Increased Risk of High-Grade Anal Neoplasia Associated with a Human Papillomavirus Type 16 E6 Sequence Variant

Maria M. Da Costa,1,2 Charissa J. Hogeboom,3 Elizabeth A. Holly,3 and Joel M. Palefsky1,2

Expression of the E6 and E7 genes of human papillomavirus (HPV) type 16 have been implicated in the etiology of anogenital premalignant and malignant lesions. To evaluate whether variations in the HPV-16 E6 sequence were related to the incidence of high-grade anal neoplasia, 628 HPV-16–positive anal specimens from 193 human immunodeficiency virus (HIV)–positive and 59 HIV-negative participants were typed for variations in 15 E6 nucleotide positions. Although most participants were infected with a prototype strain, 15 (6%) carried the G131 variant, and 12 (5%) were infected with the Af1a variant. Two new variants not previously reported were identified as well. An elevated risk for high-grade anal squamous intraepithelial lesions was associated with infection by G131 variants, compared with the prototype strain (odds ratio, 3.4; 95% confidence interval, 1.1–10), after controlling for HIV status. These data provide further evidence for HPV strain variation as a factor in determining the natural history of anogenital neoplasia.

Infection with human papillomavirus (HPV) is a primary risk factor for malignancies of anogenital epithelial tissues. High-risk HPV types have been implicated in the etiology of cervical cancer [1–3] and its precursor, cervical squamous intraepithelial lesions (CSILs) [2, 4], and anal cancer [5, 6] and its putative precursor, anal squamous intraepithelial lesions (ASILs) [7–10]. However, considerable variation is present in the HPV genome, and not all HPV types are equally associated with high-grade squamous intraepithelial lesions (HSILs). More than 100 HPV types have been identified, which, by definition, differ from one another by ≥10% of their DNA sequences [11]. Of the HPV types conferring the highest risk for HSILs and invasive cancer, HPV-16 and -18 are the most strongly associated [4, 8–10, 12–17], with HPV-16 infection being the most common [15, 18].

The exact mechanism of HPV-induced malignant transformation has not been determined. However, evidence points to a critical role for the HPV E6 and E7 genes. Transformation of keratinocytes has been shown to require E6 or E7 protein expression [19, 20]. HSILs show higher levels of expression of E6 and E7 than do low-grade squamous intraepithelial lesions (LSILs) [2], and continued expression of both these proteins is required to maintain the malignant state [21, 22]. There is evidence that the HPV-16 and -18 E6 and E7 proteins cooperate to induce malignant transformation [19, 20, 23] through their ability to bind to the cell-cycle regulating proteins p53 and retinoblastoma protein, respectively [24, 25]. HPV-16 or -18 E6 not only binds to p53 but initiates its selective degradation via the ubiquitin pathway [26, 27]. The E6 and E7 proteins from low-risk HPV types do not share this capability with the same level of efficiency [28, 29].

Amino acid differences in the E6 and E7 proteins are likely to be at least partly responsible for the differential pathogenicity of the various HPV types. However, the amino acid sequences of these proteins vary not only across HPV types but within type as well. More than 25 sequence variants of HPV-16 have been reported, and a phylogeny-based classification incorporating the L1, L2, and E6 genes has been proposed by Yamada et al. [30]. Differential pathogenicity of HPV-16 variants has been reported for cervical HSILs [31–33], invasive cervical cancer [34], and anal HSILs [35]. In a study of variants in the L1 region, Hildesheim et al. [36] reported that detection of non-European variants was associated with increased risk of cervical cancer [36]. It remains unclear which HPV-16 strain variants confer the highest risk of high-grade ASILs or CSILs and cancer. To investigate this further, we analyzed 628 HPV-16–positive anal specimens collected from 252 men and women over several years of follow-up in natural history studies of ASILs. For this analysis, we characterized the distribution of 17 E6 variations across race, sex, and time. Using a novel method of E6 sequence classification, we describe the relationship between HPV-16 E6 sequence variations and anal HSILs among our study participants.

