Recurring infection with ecologically distinct HPV types can explain high prevalence and diversity

Sylvia L. Ranjeva[a,1], Edward B. Baskerville[a], Vanja Dukic[b], Luisa L. Villa[c], Eduardo Lazcano-Ponce[d], Anna R. Giuliano[e], Greg Dwyer[a], and Sarah Cobey[a]

[a]Department of Ecology and Evolution, University of Chicago, Chicago, IL 60637; [b]Department of Applied Mathematics, University of Colorado, Boulder, CO 80309; [c]Faculdade de Medicina, Universidade de São Paulo Department of Radiology and Oncology, Centro de Investigação Translacional em Oncologia, Universidade de São Paulo, 01246-000, São Paulo, Brazil; [d]Centro de Investigación en Salud Poblacional, Instituto Nacional de Salud Pública, Cuernavaca, 62100, Mexico; and [e]Center for Infection Research in Cancer, Moffitt Cancer Center & Research Institute, Tampa, FL 33612

The high prevalence of human papillomavirus (HPV), the most common sexually transmitted infection, arises from the coexistence of over 200 genetically distinct types. Accurately predicting the impact of vaccines that target multiple types requires understanding the factors that determine HPV diversity. The diversity of many pathogens is driven by type-specific or "homologous" immunity, which promotes the spread of variants to which hosts have little immunity. To test for homologous immunity and to identify mechanisms determining HPV transmission, we fitted nonlinear mechanistic models to longitudinal data on genital infections in unvaccinated men. Our results provide no evidence for homologous immunity, instead showing that infection with one HPV type strongly increases the risk of infection with that type for years afterward. For HPV16, the type responsible for most HPV-related cancers, an initial infection increases the 1-year probability of reinfection by 20-fold, and the probability of reinfection remains 14-fold higher 2 years later. This increased risk occurs in both sexually active and celibate men, suggesting that it arises from autoinoculation, episodic reactivation of latent virus, or both. Overall, our results suggest that high HPV prevalence and diversity can be explained by a combination of a lack of homologous immunity, frequent reinfections, weak competition between types, and variation in type fitness between host subpopulations. Because of the high risk of reinfection, vaccinating boys who have not yet been exposed may be crucial to reduce prevalence, but our results suggest that there may also be large benefits to vaccinating previously infected individuals.

infectious disease | mathematical model | ecology | public health | vaccination

Human papillomavirus (HPV), a major cause of genital warts and anogenital and oropharyngeal cancers (1), is the most common sexually transmitted infection (2). While the population prevalence of genital HPV is ∼40% among women and 45% among men in the United States, over 200 HPV types have been identified, and the prevalence of individual types never exceeds 10% (3). The quadrivalent vaccine has been effective against the HPV types that cause the most disease (4), and a recent nine-valent vaccine covers additional oncogenic types (5). Predicting the effects of multivalent vaccines on HPV prevalence, however, is difficult without defining the factors that underlie HPV transmission and diversity, which are poorly understood.

Epidemiological theory has shown that the abundance of many pathogens depends on the dynamics of competition for susceptible hosts (6–8). Pathogen strains compete by inducing adaptive immune responses that are specific to shared antigens, limiting the growth rates of antigenically similar competitors (8). The accumulation of specific immunity in the host population to common strains decreases the transmission of those strains and promotes the spread of rarer antigenic variants, a phenomenon known as negative frequency-dependent selection. Such immune-mediated competition can partly explain

the antigenic and genetic diversity of influenza, pneumococcus, rotavirus, norovirus, *Neisseria meningitidis*, malaria, hepatitis C, HIV, trypanosomes, and other common pathogens (6–9).

It is unclear how HPV interacts with the immune system during infection, but in principle, distinctions between HPV types may arise because of acquired immunity from B cells and T cells. HPV types are defined by a 10% threshold of dissimilarity in the L1 gene, which codes for the major capsid protein (10). The outer capsid modulates viral entry into host cells at the epithelial basement membrane (11), and the humoral response to infection is mainly type-specific anti-L1 antibodies (12). Studies of HPV in T cell-deficient people show that cellular immunity is important to control infection (13), but it is unclear how cellular immunity achieves type-specific recognition, if at all. In individuals with HPV16-related cancerous lesions, cytotoxic T lymphocytes specific to the E6 and E7 oncoproteins are correlated with reduced disease (14, 15). The specificity of the T-cell repertoire to other genes or to the majority of HPV types, however, is not well-established.

Efforts to understand the effects of immunity on HPV dynamics must begin with homologous immunity or protection against

Significance

Human papillomavirus (HPV) is the most common sexually transmitted infection. Understanding the factors underlying the prevalence and diversity of over 200 HPV types is necessary to predict the effects of multivalent vaccines. We investigate HPV transmission dynamics by fitting mechanistic models to data from unvaccinated men. We find that infection risk in a susceptible population is, on average, low and concentrated in high-risk individuals and that, rather than inducing protective immunity, HPV infection strongly increases the risk of future infection by the same type. Thus, high HPV prevalence results from frequent reinfection or persistent infections within individuals. Vaccinating boys before their first HPV exposure may be especially effective in reducing prevalence, and vaccinating previously infected men may also be effective.

repeat infection with the same HPV type. Homologous immunity would limit the prevalence of any type through negative frequency dependence. The traditional assumption is that most HPV-infected individuals permanently clear infection after 1–2 y (16–18), suggesting protective homologous immunity. In this scenario, the elevated cancer risk associated with particular HPV types results from a small fraction of persistent infections (18). Longitudinal studies, however, have shown that hosts can be infected repeatedly (19). Although type-specific antibodies may provide modest protection against future infection in women (20), serum antibody is not a marker of immune protection in men (21, 22). Thus, the strength of homologous immunity remains unclear.

