Available online at www.sciencedirect.com

SCIENCE DIRECT

Virology 337 (2005) 76 – 84

VIROLOGY

www.elsevier.com/locate/yviro

The carcinogenicity of human papillomavirus types reflects viral evolution

Mark Schiffman[a,*], Rolando Herrero[b], Rob DeSalle[c], Allan Hildesheim[a], Sholom Wacholder[a], Ana Cecilia Rodriguez[b], Maria C. Bratti[b], Mark E. Sherman[a], Jorge Morales[b], Diego Guillen[b], Mario Alfaro[b], Martha Hutchinson[d], Thomas C. Wright[e], Diane Solomon[a], Zigui Chen[f], John Schussler[g], Philip E. Castle[a], Robert D. Burk[f]

[a]*Division of Cancer Epidemiology and Genetics or Division of Cancer Prevention, National Cancer Institute, National Institutes of Health, U.S. Department of Health and Human Services, Bethesda, MA, USA*
[b]*Proyecto Epidemiológico Guanacaste, San José, Costa Rica*
[c]*Division of Invertebrate Zoology, American Museum of Natural History, New York, NY, USA*
[d]*Women and Infants' Hospital, Providence, RI, USA*
[e]*College of Physicians and Surgeons of Columbia University, New York, NY, USA*
[f]*Albert Einstein Cancer Center, Albert Einstein College of Medicine, Bronx, NY, USA*
[g]*Information Management Services, Silver Spring, MA, USA*

Received 17 February 2005; returned to author for revision 22 February 2005; accepted 1 April 2005
Available online 3 May 2005

Abstract

Persistent infections with carcinogenic human papillomaviruses (HPV) cause virtually all cervical cancers. Cervical HPV types (n > 40) also represent the most common sexually transmitted agents, and most infections clear in 1–2 years. The risks of persistence and neoplastic progression to cancer and its histologic precursor, cervical intraepithelial neoplasia grade 3 (CIN3), differ markedly by HPV type. To study type-specific HPV natural history, we conducted a 10,000-woman, population-based prospective study of HPV infections and CIN3/cancer in Guanacaste, Costa Rica. By studying large numbers of women, we wished to separate viral persistence from neoplastic progression. We observed a strong concordance of newly-revised HPV evolutionary groupings with the separate risks of persistence and progression to CIN3/cancer. HPV16 was uniquely likely both to persist and to cause neoplastic progression when it persisted, making it a remarkably powerful human carcinogen that merits separate clinical consideration. Specifically, 19.9% of HPV16-infected women were diagnosed with CIN3/cancer at enrollment or during the five-year follow-up. Other carcinogenic types, many related to HPV16, were not particularly persistent but could cause neoplastic progression, at lower rates than HPV16, if they did persist. Some low-risk types were persistent but, nevertheless, virtually never caused CIN3. Therefore, carcinogenicity is not strictly a function of persistence. Separately, we noted that the carcinogenic HPV types code for an E5 protein, whereas most low-risk types either lack a definable homologous E5 ORF and/or a translation start codon for E5. These results present several clear clues and research directions in our ongoing efforts to understand HPV carcinogenesis.
Published by Elsevier Inc.

Introduction

Laboratory and epidemiologic data suggest that persistent infections with carcinogenic human papillomaviruses (HPV) cause virtually all cervical cancers and substantial fractions of other anogenital cancers worldwide, totaling half a million anogenital cancers annually (Bosch and de Sanjose, 2003; Bosch et al., 2002; zur Hausen, 2000). In

* Corresponding author. Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, U.S. Department of Health and Human Services, Room 7066, 6120 Executive Boulevard, Rockville, MD 20852, USA. Fax: +1 301 402 0916.
E-mail address: schiffmm@mail.nih.gov (M. Schiffman).

0042-6822/$ - see front matter. Published by Elsevier Inc.
doi:10.1016/j.virol.2005.04.002

aggregate, cervical HPV types ($n > 40$) also represent the most common sexually transmitted agents, and most infections clear without sequelae in 1–2 years (Ho et al., 1998; Molano et al., 2003; Richardson et al., 2003; Woodman et al., 2001). Based on available cross-sectional and short-term prospective data, the risks of persistence and neoplastic progression differ markedly by HPV type, with genetically-related types appearing to act most alike (Munoz et al., 2003). As follow-up data accumulate from long-term, large cohort studies, we can now explore more fully the natural history and prospective risks of the individual types, as a guide to future research and prevention efforts (Lowy and Frazer, 2003; Schiffman and Kjaer, 2003). Specifically, we conducted a 10,000-woman, population-based prospective study of HPV infection and cervical neoplasia in Guanacaste, Costa Rica (Bratti et al., 2004; Herrero et al., 1997, 2000; Schiffman et al., 2000). Here, we report the strong concordance of HPV phylogenetic analysis with viral natural history and carcinogenicity for the full range of anogenital HPV types.

Results

Fig. 1 shows the phylogeny of known genital HPV types derived using a Bayesian methodology, with notation of carcinogenic risk levels assigned by the large case-control study conducted by the International Agency for Research on Cancer (Munoz et al., 2003). This phylogenetic tree based upon full genome alignments (compare to (de Villiers et al., 2004; Van Ranst et al., 1992)) suggests that at least 3 ancestral papillomaviruses are responsible for the current heterogeneous group of genital HPV genomes. The tree shows a high degree of robustness (no nodes or "branchings" have <70% support) and agreement among Bayesian and the two maximum parsimony methods (Fig. 1). The three major groups that emerged include alpha papillomavirus species α10, α8, α1 and α13; α9, α11, α7, α5 and α6; and α4, α15, α3 and α2. Interestingly, established carcinogenic types as defined by the IARC case-control data are derived from a common ancestor. Also in Fig. 1, we noted that the species α9, α11 and α7 containing carcinogenic types can code for a homologous E5 protein, whereas the others either lack a definable E5 ORF and/or translation start codon for E5. As notable exceptions, the α10 types, which cause venereal warts, can also code for an E5 protein.