Patients, Materials, and Methods

Study participants and specimens. Samples were obtained from participants in 2 different cohort studies on the natural history of...
ASILs. The men were participants in a long-term cohort study that was initiated to investigate the natural history of ASILs [8, 10]; the women were participants at the San Francisco Bay Area site of the Women’s Interagency HIV Study [37] who also agreed to participate in a study in the natural history of ASILs in women [38, 39].

Men and women participants were given a physical exam at specified intervals over the course of these prospective studies, as described elsewhere [8, 10, 38]. Anal swab samples were obtained for anal HPV testing and anal cytology. High-resolution anoscopy was performed to visualize the anal lesions, and biopsies were obtained for histopathological evaluations as described elsewhere [40]. All anal swab specimens shown to be positive for HPV-16 DNA were studied in more detail for E6 nucleotide sequence variation, as described below.

**Polymerase chain reaction (PCR).** A single-tube, nested E6 PCR was done, as described elsewhere [39]: 100-μL amplification reactions were performed with 10 mM Tris (pH 8.3), 50 mM KCl, 2.5 mM MgCl$_2$, 200 μM each dNTP, 2 μM each internal primer (CGTAACCGAAAATGCCTGAAC and GGACATCTATTCATCCTC), and 2 μM each internal primer (GGTCAAGACAGG and ACCGGTATAGATAAAAG), and 2.5 U of AmpliTaq (Applied Biosystems). To each reaction mix was added 3 μL of crude DNA preparation. PCR was performed on a thermocycler (Uno II; Biometra); conditions were 95°C for 3 min (hotstart); 30 cycles of 94°C for 30 s, 57°C for 30 s, and 72°C for 1 min; 35 cycles of 94°C for 30 s, 57°C for 30 s, and 72°C for 1 min; and a terminal extension of 72°C for 10 min.

Each set of PCRs performed contained positive controls consisting of SiHa cells (1–2 copies of P-G variant HPV-16/cell) and plasmids containing the nested E6 fragments for variants P-T, P-G, G131G, C109T, C109G, Af1a, Af2a, or Na1. Negative controls in each PCR were P3HR lymphoblastoid cells or water.

**Dot-blot hybridization.** The HPV-16 variants were typed, using the allele-specific oligonucleotide system of Wheeler et al. [41]; this system allowed samples to be tested in a 96-well plate format in a cost-effective manner. Samples were dot-blotted, hybridized, and washed, as described elsewhere [41]. In brief, 70 μL of ampiclon was added to 850 μL of 0.4 N NaOH/25 mM EDTA. Aliquots of 70 μL were loaded on Biodyne B membrane (Life Technologies) in a 96-well vacuum manifold, and the wells were then washed with 200 μL 20× SSPE (3.6 M NaCl, 0.2 M NaH$_2$PO$_4$, H$_2$O, 0.11 M NaOH, and 0.02 M EDTA). The DNA was bound by baking the membranes for 1 h in an 80°C vacuum oven.

Twenty-six E6 probes were used for hybridization [41]. Membranes were pretreated at 1 h at 65°C in 0.1× SSPE–0.5% SDS with shaking. Membranes were hybridized overnight in 5× SSPE–0.1% SDS with biotinylated probes at a concentration of 1 pmol/μL. Membranes were washed twice for 10 min in 2× SSPE–0.1% SDS with vigorous agitation. Wash temperatures were 45°C, except for probes 131/2AG, 131/2AT, 350T, 350G, 403A, and 403G, which were washed at 47°C. Membranes were then incubated with gentle agitation at room temperature for 15 min in 37°C 2× SSPE–0.1% SDS containing 30 ng/mL streptavidin-horseradish peroxidase conjugate (Vector Laboratories). Blots were washed twice (10 min each) with mild agitation at room temperature with 37°C 2× SSPE–0.1% SDS. Back-to-back paired membranes were incubated in enhanced chemiluminescence substrate reagents (Amersham) at room temperature for 1 min. The blots were exposed to Hyperfilm MP (Amersham) for 3 min. Each well was scored as positive, negative, or equivocal.

