and multiple HPV type infected ICCs and reveal HPVs pairwise combinations significantly deviating from the null expectation. Further, we identified for the first time beyond pairwise analysis a distinct co-infection pattern in African ICCs compared to any other world region. Our results suggest that this differential clustering of HPV type combinations in ICCs after stratifying by histological presentation and geographical origin might be compatible with a specific modulation of the natural history of the infection and/or of the oncogenic potential of particular HPVs multiple infections. Finally, although our results may be valuable in estimating competition between non-vaccine HPV types and warrant monitoring for post-vaccinated, it should be emphasized that clustering of multiple HPVs infections are marginal in ICCs with likely only a minor public health importance, and mostly among high sexual risk-taking behavior population groups.

Author Contributions

Conceived the project and wrote the manuscript: VNP and ST. Performed the analysis: VNP, ST, YB. Contributed in data generation, study design and data interpretation: VNP, ST, YB, LA, NG, WQ, XFB, IGB, SS. Supervised the project: SS. All authors assisted in the writing of the manuscript and had final approval of the submitted and published versions.

This article is protected by copyright. All rights reserved.
Acknowledgments

We thank Jo Ellen Klaustermeier for excellent technical assistance. The study is part of the ICO SURVEY RIS HPVTT and has been partly supported by Spanish public grants from the Instituto de Salud Carlos III (grants numbers FIS PI030240, FIS PI061246, RCESP C03/09, RTICESP C03/10, RTIC RD06/0020/0095, and CIBERESP), Agencia de Gestio d’Ajuts Universitaris i de Recerca (AGAUR 2005SGR 00695 and 2009SGR126), Marato de TV3 Foundation (051530), The European Society of Clinical Microbiology and Infectious Diseases (ESCMID Research Grant 2015), Spanish Ministry of Health (PERIS SLT002/16/00496 and FIS PI17/00092) and GlaxoSmithKline Biologicals, Sanofi Pasteur MSD, and Merck. None of these agencies had any role in the interpretation of the results or in the preparation of this manuscript.

Potential conflict of interest

Laia Alemany and Cancer Epidemiology Research Program has received sponsorship and grants from Merck. Xavier F. Bosch has received occasional personal and travel grants from Merck, Hologic, Qiagen, Sanofi Pasteur MSD and Seegene.

References


3. Einstein MH, Schiller JT, Viscidi RP, Strickler HD, Coursaget P, Tan T, Halsey N,


9. Dickson EL, Vogel RI, Geller MA, Downs LS. Cervical cytology and multiple type


16. Wang Y, Naumann U, Wright ST, Warton DI. mvabund - an R package for model-


42. Xi LF, Edelstein ZR, Meyers C, Ho J, Cherne SL, Schiffman M. Human papillomavirus types 16 and 18 DNA load in relation to coexistence of other types, particularly those in the same species. *Cancer Epidemiol Biomarkers Prev*


49. Li N, Franceschi S, Howell-Jones R, Snijders PJF, Clifford GM. Human


Table 1. Characteristics of the ICC cases included in the study.