Evidence for competition between HPV types is also conflicting. Immune-mediated competition has been invoked to explain small increases in the prevalence of nonvaccine HPV types after HPV vaccination (23) and to explain weak cross-protection from the vaccine against related nonvaccine types (24). Mathematical models show that type competition is consistent with observed patterns of HPV prevalence (25, 26). Nevertheless, there is little empirical evidence for intertype competition (27), as shown in part by elevated rates of multiple-type compared with single-type infections and frequent concurrent acquisition of HPV types (28).

Meanwhile, the risk of HPV infection depends on differences in demographic and behavioral risk factors between host subpopulations. For example, a host's number of sexual partners strongly affects infection risk (1, 18, 29), and some evidence suggests that oncogenic and nononcogenic types are differently sensitive to the numbers of partners among hosts (30). However, detailed comparisons of risk factors for infection with each HPV type have not been performed. Different risk factors would suggest differences between types among host subpopulations that could explain type prevalence.

To investigate the factors determining HPV prevalence and diversity, we fitted mechanistic models of HPV dynamics to an extensive longitudinal dataset. Mechanistic models have long been used in infectious disease ecology to quantify the biological processes that underlie pathogen dynamics, but HPV has received comparatively little attention (17). HPV models have generally focused on qualitative dynamics (25, 31–33) and on predictions for health policy (34, 35), relying on informal methods to estimate parameters. Detailed longitudinal studies present an opportunity to use mechanistic models to rigorously test hypotheses with robust statistical methods. Here, we use such methods to show that the prevalence and diversity of HPV types are best explained by the combined effects of a lack of competition within and between types, high rates of reinfection or persistence within individuals, and modest differences in high-risk subpopulations between types.

Results

Low-Prevalence HPV Types Coexist over Time. We fitted models to data from the HPV in Men (HIM) Study, which tracks genital HPV infection and demographic and behavioral traits in unvaccinated men sampled at 6-mo intervals over 5 y (36). In these data, the mean prevalence of all HPV is 65%, but no type has a mean prevalence greater than 10% (Fig. 1A), and the prevalence of individual types is roughly constant over time (Fig. 1B), showing coexistence. The rank prevalence of HPV types is also constant among geographical locations, although absolute prevalence is higher in Brazil than in the United States or Mexico (*SI Appendix*, Fig. S4).

We analyzed the five HPV types with the highest mean prevalence: HPV62, HPV84, HPV89, HPV16, and HPV51. We also analyzed HPV6, the ninth most prevalent type, because it is included in the quadrivalent vaccine and is highly associated with genital warts (1). We accounted for previously identi-
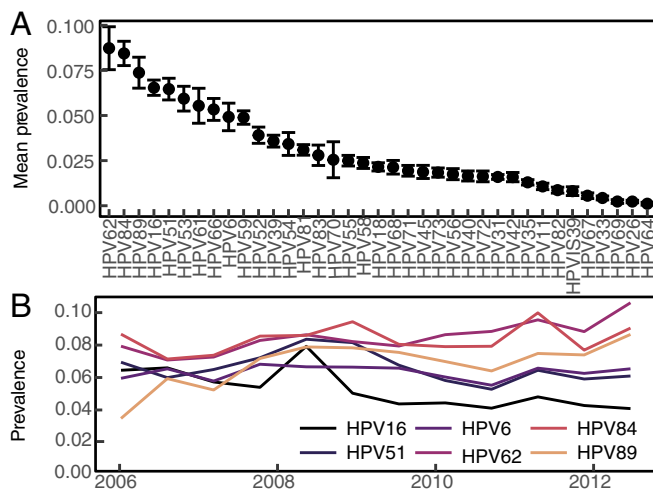

Fig. 1. (*A*) Average prevalence of HPV types in the study population, with SEs calculated across visits. (*B*) Prevalence over time of the six types analyzed.

fied risk factors for HPV infection in the HIM Study dataset (29, 36). The risk factors for any HPV type include markers of increased exposure to infected sexual partners as well as nonsexual behaviors, such as tobacco use, and demographic traits. Several factors, notably circumcision and sexual orientation, differ in their effects among HPV types (37, 38). Type-specific differences may reflect meaningful ecological distinctions. We therefore modeled the effects of a diverse set of risk factors.

Past Infection with HPV Confers Minimal Protection Against Infection with the Same Type. Our models test three hypotheses about HPV infection. Under the simplest or "memoryless" model, infection risk depends only on the effects of host- and HPV type-specific risk factors (Eq. **1**), with no consideration of immunity. Two more complex models account for the effects of previous infection with the same type (see *Materials and Methods* for model details).

Our three models differ only in their assumptions about the instantaneous per capita infection risk, or "force of infection" (39), for host i with HPV type j at time t given by λ_{ijt}.

i) In the memoryless model, the force of infection depends only on host-specific risk factors and the type-specific baseline force of infection, λ_{0_j}:

$$\lambda_{ijt} = \lambda_{0_j} f(\overrightarrow{\alpha_j} C_{it}). \quad [1]$$

The vector $\overrightarrow{\alpha_j}$ scales the effect of each of the M covariates on the force of infection, where C_{it} is the covariate matrix:

$$f(\overrightarrow{\alpha_j} C_{it}) = e^{\alpha_{j,1} C_{1it} + \alpha_{j,2} C_{2it} + \alpha_{j,3} C_{3it} + \ldots + \alpha_{j,M} C_{Mit}}. \quad [2]$$

Each model included five continuous or ordinal covariates and six binary covariates ($M = 11$ individual-level risk factors).