Strongly and weakly positive probe readings were considered to be a positive reaction. When there was evidence for background binding of probes, as defined by ≥2 nucleotides testing positive at a given position for a control sample, that set was reprobed to distinguish samples with a false-positive reaction from those with mixtures of variants. Samples that appeared to contain a previously unknown variant or that contained repeated ambiguity at a nucleotide position over multiple visits were sequenced. In total, 13 samples were sequenced. Samples with ≤2 of the 26 probes testing positive were considered to be false positives for HPV-16 and were excluded from analysis, with false positivity possibly resulting from binding to other HPV types, such as HPV-31 and -58 [41].

**Nucleotide sequence identification.** The nested E6 fragment of chosen samples was gel-purified and sequenced by double-stranded PCR sequencing, using the internal nested primers. First the PCR was repeated, and the entire ampiclon was run on an agarose gel. The PCR product was excised from the gel, placed in a spin-X tube (Costar), and frozen at −70°C for ≥1 h. Samples were spun in a microfuge at 14 K for 10 min, and the filter unit of the tube was discarded. Eluates were precipitated with NH$_4$OAc/EtOH, spun, decanted, washed with 70% ethanol, spun, decanted, and air-dried. Samples were suspended in 25 μL of water and sequenced by the Biomolecular Resource Center at University of California, San Francisco, using automated sequencers (PE Biosystems).

**E6 variant classification.** Results of the 26 probes were converted to a 15-base nucleotide “sequence” by a modified version of an SAS program (provided by K. Altobelli, University of New Mexico, Albuquerque). Each sample was assigned a variant name and classified into a lineage on the basis of the work of Yamada et al. [30, 42]. In the Yamada classification system, a phylogenetic tree suggests that 5 previously identified HPV-16 lineage groups (designated AA [Asian American], Af1 [African-1], Af2 [African-2], As [Asian], and E [European]) could be broadly grouped into 2 larger classes on the basis of sequence similarities, the European (E) class, containing lineages E and As, and a non-European (A) class, containing the AA, Af1, and Af2 lineages. These lineages and the variants that are grouped into them are presented in figure 1, in which it can be seen that 3 single nucleotide polymorphisms (SNPs) among the 26 we tested are invariant within lineage but differ across lineages (nucleotides 145, 286, and 289). This characteristic allowed us to impute the broad lineage category (E or A) from data at positions 145, 286, and 289 when the specific variant could not be ascertained for a given sample because data on ≥1 SNPs were missing or ambiguous.

In addition to phylogenetic-based classification of nucleotide variations, additional analysis was based on variations in the inferred amino acid sequence. The 15 nucleotides tested correspond to 12 amino acids in the translated protein. Among the SNPs tested, 5 were silent when translated to protein (figure 1). The amino acid sequence is numbered so that residue 1 corresponds with nucleotides 104–6.

Because of the large number of E6 sequence variations and the uneven distribution of sequence variants across cohort and human immunodeficiency virus (HIV) status, the use of the E6 variant itself or each individual nucleotide or amino acid substitution as a
predictor was not possible. Initial exploratory analysis therefore used a classification and regression tree (CART) [43] to indicate where the strongest relationships might lie. Insights gained from the CART analysis were used to structure further analyses by reducing E6 sequence variations to 3 polymorphic amino acids.

**Statistical analysis.** Data consisted of all samples obtained over the course of observation of each study participant. To control for the variable number of samples, all analyses used summary variables that condensed the repeated observations over time for each participant. Because some participants were observed more frequently or for a longer time period than others, analyses also considered the number of visits or length of observation time.