The population-based results of the Guanacaste Project (Fig. 2, explained in Materials and methods section) demonstrated wide variability in type-specific prevalence (Table 1). Notably, the more likely a viral type was to persist, the more prevalent it was in the population (Spearman correlation coefficient 0.46, $P = 0.005$). The order of the types in Table 1 matches the distribution of types in the phylogenetic tree in Fig. 1; thus, proximity in the table suggests phylogenetic relatedness. For example, at the top are listed the closely related types from species α10, HPV6 and HPV11, which are not associated with cervical cancer, but do produce venereal warts and laryngeal papillomatosis (Kashima et al., 1996). Of note, they are far less common at the cervix than is usually recognized.

Many HPV types, which share a common sexual route of transmission, were found as mixed infections with prevalent CIN3/cancer, resulting in generally elevated but confounded estimates of the fraction of CIN3/cancer attributable to each of them (Table 1). However, the hierarchically-adjusted fractions showed that prevalent cases could be largely attributed to HPV types within species α9 and α7. HPV16 had the highest prevalence and explained the greatest fraction of CIN3/cancer of any type. Additional PCR testing demonstrated HPV33 (α9) in 1 of 2 CIN3/cancer cases originally typed as HPV11 alone, and HPV16 in the case originally typed as HPV71 alone. We did not "correct" the data based on additional testing, to minimize the chance of biasing results toward our expectations.

With prevalent cases of CIN2, CIN3, and cancer excluded, the cohort was followed for viral persistence and incident neoplasia (CIN1 was considered a viral effect, not a neoplastic outcome.) Among the infections that did persist, the risk of CIN3/cancer diagnosed during prospective follow-up was convincingly elevated primarily for HPV16. One-third of women with persistent HPV16 infection had incident CIN3/cancer. Therefore, in addition to 42 (13.9%) HPV16-infected women with prevalent CIN3/cancer, 18 developed incident CIN3/cancer. In contrast, the risks of CIN3/cancer for other types in species α9 and for types in species α7, α5, and α6 were lower. We observed no cases of incident CIN3/cancer for types in species α3, α8, α10, α13, and α15. In other words, within the limits of the size of this study, the risk was not appreciably elevated compared to the risk in initially HPV-negative women. Of note, the population prevalences were as high for some of these low-risk types, including HPV61, HPV62, HPV81, HPV83 (α3) and HPV71 (α15) as for some of the types for which we did observe CIN3/cancer. Therefore, we can be relatively confident of the low risk posed by these common types that rarely if ever led to CIN3/cancer.

Fig. 3 graphically summarizes the persistence and progression data for the most prevalent species, with HPV16 shown separately from the rest of α9. Although CIN3 served as our stricter surrogate endpoint for cancer risk, this summary figure includes reviewed, histologic cases of CIN2 to provide a clinical perspective because CIN2 is usually treated. Persistence without CIN3/cancer was common in several species but only HPV16 and, secondarily, the other types in α9, and those in α7, α5, and α6 were at elevated risk of progression given persistence. In general, the same types that caused CIN3 and cancer were evident in cases of CIN2. Although the small numbers precluded statistical analysis, it is noteworthy that the types in species α7 were over-represented in follow-up invasive cancers. Of the 9 cancers diagnosed during follow-up (i.e.,