ii) In the homologous immunity model, protective immunity reduces the probability of reinfection ($p_{\text{subsequent}}$) relative to the probability of an initial infection (p_{first}) with type j:

$$p_{\text{first}} = (1_{t+\Delta t}|0_t) = \left(1 - e^{-\lambda_{0_j} f(\overrightarrow{\alpha_j} C_{it})\Delta t}\right), \quad [3]$$

$$p_{\text{subsequent}} = p_{\text{first}}(1 - (1 - d)e^{-w(t-t_{\text{clr}})}).$$

The strength of immunity d is constrained to fall between zero and one, so that d reduces the probability of reinfection
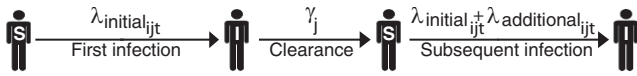
Fig. 2. Dynamics for one individual host *i* and one HPV type *j* under the additional risk model (Eq. **4**). S and I denote susceptible and infected, respectively. The duration of each infection is drawn from a gamma distribution with mean $1/\gamma_j$.

over the time interval Δt. Note that, for $d = 1$, the homologous immunity model is identical to the memoryless model. After the previous infection clears at time t_{clr}, immunity wanes at rate w.

iii) In the additional risk model, the risk of an initial infection is determined as in the memoryless model (Eq. **1**), but the risk of reinfection is allowed to be higher. The force of infection thus includes an additional risk factor, which is distinct from the host-specific covariate effects (such as the number of recent sexual partners) and describes only the effect of previous infection (Eq. **4** and Fig. 2):

$$\lambda_{ijt} = \lambda_{0_j} f(\vec{\alpha_j} \boldsymbol{C_{it}}) + \mathrm{I}_{\text{prev}} \, d_{jc_i} \, e^{-w_j(t - t_{\text{clr}})}. \qquad [4]$$

The additional risk d_{jc_i} depends on the infecting type j and on the sexual subclass c_i of host i. The variable I_{prev} indicates whether there was a previous infection with type j.

We included an observation model that determines the probability of the observed data using the sensitivity and specificity of the HPV genotyping test. We fitted our models to the data using a likelihood-based nonlinear fitting routine (*SI Appendix*) (40). For model selection, we used the corrected Akaike Information Criterion (AICc), which balances the better fit of more complex models against the parsimony of simpler models, while including a correction for finite sample size (41). The best model has the smallest AICc value.

We first tested whether infection with an HPV type reduces the risk of subsequent infection with the same type, allowing for variation in the degree and duration of protective homologous immunity [3]. Notably, the homologous immunity model reduced to the memoryless model for each HPV type ($d = 1$) (Eq. **3**) at the point estimate of the maximum likelihood. The likelihood of the homologous immunity model for each type was thus effectively identical to that of memoryless model (*SI Appendix*, Table S1). Small discrepancies in the likelihoods in the two models arise from the Monte Carlo error associated with the use of simulation-based estimation of the likelihood (*SI Appendix*). The additional parameter in the homologous immunity model therefore increased model complexity without meaningfully improving the likelihood, yielding higher AICc scores. There is thus no evidence for homologous immunity against any of the HPV types that we studied, suggesting that intratype competition is weak or absent.

Because separate types are less closely related, competition between types should be weaker still. Previous work has nevertheless speculated that there may be cross-immunity between HPV types. In particular, virus-like particles of HPV16 can induce a low level of neutralizing antibodies against HPV31, and clinical trial data suggest that current vaccines, which immunize against HPV16, provide partial protection against HPV31 (42, 43). We therefore tested for competition between HPV16 and HPV31 by fitting a model in which previous infection with HPV31 affects the risk of infection with HPV16 (*SI Appendix*). Our estimate of the additional risk of HPV16 infection given previous HPV31 infection is centered around 1 (1.3; 95% confidence interval: 0.5, 2.0), suggesting that the two types do not compete. This lack of an effect may be partly caused by the low prevalence of HPV31 (Fig. 1) and correspondingly low statistical power, but

the lack of even a trend toward competition nevertheless suggests that there is no interaction. We therefore conclude that neither intratype nor intertype competition has strong effects on HPV dynamics.

Past Infection with HPV Strongly Increases the Risk of Future Infection with the Same Type. The additional risk model fits the data vastly better than the homologous immunity and memoryless models for all six HPV types (*SI Appendix*, Table S1). Here, we included individuals who were infected at baseline, but excluding such individuals yielded the same result (*SI Appendix*). To determine whether the additional risk was caused by repeated exposure to the same HPV type by infected sexual partners, we fitted the magnitude of the additional risk d separately for three sexual subclasses: celibate individuals, individuals reporting one recent sexual partner, and individuals reporting multiple recent sexual partners. We included only people who remained in the same sexual subclass for at least 3 consecutive years. Celibate individuals were defined as reporting no recent receptive or insertional anal, vaginal, or oral sex with male or female partners (*SI Appendix*). For all six types, previous infection significantly raises the risk of subsequent homologous infection in each sexual subclass (Fig. 3*A*). The high additional risk experienced by celibate individuals (d_{celibate}) strongly suggests that serial infections are driven by factors other than sexual transmission. Crucially, our estimate of the additional risk is uncorrelated with our estimate of baseline infection risk, λ_{0_j} (*SI Appendix*, Fig. S1), showing that the high estimated additional risk does not reflect statistical nonidentifiability between additional risk and baseline risk. Estimates of the additional risk between different sexual subclasses have the same magnitude (Fig. 3*A*), showing that variation in the additional risk across sexual subclasses is minimal.