Data analysis focused on 2 areas: (1) characterization of E6 variant distribution across race, sex, and time, and (2) assessment of the

---

**Figure 1.** Classification of human papillomavirus type 16 E6 variants and distribution among all study participants. *Amino acids are numbered under the assumption that translation begins at nucleotide 104. **Includes multiple samples per participant. ***Variant classes: AA, Asian American; Af1, African-1; Af2, African-2; As, Asian; E, European; NA1, North American-1.
potential association of the E6 variant with the development or progression of ASILs. In analyses that focused on E6 variant distribution, participants who had > 1 variant were classified further as to whether the variants belonged to a single class or lineage versus multiple classes or lineages. The distribution of E6 variants across participants and their persistence over time were characterized by simple descriptive statistics. Analysis of the relationship between the E6 variant and the race and sex of the study participants used only the broad lineage classification (E, A, or mixed/unknown) to categorize each participant. The potential association of E6 lineage (E vs. A) with race or sex was evaluated by use of Pearson’s χ² or Fisher’s exact test, depending on the sample size.

For the analysis of the relationship between the E6 variant and anal lesions, E6 variants were classified into a small number of categories on the basis of 3 polymorphic amino acids that accounted for most of the sequence variation. Participants who had E6 variants that fell into ≥ 1 of these categories were excluded from this analysis. The grade of anal lesion at each visit was defined as the highest grade observed by either cytology or histology among the samples obtained at that visit. The evolution of anal lesions over time for each participant was summarized by considering: (1) the grade of anal lesion at the participant’s baseline visit, (2) the highest grade of anal lesion ever seen during the course of observing the participant, (3) the highest grade of anal lesion observed prior to the first evidence of HPV-16 infection (if available), and (4) the highest grade of anal lesion observed at or after the first observation of HPV-16 infection.

Since not all participants were observed prior to their first exposure to HPV-16 or their first abnormal anal sample, we considered the following questions: (1) among participants with normal anal cytology and no sign of ASILs at high-resolution anoscopy prior to the first evidence of HPV-16 infection, were any E6 variants predictive of subsequent development of anal cytologic or histologic abnormalities; (2) among all participants, were any E6 variants predictive of progression to a higher grade of anal lesion subsequent to the first evidence of HPV-16 infection; and (3) were participants who developed ASILs (LSILs or HSILs) anytime during the course of observation more likely to be infected with one E6 variant than with another? The potential association of the E6 variant with anal lesions was assessed by incorporating the variant categories in a logistic model that controlled for cohort, HIV status, and number of visits. In cases where the E6 variant was potentially confounded by other factors, stratified analyses also were done.

We performed additional analyses that attempted to replicate, as closely as possible, the analyses reported by others in order to compare results. In an analysis modeled after that of Xi et al. [32, 35], we classified E6 variants as either “prototype” (including P-T, P-G, G131, and others) or “non-prototype” (including the AF1, AF2, and AA variants). These categories were entered into a Cox proportional hazards regression model for the incidence of anal HSILs. Additional predictors that were incorporated as appropriate included HIV status, CD4 cell count, baseline abnormality, infection with other HPV types, age, race, number of sex partners, history of anal intercourse, and number of visits. Analyses modeled after that of Ethington et al. [33] selected all participants who had an abnormal cytologic test result and who also were infected with a G131 or P350 E6 variant and had a biopsy within 6 months of the first abnormal cytologic test result. Odds ratios (ORs) were computed for the association between the E6 variant and an HSILs biopsy at the time of the first abnormal cytology and also at any time during the course of observation. Tests of association and logistic regressions were performed with SAS software (version 6.12; SAS Institute). Exploratory analyses using CART were performed with the tree function in S-plus (Insightful).