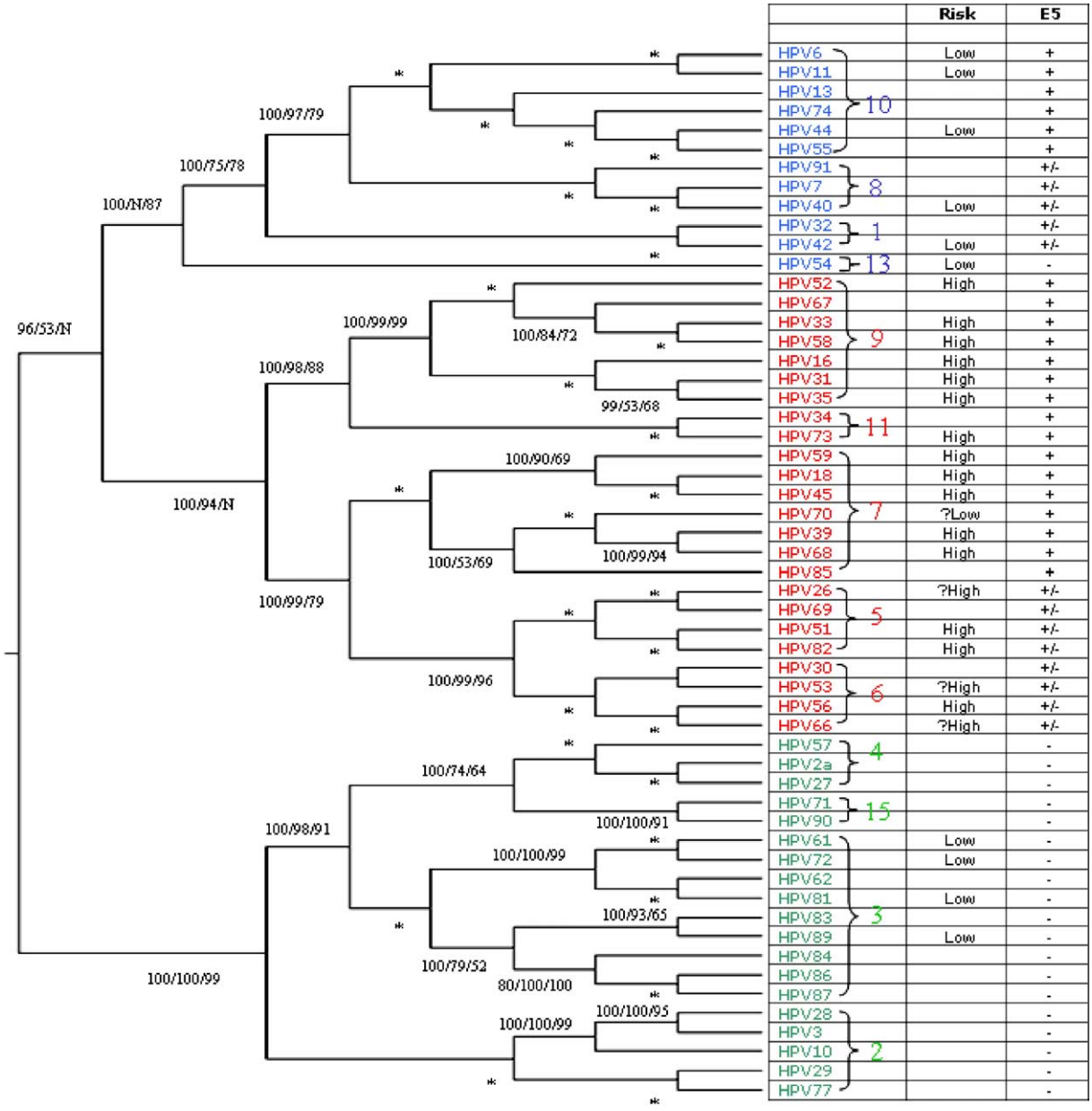

Fig. 1. Phylogenetic analysis of anogenital HPV types. Numbers on or near branches (nodes) indicate support indices from 100 bootstrap estimations using each of the methods in the following order: Bayesian credibility value, parsimony bootstrap percentage based on nucleotide alignment, and parsimony bootstrap percentage based on amino acid alignment. An asterisk indicates 100% agreement (all 100 estimations concordant for each method for that branching). N reflects disagreement between a method and the reference Bayesian tree at a given node. "Risk" refers to the cancer risk categorization assigned by the IARC case-control data (Munoz, et al., 2003); "?High" or "?Low" were assigned when the data were ambiguous. A homologous E5 gene is depicted as present (+) or absent (−). We indicated a value of (+/−) when an E5 ORF was identified but did not have a translation start codon present. The number in brackets indicate the α species of the types.

missed during the baseline screening), three showed persistent HPV18 (one demonstrated by repeated testing of the previously-negative enrollment specimen), one persistent HPV45, one HPV18 only at diagnosis, and three persistent HPV16. One tested HPV negative at both times.

In an ancillary analysis we divided the population into two age groups, <30 and 30+, in order to see whether age modified the trends we observed. HPV persistence rates increased generally with age, as detailed in a more extensive analysis published elsewhere (Castle et al., in press). However, the ancillary analysis showed that relative, inter-typic and inter-species differences in persistence and risk of CIN3 were not appreciably altered by broad age stratification.

Discussion

The combination of phylogenetic analysis and population-based prospective epidemiology provided interesting new insights. Phylogenetic grouping predicted the natural history and carcinogenicity of individual HPV types, and corroborated much of the recent cross-sectional data from

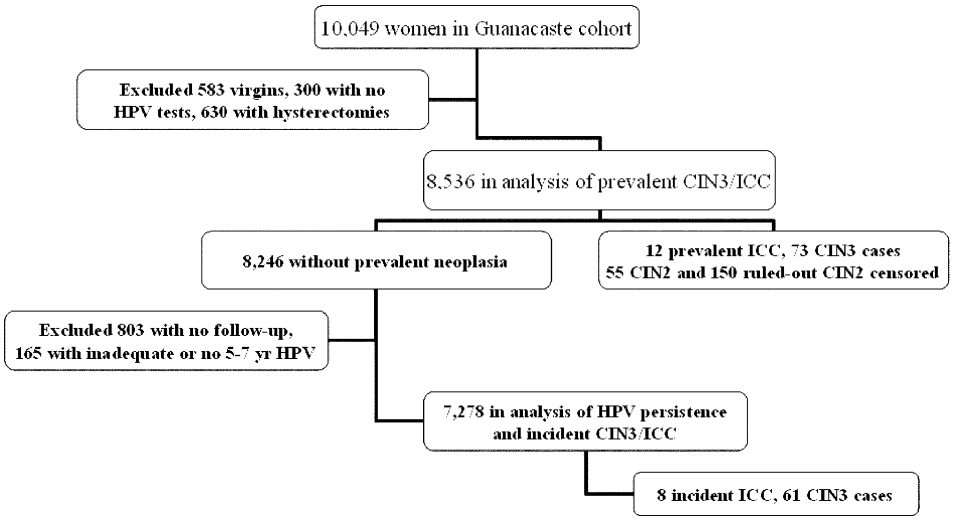

Fig. 2. Guanacaste study population. Of the 10,049 women enrolled in the Guanacaste cohort, we excluded 583 virgins, 300 women without enrollment HPV results, and 630 women with hysterectomies. Among 8536 women screened at enrollment, there were 55 CIN2, 73 CIN3 and 12 invasive cancer cases. In the prospective cohort, we excluded women without any follow-up PCR results (803 no follow-up, 165 inadequate or no PCR results for collected specimens). We also excluded 290 women with possible CIN 2 or worse at enrollment (including the 140 with confirmed, prevalent ≥CIN2). There were 60 cases of incident histologic CIN3 and 9 cases of cancer that occurred despite surveillance. For those censored during follow-up but prior to 5 years due to possible (even if not confirmed) incident CIN2, CIN3, or cancer (n = 165) or for any other reason (e.g., death from other causes, relocation, n = 179), we analyzed the last specimen prior to censoring and/or appropriate treatment.