To quantify the impact of previous infection, we calculated the effect of previous infection on the total risk of a subsequent homologous infection at $t = 0$, 1, and 3 y after clearing the previous infection. Even several years after the initial infection is cleared, the additional risk caused by previous infection accounts for more than 90% of the force of infection (*SI Appendix*, Fig. S3). Moreover, immediately after infection, the 1-y probability of reinfection with HPV16 (Eq. **5**) is, on average, 20.4-fold higher
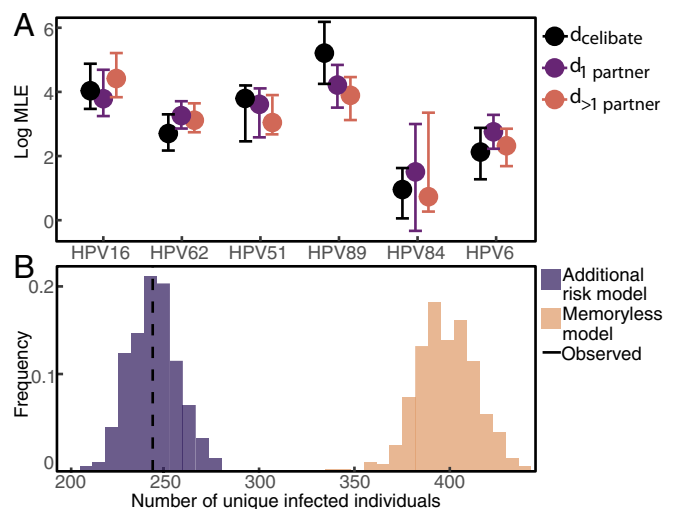

Fig. 3. (*A*) Log estimate of the additional risk d by sexual subclass for each HPV type [maximum likelihood estimate (MLE) and 95% confidence interval]. (*B*) Predicted number of unique individuals infected with HPV16 during follow-up across 1,000 simulations under the additional risk model compared with the memoryless model and the data.

than the probability of infection in a naive individual. The average increase is 19.1- to 20.5-fold among HPV types. Even 3 y after clearing an infection, the probability of reinfection with HPV16 in the following year remains 13.5-fold higher (7.4- to 20.5-fold among types).

The strong additional risk independent of sexual subclass is consistent with two major biological explanations: repeat infections, presumably caused by autoinoculation between anatomical sites, or episodic reactivation of latent infection. Because our models account for 11 different risk factors other than previous infection, it is unlikely that the increased risk is because of confounding by unmeasured individual-level risk factors. Completely ruling out unmeasured covariates is impossible without controlled experiments. As an initial test, however, we repeated our estimation of the additional risk d using a model that includes additional measured covariates (17 in total). In this more complex model, the additional risk d still accounts for more than 90% of the force of infection after the initial infection, and the model with more covariates did not provide a better fit to the data than the best fit model with 11 covariates (*SI Appendix*). This result suggests that the increased risk is not simply because of unmeasured individual-level risk factors, while emphasizing the inability of traditional risk factors to explain the vast majority of HPV infections.

Modest Differences in Host-Specific Risk Factors Suggest Ecological Differences Between HPV Types and Highlight High-Risk Subpopulations. Although the additional risk conferred by past infection is substantial, a model without host-specific risk factors fits the data far worse than the full additional risk model (Eq. **4**) for all types (*SI Appendix*). Moreover, the effects of the host-specific risk factors vary among HPV types. To understand this variation, we inserted our estimates of the baseline force of infection and the covariate effects for each HPV type into the best fit model to calculate the distribution of infection risk in individuals who have never been infected. The expected time to infection ($1/\lambda_{ijt}$), a measure of risk, is generally low but varies by orders of magnitude among individuals in the naive population (*SI Appendix*, Fig. S2). Thus, the risk of initial infection with any type is concentrated in a few high-risk individuals. This feature captures a major pattern in the data. Under the additional risk model, the population prevalence of an HPV type arises from repeated infections in relatively few individuals. In the memoryless model, by contrast, the initial infection risk is higher, previous infection has no effect on subsequent risk, and prevalence arises from fewer infections in more individuals. Simulations from the maximum likelihood estimates of the two models show that both reproduce the population-level prevalence in the data across visits (*SI Appendix*, Fig. S9). However, only the additional risk model accurately predicts the fraction of individuals who ever experience infection, which the memoryless model overestimates (Fig. 3B). Unmeasured assortative mixing (44) and simultaneous exposure to multiple HPV types through sexual partnerships (45) can confound associations between HPV incidence and past exposures. Residual confounding may be a problem, because the initial infection risk is low and sensitive to host-specific risk factors. However, such confounding would be minimized in celibate individuals. We, therefore, reestimated the magnitude of the additional risk d_{celibate} by fitting the additional risk model (Eq. **4**) to data from celibate individuals only. In the celibates-only analysis, the risk conferred by previous infection was still high (*SI Appendix*) and consistent with our previous estimates. The high additional risk inferred for the celibate individuals and the consistency of the additional risk for all three sexual subclasses suggest that the additional risk reflects autoinoculation or reactivation rather than confounding from unmeasured sexual activity.

The effects of each covariate also show interesting similarities and differences between types (Fig. 4). Having more recent female partners strongly increases risk for all HPV types, emphasizing the importance of heterosexual transmission. The addition of a single female sexual partner raises instantaneous infection risk by 80–120% among types compared with the risk in individuals with no recent female partners. The number of male partners, however, has divergent effects. The addition of a single male sexual partner reduces instantaneous risk with HPV16 by 50% but increases the risk of infection with types HPV84 and HPV89 by 20 and 70%, respectively. Most other covariates were significant for only some types, indicating subtle differences between types. Our results are largely consistent with previous findings, with differences likely arising from the greater statistical power of our inference technique (*SI Appendix*).

Discussion

Understanding the dynamics of HPV transmission is crucial to explain the coexistence of over 200 low-prevalence HPV types and to accurately predict the impact of multivalent HPV vaccines. Our results show little evidence of competition within and between HPV types, and instead, they show high rates of reinfection or persistence by the same type. The inferred lack of both homologous and heterologous immunity implies that HPV-type diversity cannot be explained by negative frequency-dependent selection, which promotes the coexistence of other immunologically distinct pathogen strains. Our results instead suggest that high HPV prevalence results from continuous or repeated infection with multiple, apparently independent virus types. Types differ slightly in their risk factors, suggesting that ecological niche partitioning could further promote coexistence.