<table>
<thead>
<tr>
<th>ICCs</th>
<th>HPV positive</th>
<th>Multiple HPVs</th>
<th>Crude OR [95%CI]</th>
<th>Adjusted OR [95%CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at diagnosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>51.0 (13.17)</td>
<td>51.6 (13.8)</td>
<td>1.00 [0.996-1.01]</td>
<td></td>
</tr>
<tr>
<td>Median (Min-Max)</td>
<td>50 (17-97)</td>
<td>51 (24-93)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANOVA P-value</td>
<td></td>
<td></td>
<td>0.318</td>
<td></td>
</tr>
<tr>
<td>LRT P-value</td>
<td></td>
<td></td>
<td>0.315</td>
<td></td>
</tr>
<tr>
<td>Age at diagnosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65+</td>
<td>1,226 (14.0)</td>
<td>73 (6.0)</td>
<td>Ref.</td>
<td>Ref.</td>
</tr>
<tr>
<td>55-64</td>
<td>1,501 (17.1)</td>
<td>102 (6.8)</td>
<td>1.15 [0.84-1.57]</td>
<td>1.17 [0.86-1.60]</td>
</tr>
<tr>
<td>45-54</td>
<td>2,123 (24.2)</td>
<td>119 (5.6)</td>
<td>0.94 [0.69-1.27]</td>
<td>0.96 [0.71-1.30]</td>
</tr>
<tr>
<td>35-44</td>
<td>1,806 (20.6)</td>
<td>88 (4.9)</td>
<td>0.81 [0.59-1.11]</td>
<td>0.83 [0.61-1.15]</td>
</tr>
<tr>
<td>25-34</td>
<td>635 (7.2)</td>
<td>40 (6.3)</td>
<td>1.06 [0.71-1.58]</td>
<td>1.14 [0.76-1.70]</td>
</tr>
<tr>
<td>&lt;=24</td>
<td>48 (0.5)</td>
<td>5 (10.4)</td>
<td>1.84 [0.71-4.78]</td>
<td>1.96 [0.75-5.13]</td>
</tr>
<tr>
<td>Chi-square P-value</td>
<td></td>
<td></td>
<td>0.162</td>
<td></td>
</tr>
<tr>
<td>Trend test P-value</td>
<td></td>
<td></td>
<td>0.365</td>
<td>0.435</td>
</tr>
<tr>
<td>LRT P-value</td>
<td></td>
<td></td>
<td>0.180</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Missing</td>
<td>1,441</td>
<td>160</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Geographic origin

<table>
<thead>
<tr>
<th>Region</th>
<th>ICCs</th>
<th>(%)</th>
<th>HPV positive</th>
<th>(%)</th>
<th>Multiple HPVs</th>
<th>(%)</th>
<th>Crude OR [95%CI]</th>
<th>Adjusted OR [95%CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>525</td>
<td>(6.0)</td>
<td>101 (19.2)</td>
<td>Ref.</td>
<td>Ref.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Europe</td>
<td>2,031</td>
<td>(23.1)</td>
<td>134 (6.6)</td>
<td>0.30</td>
<td>[0.22-0.39]</td>
<td>0.36</td>
<td>[0.27-0.49]</td>
<td></td>
</tr>
<tr>
<td>Asia</td>
<td>2,751</td>
<td>(31.3)</td>
<td>147 (5.3)</td>
<td>0.24</td>
<td>[0.18-0.31]</td>
<td>0.31</td>
<td>[0.23-0.42]</td>
<td></td>
</tr>
<tr>
<td>America</td>
<td>3,473</td>
<td>(39.6)</td>
<td>205 (5.9)</td>
<td>0.26</td>
<td>[0.20-0.34]</td>
<td>0.36</td>
<td>[0.27-0.49]</td>
<td></td>
</tr>
<tr>
<td>Chi-square P-value</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LRT P-value</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Period of diagnosis

<table>
<thead>
<tr>
<th>Period</th>
<th>ICCs</th>
<th>(%)</th>
<th>HPV positive</th>
<th>(%)</th>
<th>Multiple HPVs</th>
<th>(%)</th>
<th>Crude OR [95%CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1940 - 1989</td>
<td>3,513</td>
<td>(40.0)</td>
<td>202 (5.8)</td>
<td>Ref.</td>
<td>Ref.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1990 - 2009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>-------------</td>
<td>-------</td>
<td>-------</td>
<td>-----</td>
<td>-------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5,267</td>
<td>(60.0)</td>
<td>385</td>
<td>(7.3)</td>
<td>1.29</td>
<td>[1.08-1.54]</td>
<td></td>
</tr>
<tr>
<td>Chi-square P-value</td>
<td>0.296</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LRT P-value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.004</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td><strong>Histology</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>8,064</td>
<td>(91.8)</td>
<td>530</td>
<td>(6.6)</td>
<td>Ref.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chi-squared test P-value</td>
<td>0.154</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LRT P-value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.154</td>
<td>0.165</td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>8,780</td>
<td>(100.0)</td>
<td>587</td>
<td>(6.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: ICC, Invasive cervical cancer; HPV, Human Papillomavirus; Crude, crude odds ratio (OR); adjusted, ORs adjusted by age at diagnosis, period of diagnosis and geographic origin; N, Number of HPV/DNA-positive cases, SD: standard deviation; LRT: log-likelihood ratio test.