case-control studies and case series of cervical cancer (Munoz et al., 2003). As the most important novel epidemiologic contributions, 1) we showed that population prevalence of individual types is correlated with viral persistence and 2) we distinguished viral persistence from neoplastic progression given viral persistence. Persistence and progression are highly associated (Ho et al., 1995; Nobbenhuis et al., 1999; Schlecht et al., 2001) and large numbers of subjects must be followed for several years to disentangle the two. We observed that the types and species with presumed low cancer risk defined *a priori* by their absence (i.e., they are not found alone) in cancer case series showed varying degrees of persistence, but did not progress to CIN3/cancer. In contrast, the most carcinogenic HPV types concentrated in species α9 and α7 were distinguished by elevated risk of progression given persistence, rather than persistence alone. We also showed clearly that the role of E5, which is a transforming protein in some papillomaviruses (DiMaio and Mattoon, 2001), deserves further study regarding its activities in infected cervical cells.

HPV16 was uniquely carcinogenic by all important standards of risk, i.e., attributable fraction of prevalent CIN3/cancer, probability of persistence, and probability of incident CIN3/cancer given persistence. It has been reported previously based on cross-sectional evidence that approximately 50% of cervical cancers, and an even higher percentage of non-cervical HPV-induced cancers (e.g., HPV-positive oropharyngeal cancers), are caused by HPV16 (Bosch and de Sanjose, 2003; Herrero et al., 2003). To this, we add that persistent HPV16 infection represents a carcinogen with a very high positive predictive value of serious neoplasia, deserving clinical evaluation and follow-up. Few other exposures cause precancer/cancer in approximately 20% of those exposed. HPV16 is now the primary target of HPV vaccine trials (Koutsky et al., 2002; Lowy and Frazer, 2003), and our data indicate it should be distinguished from other carcinogenic types in diagnostic kits, even if the other types are not individually characterized. In previous work, we and others have observed that most resolving HPV16 infections cleared by 1–2 years after enrollment (Ho et al., 1998; Richardson et al., 2003), suggesting a practical endpoint to surveillance of persistently infected women without evident lesions (yet). HPV18 also deserves special attention because of its established association with glandular lesions that are difficult to detect by cytology (Andersson et al., 2001), shown perhaps by the presence of HPV18 (and possibly HPV45, which is closely related) among the few missed follow-up cancers in our study.

Despite the large size of this cohort, the stability of estimates for the less common types was low and individual confidence intervals were very broad. Also, it is possible that some of the risk estimates for probability of persistence and probability of CIN3/cancer could have been affected by differential sensitivity of HPV testing method for different HPV types (Gravitt et al., 2000). Specifically, the primer set we used does not efficiently amplify some types, including HPV68, which showed low rates of persistence. Furthermore, our type-specific HPV-infected subcohorts were defined at enrollment, without distinction between new (incident) and already persistent infections. Perhaps, on average, some of the types had already persisted longer than others and were, therefore, at altered risk of subsequent extended persistence and progression to CIN3/cancer. In ongoing work, we are following newly-infected women, to gain a more complete

Table 1
Population-based HPV natural history and CIN3/cancer risk data from Guanacaste

Species (alpha)	Type	# of infected women	# Prevalent CIN3/cancer				# and % of persistence among women <CIN2 at enrollment[c]		# and % of CIN3/ICC among women wit persistent type-specific HPV[c,d]	
			#	AF[a]	$\#_{adj}$[b]	AF_{adj}[b]				
10	6	50	1	0.6%	0	0.0%	4	10.0%	0	0.0%
10	11	24	2	2.1%	2	2.3%	0	0.0%	N/A	
10	74v	15	0	0.0%	0	0.0%	1	7.7%	0	0.0%
10	55	20	0	0.0%	0	0.0%	4	22.2%	0	0.0%
8	40	12	0	0.0%	0	0.0%	0	0.0%	N/A	
1	32	29	1	0.8%	0	0.0%	3	12.0%	0	0.0%
13	54	37	3	3.1%	0	0.0%	5	17.2%	0	0.0%
9	52	135	6	5.6%	0	0.0%	13	13.1%	4	30.8%
9	67	14	0	0.0%	0	0.0%	0	0.0%	N/A	
9	33	59	4	4.0%	1	1.2%	13	27.7%	3	23.1%
9	58	170	12	12.4%	10	11.1%	20	16.5%	1	5.0%
9	16	302	42	47.6%	42	47.6%	56	28.9%	18	32.1%
9	31	121	10	10.5%	9	10.2%	13	14.9%	2	15.4%
9	35	36	1	0.8%	0	0.0%	2	6.9%	1	50.0%
11	73	38	0	0.0%	0	0.0%	1	3.7%	0	0.0%
7	59	31	2	2.0%	0	0.0%	2	8.3%	1	50.0%
7	18	111	10	10.6%	5	5.7%	13	15.7%	4	30.8%
7	45	66	3	2.8%	0	0.0%	5	9.4%	1	20.0%
7	70	177	3	1.5%	0	0.0%	22	16.1%	1	4.6%
7	39	69	2	1.6%	0	0.0%	8	15.1%	0	0.0%
7	68	29	1	0.8%	1	1.2%	1	4.4%	0	0.0%
7	85	59	2	1.7%	0	0.0%	5	10.9%	0	0.0%
5	26	16	1	1.0%	0	0.0%	0	0.0%	N/A	
5	51	166	6	5.2%	3	3.4%	6	4.9%	2	33.3%
5	82v	31	1	0.8%	0	0.0%	3	12.5%	0	0.0%
6	53	200	3	1.2%	0	0.0%	10	6.6%	2	20.0%
6	56	75	3	2.7%	3	3.4%	9	14.8%	1	11.1%
6	66	68	1	0.4%	0	0.0%	5	9.4%	0	0.0%
15	71	204	1	0.0%	1	1.1%	29	16.9%	0	0.0%
3	61	208	2	0.0%	0	0.0%	25	13.9%	0	0.0%
3	72	25	0	0.0%	0	0.0%	5	21.7%	0	0.0%
3	62	150	2	0.6%	0	0.0%	7	5.6%	0	0.0%
3	81	103	2	1.2%	0	0.0%	10	11.6%	0	0.0%
3	83	100	2	1.2%	0	0.0%	11	13.9%	0	0.0%
3	89	21	0	0.0%	0	0.0%	0	0.0%	N/A	
3	84	58	1	0.5%	0	0.0%	4	8.2%	0	0.0%
	? type	378	6	2.8%	6	6.0%		N/A	3	0.9%[e]
	Neg	6182	2	N/A	2	N/A		N/A	21	0.4%[e]