Our results add to growing evidence that HPV infection, in contrast to vaccination, does not confer protective homologous immunity in men. Although we allowed the strength and duration of immunity to vary after initial infection, we could not detect any protection against reinfection with the same type. Because sterilizing immunity reduces infection rates, models that assume that infection induces protective immunity in men (33–35, 46) likely underestimate vaccine effectiveness.

Our conclusion that there is no homologous immunity in men supports the hypothesis that genital infection differs between men and women. Although the durations and distributions of types in genital HPV infection are similar in men and women
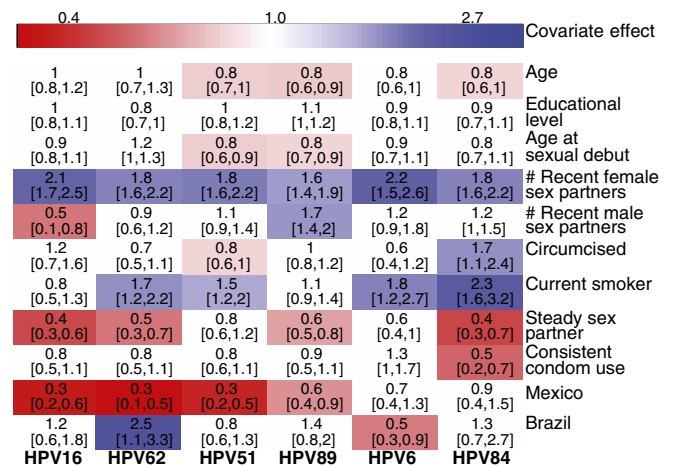

0.4		1.0		2.7		Covariate effect
1 [0.8,1.2]	1 [0.7,1.3]	0.8 [0.7,1]	0.8 [0.6,0.9]	0.8 [0.6,1]	0.8 [0.6,1]	Age
0.8 [0.8,1.1]	0.8 [0.7,1]	1 [0.8,1.2]	1.1 [1,1.2]	0.9 [0.8,1.1]	0.9 [0.7,1.1]	Educational level
0.9 [0.8,1.1]	1.2 [1,1.3]	0.8 [0.6,0.9]	0.8 [0.7,0.9]	0.9 [0.7,1.1]	0.8 [0.7,1.1]	Age at sexual debut
2.1 [1.7,2.5]	1.8 [1.6,2.2]	1.8 [1.6,2.2]	1.6 [1.4,1.9]	2.2 [1.5,2.6]	1.8 [1.6,2.2]	# Recent female sex partners
0.5 [0.1,0.8]	0.9 [0.6,1.2]	1.1 [0.9,1.4]	1.7 [1.4,2]	1.2 [0.9,1.8]	1.2 [1,1.5]	# Recent male sex partners
1.2 [0.7,1.6]	0.7 [0.5,1.1]	0.8 [0.6,1]	1 [0.8,1.2]	0.6 [0.4,1.2]	1.7 [1.1,2.4]	Circumcised
0.8 [0.5,1.3]	1.7 [1.2,2.2]	1.5 [1.2,2]	1.1 [0.9,1.4]	1.8 [1.2,2.7]	2.3 [1.6,3.2]	Current smoker
0.4 [0.3,0.6]	0.5 [0.3,0.7]	0.8 [0.6,1.2]	0.6 [0.5,0.8]	0.6 [0.4,1]	0.4 [0.3,0.7]	Steady sex partner
0.8 [0.5,1.1]	0.8 [0.5,1.1]	0.8 [0.6,1.1]	0.9 [0.5,1.1]	1.3 [1,1.7]	0.5 [0.2,0.7]	Consistent condom use
0.3 [0.2,0.6]	0.3 [0.1,0.5]	0.3 [0.2,0.5]	0.6 [0.4,0.9]	0.7 [0.4,1.3]	0.9 [0.4,1.5]	Mexico
1 [0.6,1.8]	2.5 [1.1,3.3]	0.8 [0.6,1.3]	1.4 [0.8,2]	0.5 [0.3,0.9]	1.3 [0.7,2.7]	Brazil
HPV16	**HPV62**	**HPV51**	**HPV89**	**HPV6**	**HPV84**	

Fig. 4. Inferred effect of host-specific covariates (maximum likelihood parameter estimate and the 95% confidence interval). Colored cells denote statistically significant positive (blue) and negative (red) effects.

(1, 3), the prevalence of genital HPV is higher in men (3). Acquired immunity has been proposed to explain declining cervical HPV prevalence by age in women in some countries (47, 48), whereas HPV prevalence does not change with age in men (3, 36). Furthermore, the seroprevalence of some types is higher in women than in men from the same source population (49), although homologous protection in women is still likely to be weak (20).

The high risk of recurring infection is consistent with either autoinoculation across anatomic sites or episodic reactivation of latent infection. Studies showing type-concordant HPV infection across anatomic sites support the importance of autoinoculation. First, HPV DNA has been detected on the fingers of patients with genital warts (50), suggesting transmission between the hands and the genitals. Second, an analysis of a subcohort of the HIM Study showed a 1.5- to 15-fold increase in the risk of anal HPV infection after genital infection (38), and a similar analysis of women found a 20-fold increase in the risk of recurrent anal HPV infection given infection at the cervix (51). Moreover, 63% of the cases of recurrent anal HPV occurred without a self-reported history of anal sex (51). Significantly, the magnitude of the effect in these studies is consistent with our estimate that previous infection yields a 20-fold increase in the risk of subsequent infection.