1 HPV positive by PCR with SPF-10 primers followed by DEIA and LIPA25 genotyping of 25 HPV types (HPV6, 11, 16, 18, 31, 33, 34, 35, 39, 40, 42, 43, 44, 45, 51, 52, 53, 54, 56, 58, 66, 68, 70, and 74).

2 Column percent.

3 Row percent of multiple infections among HPV/DNA-positive ICC cases.

4 Asia including Oceania.

5 Diagnosis of adenocarcinoma, adenosquamous carcinoma and to a lesser extent undifferentiated and neuroendocrine tumors.
Table 2. Multivariate analysis of all 25 HPV types available among co-infected ICCs, and HPV type specific prevalence distribution for eight HPV types showing significant clustering by histology, period of diagnosis or geographic origin of the samples.

<table>
<thead>
<tr>
<th>ICCs</th>
<th>Histology</th>
<th>Period of diagnosis</th>
<th>Geographic origin</th>
<th>Prevalence among HPV co-infected ICCs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LRT</td>
<td>P-value</td>
<td>LRT</td>
<td>P-value</td>
</tr>
<tr>
<td>58.6</td>
<td>0.001</td>
<td>43.35</td>
<td>0.008</td>
<td>201.24</td>
</tr>
<tr>
<td>29.7</td>
<td>0.001</td>
<td>43.35</td>
<td>0.008</td>
<td>201.24</td>
</tr>
<tr>
<td>18.8</td>
<td>0.003</td>
<td>43.35</td>
<td>0.008</td>
<td>201.24</td>
</tr>
<tr>
<td>39.6</td>
<td>0.003</td>
<td>43.35</td>
<td>0.008</td>
<td>201.24</td>
</tr>
<tr>
<td>1.0</td>
<td>0.006</td>
<td>43.35</td>
<td>0.008</td>
<td>201.24</td>
</tr>
<tr>
<td>2.0</td>
<td>0.006</td>
<td>43.35</td>
<td>0.008</td>
<td>201.24</td>
</tr>
<tr>
<td>12.2</td>
<td>0.006</td>
<td>43.35</td>
<td>0.008</td>
<td>201.24</td>
</tr>
<tr>
<td>13.2</td>
<td>0.006</td>
<td>43.35</td>
<td>0.008</td>
<td>201.24</td>
</tr>
<tr>
<td>0.0</td>
<td>0.006</td>
<td>43.35</td>
<td>0.008</td>
<td>201.24</td>
</tr>
<tr>
<td>0.25</td>
<td>0.006</td>
<td>43.35</td>
<td>0.008</td>
<td>201.24</td>
</tr>
<tr>
<td>30.7</td>
<td>0.006</td>
<td>43.35</td>
<td>0.008</td>
<td>201.24</td>
</tr>
<tr>
<td>8.2</td>
<td>0.006</td>
<td>43.35</td>
<td>0.008</td>
<td>201.24</td>
</tr>
<tr>
<td>5.9</td>
<td>0.006</td>
<td>43.35</td>
<td>0.008</td>
<td>201.24</td>
</tr>
</tbody>
</table>

Abbreviations: ICC, Invasive cervical cancer; Asia¹, Asia including Oceania; LRT: log-likelihood ratio test; Underlined values indicate significant or marginally significant P-values; Bolded values indicate the highest geographic prevalence values observed. HPV types in descending order regarding the LRT values by geographic origin.
Pairwise HPV-type combinations
(Adjusted analysis - All infections)

Pairwise HPV-type combinations
(Adjusted analysis - Only infections with 2 HPV types)
N (total)=587
PP(HPV16) 58.4%
[54.3-62.5]

HPV31-positive
PP 32.1% [22.2-43.4]
N=81

HPV31-negative
PP 62.6% [58.3-66.9]
N=506

HPV45-positive
PP 42.3% [32.7-52.4]
N=104

HPV45-negative
PP 67.9% [63.1-72.5]
N=402

HPV16 positive cases