[a] The fraction of prevalent cases of CIN3/cancer attributable to an HPV type was calculated as AF = % cases positive for that type × (1 − 1/OR) where OR represents the odds ratio associating that type and prevalent risk of CIN3/cancer. OR for a type = (number of cases positive for that type divided by number of cases negative for that type) divided by (number of controls positive for that type divided by number of controls negative for that type). Two types (71 and 61) had slightly negative AF, which were set to zero.

[b] In the hierarchical analysis of attributable fraction, women with types accounting for a higher fraction of CIN3/cancer were excluded. For example, women with the type accounting for the highest fraction of CIN3/cancer (i.e., HPV16) were excluded to find the next most important type (i.e., HPV58), and women with either type were excluded to find the third, etc.

[c] Women with prevalent or incident CIN2 were excluded to provide diagnostic certainty, because they represented equivocal cases. As a result, the percentages in the table can not be exactly calculated directly from the counts of numbers of infections and outcomes.

[d] Four cases of CIN3 involved persistence of two types: 52/59, 16/33, 16/51, 51/70.

[e] Some cases of incident CIN3 arose without observed viral persistence of a defined type (i.e., they were HPV negative or had an undefined type at enrollment). In some cases, we possibly missed the period of persistence by infrequent (5–7 year interval) measurement.

understanding of the comparative natural histories of incident infections with different HPV types.

Despite these acknowledged limitations, the consistency of the genealogical inference represented by Fig. 1 with the patterns of natural history data was remarkable. The HPV genome is small enough (8 Kb) to permit a comprehensive analysis of all its components and functions on a population level. Through the study of HPV types and variants, we wish to define the critical, specific viral sequences and polymorphisms associated with immune evasion, persistence, transformation (Fehrmann and Laimins, 2003), and possible interactions between HPV and humans (Wang and Hildesheim, 2003). In particular, we want to identify and understand the genetic determinants that set HPV16 apart from other closely-related types in species α9 and from long-persisting but benign types in other species groups

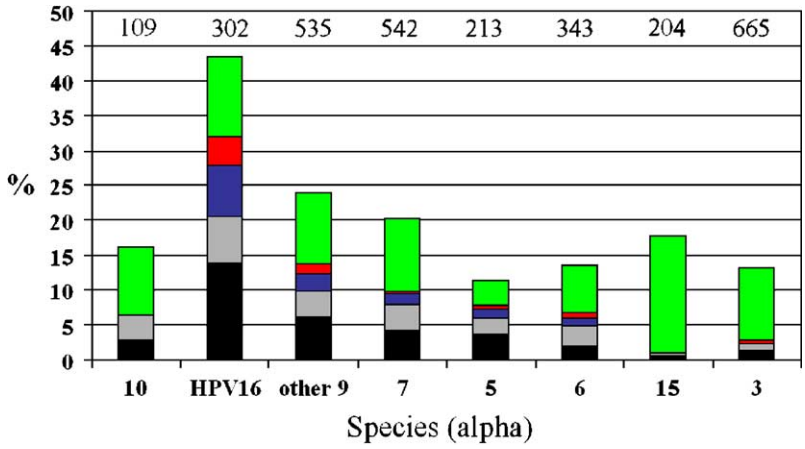

Fig. 3. Outcomes of infections in the 5–7 year study, by species. For the most prevalent species, we graphed the additive percentages of prevalent CIN3/cancer (black), prevalent CIN2 (gray), incident CIN3/cancer given viral persistence (blue), incident CIN2 given viral persistence (red) and viral persistence without progression to CIN2/CIN3/cancer (green). To highlight its uniqueness, we showed HPV16 separately from the rest of species α9. The addition of prevalent and incident CIN3/cancer yields a rough estimate of the absolute risk of CIN3/cancer for each species. Prevalent cases were assigned hierarchically (see Materials and methods and Table 1). The percentage of incident CIN2/CIN3/cancer was calculated among women with follow-up. We noted above each bar the number of infections with types in that species.

(e.g., α3). HPV molecular epidemiology now presents an unparalleled opportunity to study carcinogenesis from the molecular to the population level.

Materials and methods

Phylogenetic analysis

Papillomavirus evolution is inferred from phylogenetic analysis, in which the relatedness of viral DNA and translated proteins from individual types is analyzed. Types with smaller differences between them are presumed to be more closely related, as displayed on the branchings of evolutionary trees. We derived a phylogenetic tree using standard approaches. Specifically, we used independent Bayesian (Huelsenbeck, 2004) and maximum parsimony methods (Swofford, 1998). The HPV genome is quite small, permitting analysis based on most of its genes. All component trees were based on the alignment of concatenated early and late open reading frames (E6, E7, E1, E2, L2 and L1 ORF). The Bayesian tree was calculated from a mixed data set containing alignments of both amino acid and nucleotide sequences. Only informative sites were kept for the two parsimony analyses, one based on nucleotide sequences and the other based on amino acid sequences. Bovine papillomavirus 1 was used as the "referent outgroup", meaning an independent point of comparison to the human papillomaviruses under study. We determined the presence of an E5 ORF since it is the only ORF in the genome of genital HPVs showing significant heterogeneity in its presence. (We consider the E5 ORF identified in HPV16 as the bona fide E5, since the protein has been identified in lesions (Kell et al., 1994)). HPV genomes were examined for the presence of a homologous E5 ORF based on the length and sequence similarity of codon region between E2 and L2 ORFs by pairwise alignment with the HPV16 E5 amino acid sequence. In addition, we also checked the annotated file of each HPV type in GenBank, NCBI.