Apparent reinfection could instead be caused by the reactivation of latent virus. Whether HPV persists latently remains unknown (17, 52), but an animal model of cottontail oral papillomavirus provides evidence of a latent reservoir (53). Furthermore, sexually inactive HIV-positive women have a higher risk of recurrent HPV infection compared with HIV-negative controls (54), and among HIV-positive females, higher CD4$^+$ T-cell counts are negatively correlated with recurrent infection risk (55). The latter two studies have been interpreted as evidence of reactivation of latent infection (52), but either effect may instead result from increases in susceptibility to autoinoculation. Furthermore, the studies in HIV-positive women suggest that reactivation requires immune suppression and may therefore be rare in healthy individuals.

Vaccine efficacy trials suggest that reinfection is more common than reactivation. Women with previous HPV-related disease who received the quadrivalent vaccine were protected against HPV-related lesions (56). Similarly, one study showed 100% efficacy of the HPV vaccine against HPV-related disease in individuals with serological evidence of past HPV6, -11, -16, or -18 infection (57). HPV vaccines prevent infection by inducing antibodies that block viral entry, and such antibodies would likely not prevent reactivation. If indeed vaccines do not affect the disease course of reactivated infections, then these trials suggest that the vaccine diminishes the incidence of HPV lesions by blocking true reinfection. Vaccinating children before sexual debut clearly reduces infection rates (4), but if vaccination prevents autoinoculation, then vaccinating previously infected individuals may also reduce HPV prevalence.

The type-specific effects of demographic and behavioral risk factors that we observed suggest that modest differences exist in the host subpopulations supporting each HPV type. For all types, the major determinant of infection risk is the number of recent female sex partners, suggesting a central role for heterosexual transmission. Additionally, current smoking increases risk in most types, consistent with other studies (3, 29). Smoking can suppress mucosal and cellular immunity (58), but its effects may be confounded with other high-risk behaviors that were not documented in the HIM Study. Although shared risk factors account for most of the initial infection risk for each type, there are important distinctions too. For instance, circumcision has a type-specific effect on the risk of HPV infection. Randomized, controlled trials of male circumcision have

shown that circumcision protects against any HPV infection (59), whereas observational studies have documented either no effect or increased risk (3, 29). These conflicting results may have arisen from differences between types, potentially reflecting adaptation to different epithelial environments in the male genital tract.

Our modeling approach has several limitations: notably, that we cannot distinguish between reactivation and autoinoculation and that unmeasured covariates could have affected our results. While we identified celibate individuals based on detailed self-reporting on anal, vaginal, and oral sex, celibacy did not account for all forms of nonpenetrative sexual contact, such as kissing and masturbation. The construction of covariates from risk factor data, a necessary feature of any quantitative study of individual-level risk factors, may have led to bias (*SI Appendix*), and we cannot completely rule out the possibility that unobserved properties of the sexual contact network affected our results. In addition, our model of homologous immunity assumes that protection arises immediately after infection, but protection may be lagged (32). The poor fit of the homologous immunity model nevertheless suggests that any such effects are weak, but future investigations should allow for more complex forms of homologous immunity. We lack information on genetic differences in host susceptibility and on HPV exposure or changes in risk factors that occurred before the study. For example, we would have little power to detect immunity if a large portion of individuals developed sterilizing immunity before entering the study. Finally, the data track genital HPV infection, ignoring other sites.

These caveats notwithstanding, the strong effect of initial infection on subsequent infection risk that we inferred is important for the design of epidemiological studies and models to inform public health policy. Our results imply that the HPV vaccine has the potential to lower HPV prevalence far more than previously expected.

Materials and Methods

Data. The HIM Study is a multinational cohort study of HPV infection in unvaccinated men ages 18–70 y old. The study enrolled 4,123 participants between 2005 and 2009 and tracked men longitudinally over 5 y. Men were recruited from three cities: Tampa, Florida; Cuernavaca, Mexico; and Sao Paulo, Brazil. Detailed study methods are described elsewhere (36). The human subjects committees of the University of South Florida, Tampa, Florida; Centro de Referência e Treinamento em Doenças Sexualmente Transmissíveis e AIDS, Sao Paulo, Brazil; and Instituto Nacional de Salud Pública de México, Cuernavaca, Mexico approved all HIM Study procedures. All participants provided written informed consent. The data for each individual consist of binary time series of infection status with each HPV type for up to 10 clinic visits (median = 10 visits). More information is in *SI Appendix*.

Model of HPV Dynamics. Our models are two-state discrete time partially observed Markov processes, in which an individual is either infected (one) or uninfected (zero) at any time with an HPV type. Infection of individual i with HPV type j occurs at rate λ_{ijt}, such that the probability of infection over $\triangle t$ is

$$p(1_{t+\triangle t}|0_t) = 1 - e^{-\lambda_{ijt}\triangle t}. \qquad [5]$$

Because we cannot directly observe infection events, the measurement model relates each observation Y_{ijt} to the latent state X_{ijt}:

$$Y_{ijt} \sim \text{Bernoulli}(X_{ijt}(1 - P_{FP}) + (1 - X_{ijt})P_{FN}), \qquad [6]$$

where P_{FP} and P_{FN} are the rates of false positives and false negatives, respectively. The duration of infection was drawn from a gamma distribution, where the shape (k_j) and scale (θ_j) were fixed to match the empirical distribution of infection durations in the data for type j (*SI Appendix*, Fig. S8 and Table S7).

Statistical Inference and Parameter Estimation. We used iterated filtering (maximum likelihood via iterated filtering) (40) implemented in the R

package panelPomp (version 0.3.1), an extension of the pomp package (60). Details are in *SI Appendix*.

Data Accessibility. The data and code are available at https://github.com/cobeylab/HPV-model.