Epidemiologic methods

The Guanacaste study population is outlined in Fig. 2. The Guanacaste Project enrolled 10,049 (93.6% of eligible) women by random sampling of the high-risk Costa Rican province in 1993-1994, following approval by Costa Rican and U.S. ethical boards and individual written informed consent (Bratti et al., 2004; Herrero et al., 1997). A few, intensively trained nurses screened adult women (age range 18–97) using cytology and cervicography (magnified cervical images). The cytologic specimens were collected with cervical brooms to create split-sample cytology (conventional then liquid-based). Cells for HPV DNA testing were collected using Dacron swabs placed into Digene specimen transport medium, and kept at −70° for long-term storage.

For both the cross-sectional (enrollment) and prospective parts of this analysis, we excluded women with prior hysterectomies (n = 630), who were virgins (n = 583), who refused a pelvic exam (n = 291), or who were missing enrollment HPV test results (n = 9). The remaining women were categorized into several groups based on the results of the screening exam; the groups were followed differently according to perceived risk of developing CIN3/cancer. We actively re-screened every 6–12 months those women considered at elevated risk because of equivocal or mildly abnormal cytology, HPV DNA positivity using the Hybrid Capture Tube Test (Digene Corporation, Gaithersburg, MD), a positive cervigram, or a lifetime report of >4 sexual partners. A randomly-chosen, comparison group with entirely normal findings at enrollment was also followed

actively every year, while the remaining women with normal screening results (n = 5,620) were assigned to passive follow-up at 5–7 years, with 4,903 (87.2%) participating.

For the prospective analysis, the final analysis group excluded women diagnosed with CIN2 or more severe (≥CIN2) at enrollment (n = 140) or missing follow-up PCR results for any reason (n = 1,118). Missing follow-up was strongly related to younger age, due to younger women leaving Guanacaste for work ($P < 0.0001$, Pearson χ^2).

Because women were re-screened on different schedules related to risk of CIN3/cancer, we concentrated on long-term viral persistence, which we could examine for the entire cohort. We chose the follow-up specimens from return visits 5–7 years after enrollment except for women referred to colposcopy earlier for possible incident CIN2 or worse. From them, we tested the last available specimen. Women had a mean of 3.2 visits (standard deviation = 2.0 visits), median of 2.0 visits, and a range of 2-10 visits including enrollment and follow-up specimen collection visits. The mean follow-up of the 7278 women in the prospective analysis was 5.6 +/− 1.2 years (median 5.1).

All women with possible CIN2 or worse at any time, detected by any screening technique including nurse concern on gross examination (and whether ultimately confirmed on initial or final histologic review) were referred to colposcopy with guided biopsy of visible lesions. Women diagnosed locally on initial histologic evaluation as CIN 2 or worse were treated by LEEP or by inpatient surgery if needed; women treated with LEEP were re-examined at six-month intervals to insure treatment success.

Although CIN2 was treated if diagnosed locally, it is not an optimal surrogate endpoint for cancer risk because of diagnostic variability and biologic variability. We wished to be as certain as possible that we were studying risk for cancer or its immediate precursor (Stoler and Schiffman, 2001). Therefore, histologic CIN2 was kept separate from the primary CIN3/cancer analytic case group. CIN1 was considered a morphologic consequence of HPV infection, not serious neoplasia, and was not treated during follow-up (Wright et al., 2003).

For study purposes (not clinical care), histology diagnosed locally as CIN2 or worse locally was reviewed by U.S. pathologists (M.E.S. and T.C.W. for the enrollment cases; M.E.S. and D.S. for the follow-up cases). The few cases of glandular lesions were combined with comparably severe squamous in situ or invasive carcinoma. The final assignment of cases as invasive cancer, CIN3, CIN2, <CIN2 was made by an algorithm based on independent masked reviews. Only a few very difficult cases were adjudicated by joint review, occasionally with consideration of cytologic slides as well as histology.

Statistical methods

For this analysis, we thoroughly examined each woman at enrollment and 5–7 years later, to ascertain the HPV type-specific risk of prevalent neoplasia, viral persistence, and incident cervical neoplasia. Except for exclusion of women with no follow-up from prospective rates, HPV prevalences and risks were calculated without adjustment for loss-to-follow-up because of the very high participation. In calculating the fractions of prevalent CIN3/cancer attributable to individual HPV types, we accounted for frequent multiple infections in cases (41% of the 85 prevalent CIN3/cancer cases had multiple infections, yielding 138 case-associated infections in all) by hierarchical analysis: HPV types were sorted according to attributable risk of CIN3/cancer. When analyzing the contribution of a given type, women with types accounting for a higher fraction of CIN3/cancer were excluded. For example, women with the type accounting for the highest fraction of CIN3/cancer (i.e., HPV16) were excluded to find the next most important type (i.e., HPV58), and women with either type were excluded to find the third, etc. The hierarchical approach might have underestimated the risk of some types with low attributable fractions (e.g., HPV52), thus, crude and hierarchically-adjusted estimates are both shown.

HPV testing methods

HPV DNA was detected using MY09/M11 L1 consensus primer PCR (MY09/11 PCR) with AmpliTaq Gold polymerase as previously described (Castle et al., 2002; Herrero et al., 2000; Qu et al., 1997). Briefly, an aliquot of the STM specimen was lysed, and the specimen DNA was precipitated by ammonium acetate/ethanol solution and then pelleted by centrifugation. The DNA pellet was suspended in 10 mM Tris, pH 8.0 with 0.1 mM EDTA and stored frozen until used. Thermocyling conditions included initial denaturation at 95 °C for 9 minutes; thereafter, each cycle consisted of 95 °C for 60 seconds, 55 °C annealing for 60 seconds, and extension at 72 °C for 60 seconds for a total 40 cycles with a final extension at 72 °C for 5 minutes. A 100-cell copy SiHa HPV DNA positive control, a 2-cell copy SiHa HPV DNA positive control, and a 100-cell copy of HuH7 HPV DNA negative control were used per every 48 specimens tested.

PCR products were analyzed by gel electrophoresis and then transferred to nylon filters. The filters were hybridized overnight with radiolabeled generic probes for HPV (HPV 11, 16, 18, 51, 73 and 81 combined). Thereafter, HPV PCR products were typed using dot blot hybridization with biotinylated type-specific oligonucleotide probes for HPV types: 2, 6, 11, 13, 16, 18, 26, 31-35, 39, 40, 42-45, 51-59, 61, 62, 64, 66-74v, 81-85, 82v (AE2), 89, AE9, and AE10. Probes for HPV Types 2, 13, 34, 42-4, 57, 64, 69, 74, 82, and AE9 were combined in dot blot hybridizations for detection of rare types (dbmix). A specimen was considered HPV positive, but uncharacterized, if it tested positive for HPV DNA by the radiolabeled generic probe mix but was not positive for any type specific probe.

Acknowledgments

The Guanacaste Project was supported by the National Cancer Institute, National Institutes of Health, Department of Health and Human Services contracts NO1-CP-21081, NO1-CP-33061, NO1-CP-40542 and NO1-CP-506535. We are grateful for the support of Costa Rican health authorities and for the decade-long dedication of the superb project staff. Dr. Burk was supported by National Cancer Institute grant CA78527.

References

Andersson, S., Rylander, E., Larsson, B., Strand, A., Silfversvard, C., Wilander, E., 2001. The role of human papillomavirus in cervical adenocarcinoma carcinogenesis. Eur. J. Cancer 37, 246–250.

Bosch, F.X., de Sanjose, S., 2003. Chapter 1: Human papillomavirus and cervical cancer–burden and assessment of causality. J. Natl. Cancer Inst. Monogr. No. 31, 3–13.

Bosch, F.X., Lorincz, A., Munoz, N., Meijer, C.J., Shah, K.V., 2002. The causal relation between human papillomavirus and cervical cancer. J. Clin. Pathol. 55, 244–265.

Bratti, M.C., Rodriguez, A.C., Schiffman, M., Hildesheim, A., Morales, J., Alfaro, M., Guillen, D., 2004. Description of a seven-year prospective study of HPV infection and cervical neoplasia among 10,000 women in Guanacaste. Costa Rica. J. P. A. H. O. 15, 75–89.

Castle, P.E., Schiffman, M., Gravitt, P.E., Kendall, H., Fishman, S., Dong, H., Hildesheim, A., Herrero, R., Bratti, M.C., Sherman, M.E., Lorincz, A., Schussler, J.E., Burk, R.D., 2002. Comparisons of HPV DNA detection by MY09/11 PCR methods. J. Med. Virol. 68, 417–423.

Castle, P.E., Schiffman M., Herrero R., Hildesheim A., Rodriguez A.C., Bratti C., Sherman M.E., Wacholder S., Tarone R., Burk R.D., in press. A prospective study of age trends in cervical HPV acquisition and persistence in Guanacaste, Costa Rica. J Infect Dis.

de Villiers, E.M., Fauquet, C., Broker, T.R., Bernard, H.U., zur, H.H., 2004. Classification of papillomaviruses. Virology 324, 17–27.

DiMaio, D., Mattoon, D., 2001. Mechanisms of cell transformation by papillomavirus E5 proteins. Oncogene 20, 7866–7873.

Fehrmann, F., Laimins, L.A., 2003. Human papillomaviruses: targeting differentiating epithelial cells for malignant transformation. Oncogene 22, 5201–5207.

Gravitt, P.E., Peyton, C.L., Alessi, T.Q., Wheeler, C.M., Coutlee, F., Hildesheim, A., Schiffman, M.H., Scott, D.R., Apple, R.J., 2000. Improved amplification of genital human papillomaviruses. J. Clin. Microbiol. 38, 357–361.

Herrero, R., Schiffman, M.H., Bratti, C., Hildesheim, A., Balmaceda, I., Sherman, M.E., Greenberg, M., Cardenas, F., Gomez, V., Helgesen, K., Morales, J., Hutchinson, M., Mango, L., Alfaro, M., Potischman, N.W., Wacholder, S., Swanson, C., Brinton, L.A., 1997. Design and methods of a population-based natural history study of cervical neoplasia in a rural province of Costa Rica: the Guanacaste Project. Rev. Panam. Salud Publica 1, 362–375.

Herrero, R., Hildesheim, A., Bratti, C., Sherman, M.E., Hutchinson, M., Morales, J., Balmaceda, I., Greenberg, M.D., Alfaro, M., Burk, R.D., Wacholder, S., Plummer, M., Schiffman, M., 2000. Population-based study of human papillomavirus infection and cervical neoplasia in rural Costa Rica. J. Natl. Cancer Inst. 92, 464–474.