1. Giuliano AR, et al. (2015) EUROGIN 2014 roadmap: Differences in human papillomavirus infection natural history, transmission and human papillomavirus-related cancer incidence by gender and anatomic site of infection. *Int J Cancer* 136:2752–2760.
2. Weinstock H, Berman S, Cates W (2004) Sexually transmitted diseases among American youth: Incidence and prevalence estimates, 2000. *Perspect Sex Reprod Health* 36:6–10.
3. Han JJ, et al. (2017) Prevalence of genital human papillomavirus infection and human papillomavirus vaccination rates among US Adult men. *JAMA Oncol* 3:810–816.
4. Markowitz LE, et al. (2013) Reduction in human papillomavirus (HPV) prevalence among young women following HPV vaccine introduction in the United States, National Health and Nutrition Examination Surveys, 2003-2010. *J Infect Dis* 208:385–393.
5. Joura EA, et al. (2015) A 9-valent HPV vaccine against infection and intraepithelial neoplasia in women. *N Engl J Med* 372:711–723.
6. Lipsitch M, O'Hagan JJ (2007) Patterns of antigenic diversity and the mechanisms that maintain them. *J R Soc Interface* 4:787–802.
7. Taylor BS, Sobieszczyk ME, McCutchan FE, Hammer SM (2008) The challenge of HIV-1 subtype diversity. *New Engl J Med* 358:1590–1602.
8. Cobey S (2014) Pathogen evolution and the immunological niche. *Ann NY Acad Sci* 1320:1–15.
9. Mackinnon MJ, Marsh K (2010) The selection landscape of malaria parasites. *Science* 328:866–871.
10. de Villiers EM, Fauquet C, Broker TR, Bernard HU, zur Hausen H (2004) Classification of papillomaviruses. *Virol* 324:17–27.
11. Buck CB, Day PM, Trus BL (2013) The papillomavirus major capsid protein L1. *Virol* 445:169–174.
12. Passmore JaS, Williamson Al (2016) Host immune responses associated with clearance or persistence of human papillomavirus infections. *Curr Obstet Gynecol Rep* 5:177–188.
13. Scott M, Nakagawa M, Moscicki AB (2001) Cell-mediated immune response to human papillomavirus infection. *Clin Diagn Lab Immunol* 8:209–220.
14. Nakagawa M, et al. (2000) Persistence of human papillomavirus type 16 infection is associated with lack of cytotoxic T lymphocyte response to the E6 antigens. *J Infect Dis* 182:595–598.
15. Piersma SJ, et al. (2007) High number of intraepithelial CD8+ tumor-infiltrating lymphocytes is associated with the absence of lymph node metastases in patients with large early-stage cervical cancer. *Cancer Res* 67:354–361.
16. Stanley M (2006) Immune responses to human papillomavirus. *Vaccine* 24(Suppl 1):16–22.
17. Ryser MD, Gravitt PE, Myers ER (2017) Mechanistic mathematical models: An underused platform for HPV research. *Papillomavirus Res* 3:46–49.
18. Moscicki AB, et al. (2012) Updating the natural history of human papillomavirus and anogenital cancers. *Vaccine* 30:F24–F33.
19. Shew ML, et al. (2015) Episodic detection of human papillomavirus within a longitudinal cohort of young women. *J Med Virol* 87:2122–2129.
20. Beachler DC, Jenkins G, Safaeian M, Kreimer AR, Wentzensen N (2016) Natural acquired immunity against subsequent genital human papillomavirus infection: A systematic review and meta-analysis. *J Infect Dis* 213:1444–1454.
21. Pamnani SJ, et al. (2016) Sequential acquisition of anal human papillomavirus (HPV) infection following genital infection among men who have sex with women: The HPV infection in men (HIM) study. *J Infect Dis* 214:1180–1187.
22. Lu B, et al. (2012) Prevalent serum antibody is not a marker of immune protection against acquisition of oncogenic HPV16 in men. *Cancer Res* 72:676–685.
23. Kahn Ja, et al. (2012) Vaccine-type human papillomavirus and evidence of herd protection after vaccine introduction. *Pediatrics* 130:e249–e256.
24. Wheeler CM, et al. (2012) Cross-protective efficacy of HPV-16/18 AS04-adjuvanted vaccine against cervical infection and precancer caused by non-vaccine oncogenic HPV types. *Lancet Oncol* 13:100–110.
25. Durham DP, Poolman EM, Ibuka Y, Townsend JP, Galvani AP (2012) Reevaluation of epidemiological data demonstrates that it is consistent with cross-immunity among human papillomavirus types. *J Infect Dis* 206:1291–1298.
26. Murall CL, McCann KS, Bauch CT (2014) Revising ecological assumptions about human papillomavirus interactions and type replacement. *J Theor Biol* 350:98–109.
27. Rositch AF, et al. (2012) Vaccine-relevant human papillomavirus (HPV) infections and future acquisition of high-risk HPV types in men. *J Infect Dis* 206:669–677.
28. Chaturvedi AK (2005) Prevalence and clustering patterns of human papillomavirus genotypes in multiple infections. *Cancer Epidemiol Biomarkers Prev* 14:2439–2445.
29. Giuliano AR, et al. (2009) Circumcision and sexual behavior: Factors independently associated with human papillomavirus detection among men in the HIM study. *Int J Cancer* 124:1251–1257.
30. Svare EI, et al. (2002) Risk factors for genital HPV DNA in men resemble those found in women: A study of male attendees at a Danish STD clinic. *Sex Transm Infect* 78:215–218.
31. Matthijsse SM, et al. (2015) The role of acquired immunity in the spread of human papillomavirus (HPV): Explorations with a microsimulation model. *PLoS One* 10:1–14.
32. Murall CL, Bauch CT, Day T (2015) Could the human papillomavirus vaccines drive virulence evolution? *Proc R Soc B Biol Sci* 282:20141069.