Herrero, R., Castellsague, X., Pawlita, M., Lissowska, J., Kee, F., Balaram, P., Rajkumar, T., Sridhar, H., Rose, B., Pintos, J., Fernandez, L., Idris, A., Sanchez, M.J., Nieto, A., Talamini, R., Tavani, A., Bosch, F.X., Reidel, U., Snijders, P.J., Meijer, C.J., Viscidi, R., Munoz, N., Franceschi, S., 2003. Human papillomavirus and oral cancer: the International Agency for Research on Cancer multicenter study. J. Natl. Cancer Inst. 95, 1772–1783.

Ho, G.Y., Burk, R.D., Klein, S., Kadish, A.S., Chang, C.J., Palan, P., Basu, J., Tachezy, R., Lewis, R., Romney, S., 1995. Persistent genital human papillomavirus infection as a risk factor for persistent cervical dysplasia. J. Natl. Cancer Inst. 87, 1365–1371.

Ho, G.Y., Bierman, R., Beardsley, L., Chang, C.J., Burk, R.D., 1998. Natural history of cervicovaginal papillomavirus infection in young women. N. Engl. J. Med. 338, 423–428.

Huelsenbeck, J.P.R.F., 2004. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.

Kashima, H.K., Mounts, P., Shah, K., 1996. Recurrent respiratory papillomatosis. Obstet. Gynecol. Clin. North Am. 23, 699–706.

Kell, B., Jewers, R.J., Cason, J., Pakarian, F., Kaye, J.N., Best, J.M., 1994. Detection of E5 oncoprotein in human papillomavirus type 16-positive cervical scrapes using antibodies raised to synthetic peptides. J. Gen. Virol. 75, 2451–2456.

Koutsky, L.A., Ault, K.A., Wheeler, C.M., Brown, D.R., Barr, E., Alvarez, F.B., Chiacchierini, L.M., Jansen, K.U., 2002. A controlled trial of a human papillomavirus type 16 vaccine. N. Engl. J. Med. 347, 1645–1651.

Lowy, D.R., Frazer, I.H., 2003. Chapter 16: Prophylactic human papillomavirus vaccines. J. Natl. Cancer Inst. Monogr. No. 31, 111–116.

Molano, M., Van den, B.A., Plummer, M., Weiderpass, E., Posso, H., Arslan, A., Meijer, C.J., Munoz, N., Franceschi, S., 2003. Determinants of clearance of human papillomavirus infections in Colombian women with normal cytology: a population-based, 5-year follow-up study. Am. J. Epidemiol. 158, 486–494.

Munoz, N., Bosch, F.X., de Sanjose, S., Herrero, R., Castellsague, X., Shah, K.V., Snijders, P.J., Meijer, C.J., 2003. Epidemiologic classification of human papillomavirus types associated with cervical cancer. N. Engl. J. Med. 348, 518–527.

Nobbenhuis, M.A., Walboomers, J.M., Helmerhorst, T.J., Rozendaal, L., Remmink, A.J., Risse, E.K., van der Linden, H.C., Voorhorst, F.J., Kenemans, P., Meijer, C.J., 1999. Relation of human papillomavirus status to cervical lesions and consequences for cervical-cancer screening: a prospective study. Lancet 354, 20–25.

Qu, W., Jiang, G., Cruz, Y., 1997. PCR detection of human papillomavirus: comparison between MY09/MY11 and GP5+/GP6+ primer systems. J. Clin. Microbiol. 35, 1304–1310.

Richardson, H., Kelsall, G., Tellier, P., Voyer, H., Abrahamowicz, M., Ferenczy, A., Coutlee, F., Franco, E.L., 2003. The natural history of type-specific human papillomavirus infections in female university students. Cancer Epidemiol. Biomarkers Prev. 12, 485–490.

Schiffman, M., Kjaer, S.K., 2003. Chapter 2: Natural history of anogenital human papillomavirus infection and neoplasia. J. Natl. Cancer Inst. Monogr. No. 31, 14–19.

Schiffman, M., Hildesheim, A., Herrero, R., Bratti, C., 2000. Human papillomavirus testing as a screening tool for cervical cancer. JAMA 283, 2525–2526.

Schlecht, N.F., Kulaga, S., Robitaille, J., Ferreira, S., Santos, M., Miyamura, R.A., Duarte-Franco, E., Rohan, T.E., Ferenczy, A., Villa, L.L., Franco, E.L., 2001. Persistent human papillomavirus infection as a predictor of cervical intraepithelial neoplasia. JAMA 286, 3106–3114.

Stoler, M.H., Schiffman, M., 2001. Interobserver reproducibility of cervical cytologic and histologic interpretations: realistic estimates from the ASCUS-LSIL Triage Study. JAMA 285, 1500–1505.

Swofford, D.L., 1998. PAUP*: Phylogenetic analysis using parsimony (*and other methods). Sinauer, Sunderland, Massachusetts.

Van Ranst, M., Kaplan, J.B., Burk, R.D., 1992. Phylogenetic classification of human papillomaviruses: correlation with clinical manifestations. J. Gen. Virol. 73, 2653–2660.

Wang, S.S., Hildesheim, A., 2003. Chapter 5: Viral and host factors in human papillomavirus persistence and progression. J. Natl. Cancer Inst. Monogr. No. 31, 35–40.

Woodman, C.B., Collins, S., Winter, H., Bailey, A., Ellis, J., Prior, P., Yates,

M., Rollason, T.P., Young, L.S., 2001. Natural history of cervical human papillomavirus infection in young women: a longitudinal cohort study. Lancet 357, 1831–1836.

Wright Jr., T.C., Cox, J.T., Massad, L.S., Carlson, J., Twiggs, L.B., Wilkinson, E.J., 2003. 2001 consensus guidelines for the management of women with cervical intraepithelial neoplasia. Am. J. Obstet. Gynecol. 189, 295–304.

zur Hausen, H., 2000. Papillomaviruses causing cancer: evasion from host-cell control in early events in carcinogenesis. J. Natl. Cancer Inst. 92, 690–698.