33. Orlando PA, Gatenby RA, Giuliano AR, Brown JS (2012) Evolutionary ecology of human papillomavirus: Trade-offs, coexistence, and origins of high-risk and low-risk types. *J Infect Dis* 205:272–279.
34. Durham DP, et al. (2016) National- and state-level impact and cost-effectiveness of nonavalent HPV vaccination in the United States. *Proc Natl Acad Sci USA* 113:5107–5112.
35. Brisson M, et al. (2016) Population-level impact, herd immunity, and elimination after human papillomavirus vaccination: A systematic review and meta-analysis of predictions from transmission-dynamic models. *Funding Can Institutes Heal Res* 1:8–17.
36. Giuliano AR, et al. (2008) The human papillomavirus infection in men study: Human papillomavirus prevalence and type distribution among men residing in Brazil, Mexico, and the United States. *Cancer Epidemiol Biomarkers Prev* 17:2036–2043.
37. Albero G, Castellsagué X, Lin H (2014) Male circumcision and the incidence and clearance of genital human papillomavirus (HPV) infection in men: The HPV infection in men (HIM) cohort study. *BMC Infect Dis* 14:75.
38. Nyitray AG, et al. (2015) The natural history of genital human papillomavirus among HIV-negative men having sex with men and men having sex with women. *J Infect Dis* 212:202–212.
39. Keeling MJ, Pejman R (2008) Formulating the deterministic SIR model. *Modeling Infectious Diseases in Humans and Animals* (Princeton Univ Press, Princeton), 2nd Ed, pp 15–52.
40. Ionides EL, Bretó C, King AA (2006) Inference for nonlinear dynamical systems. *Proc Natl Acad Sci USA* 103:18438–18443.
41. Hurvich CM, Tsai CL (1989) Regression and time series model selection in small samples. *Biometrika* 76:297–307.
42. Draper E, Bissett S, Howell-Jones R, Edwards D (2011) Neutralization of non-vaccine human papillomavirus pseudoviruses from the A7 and A9 species groups by bivalent HPV vaccine sera. *Vaccine* 29:8585–8590.
43. Kemp TJ, et al. (2011) HPV16/18 L1 VLP vaccine induces cross-neutralizing antibodies that may mediate cross-protection. *Vaccine* 29:2011–2014.
44. Lemieux-Mellouki P, et al. (2016) Assortative mixing as a source of bias in epidemiological studies of sexually transmitted infections: The case of smoking and human papillomavirus. *Epidemiol Infect* 144:1490–1499.
45. Malagón T, Lemieux-Mellouki P, Laprise JF, Brisson M (2016) Bias due to correlation between Times-at-risk for infection in epidemiologic studies measuring biological interactions between sexually transmitted infections: A case study using human papillomavirus type interactions. *Am J Epidem* 184:873–883.
46. Van de Velde N, Brisson M, Boily MC (2010) Understanding differences in predictions of HPV vaccine effectiveness: A comparative model-based analysis. *Vaccine* 28:5473–5484.
47. Bosch FX, Castellsagué X, de Sanjosé S (2008) HPV and cervical cancer: Screening or vaccination? *Br J Cancer* 98:15–21.
48. Safaeian M, et al. (2010) Epidemiological study of anti-HPV16/18 seropositivity and subsequent risk of HPV16 and -18 infections. *J Natl Cancer Inst* 102:1653–1662.
49. Markowitz L, Sternberg M, Dunne E, McQuillan G, Unger ER (2009) Seroprevalence of human Papillomavirus types 6, 11, 16, and 18 in the United States: National Health and Nutrition Examination survey 2003-2004. *J Infect Dis* 200:1059–1067.
50. Sonnex C, Strauss S, Gray JJ (1999) Detection of human papillomavirus DNA on the fingers of patients with genital warts. *Sex Transm Infect* 75:317–319.
51. Goodman MT, et al. (2010) Sequential acquisition of human papillomavirus (HPV) infection of the anus and cervix: The Hawaii HPV cohort study. *J Infect Dis* 201:1331–1339.
52. Gravitt PE (2012) Evidence and impact of human papillomavirus latency. *Open Virol J* 6:198–203.
53. Maglennon GA, McIntosh P, Doorbar J (2011) Persistence of viral DNA in the epithelial basal layer suggests a model for papillomavirus latency following immune regression. *Virology* 414:153–163.
54. Theiler RN, et al. (2010) High-risk human papillomavirus reactivation in human immunodeficiency virus-infected women: Risk factors for cervical viral shedding. *Obstet Gynecol* 115:1150–1158.
55. Strickler HD, et al. (2005) Natural history and possible reactivation of human papillomavirus in human immunodeficiency virus-positive women. *J Natl Cancer Inst* 97:577–586.
56. Joura EA, et al. (2012) Effect of the human papillomavirus (HPV) quadrivalent vaccine in a subgroup of women with cervical and vulvar disease: Retrospective pooled analysis of trial data. *Br Med J* 344:e1401.
57. Olsson SE, et al. (2009) Evaluation of quadrivalent HPV 6/11/16/18 vaccine efficacy against cervical and anogenital disease in subjects with serological evidence of prior vaccine type HPV infection. *Hum Vaccin* 5:696–704.
58. Lee J, Taneja V, Vassallo R (2012) Cigarette smoking and inflammation: Cellular and molecular mechanisms. *J Dent Res* 91:142–149.
59. Gray RH, et al. (2010) Male circumcision decreases acquisition and increases clearance of high-risk human papillomavirus in HIV-negative men: A randomized trial in Rakai, Uganda. *J Infect Dis* 201:1455–1462.
60. King AA, Nguyen D, Ionides EL (2016) Statistical inference for partially observed Markovprocesses via the R package pomp. *J Stat Softw* 69:1–43.