Results

In total, 634 anal specimens from 254 participants from our combined cohort of men and women were determined by PCR to be positive for HPV-16. Each of these samples was typed for E6 nucleotide sequence variation by hybridization with 26 probes that tested for variations in 15 SNPs. Six samples did not test positive for > 2 of 26 probes, were considered to be false positives for HPV-16, and were eliminated from analysis, leaving 628 anal samples from 211 men and 41 women. The 2 cohorts differed significantly with respect to HIV status, racial distribution, and number of samples obtained. Of the 211 men, 154 (73%) were HIV positive, and, of the 41 women, 39 (95%) were HIV positive. Most of the men were white, non-Hispanic (91%), and most of the women were either white, non-Hispanic (44%) or African American (49%). The racial/ethnic distribution of the men and women is shown in table 1. The mean number of specimens collected per man was 2.7, and the mean number collected per woman was 1.2.

Figure 1 presents the distribution of E6 variants that were detected in anal samples from the 252 participants tested. The classification system and naming convention were adapted from Yamada et al. [30, 42]. In most cases, only 1 E6 variant was detected for each participant, and the same variant was seen when repeated anal samples were obtained over time. However, in ~15% of participants, ≥ 2 variants were seen; these participants are categorized as having a mixture of variants. More than two-thirds of the participants were infected with the prototype (reference) strain P-T (39%) or its 350G variant, P-G (27%), and 5% were infected with both. The next 2 most common variants were the G131 type (6%) and the AF1a type (5%). Two novel E6 variants were detected and confirmed by sequencing. One contained the substitutions found in both G131G and A176G, and the other incorporated a 350G substitution into the AF2a variant sequence.

The distribution of E6 variants was different for the 2 cohorts. African variants were more common in the women, with 22% of the women but only 1% of the men carrying the AF1a variant. Conversely, although the prototype strains (P-T or P-G) were seen in three-quarters of the men, only half the women were infected with 1 of these strains. Cohort and racial differences are summarized in table 1. Most (93%) of the men carried an E variant class, whereas only two-thirds of the women had an E variant. Because the cohorts differed in racial composition, data were analyzed to determine whether cohort differences in E6 variant distribution were explained by race. Among the women, 33% of the white, non-Hispanic participants and 44% of the
African Americans were infected with A types \( (P = .49) \). Comparable figures for the men were not available, since only 2 men were African American. Among white, non-Hispanic participants, 33% of women but only 7% of men were infected with A types \( (P = .002) \). These findings indicate that the distribution of non-prototype variants in our sample varied by sex or by other differences between the cohorts rather than by ethnicity of the participant.

To determine whether HPV-16 variants differed in their pathogenicity, we estimated the strength of the association between the variant and the degree of ASILs. Seventeen different E6 variants were detected (figure 1); however, the pathogenicity of each individual variant could not be determined because of the large number of variants. We therefore chose 2 means to classify E6 variants for analysis. The first was to combine variants into 2 broad lineages, as shown in table 1. The second was to reduce the complexity by considering variations in the E6 protein rather than in its nucleotide sequence.

The 15 SNPs resulted in heterogeneity for only 7 of the 12 corresponding amino acid positions in the translated protein (figure 1). Furthermore, among our participants, there was little diversity at amino acid positions 25, 27, and 29, with 97%–99% of participants harboring the identical residue at these positions. Amino acids at position 10, 14, and 78 showed less identity (80%–90%), but they showed complete correlation between the residues at positions 14 and 78. Tyrosine (Y) at position 78 was always associated with aspartic acid (D) or histidine (H) at position 14, and histidine at position 78 was always associated with glutamine (Q) at position 14. Position 83 showed the highest degree of diversity, with 53% of participants harboring a variant containing leucine (L) at this position, and 39% harboring a variant containing valine (V). On the basis of these observations, we chose polymorphisms at amino acid positions 10, 14, and 83 as potential predictor variables that would capture most of the protein sequence variation in our sample.

Table 2 shows the distribution of these amino acid polymorphisms among our study participants and their corresponding variant classes. Because of the uneven distribution of amino acid polymorphisms in our sample, we were unable to test the pathogenicity of all possible variations. The primary contrasts we tested were either L or V at position 83 and RQ, GQ, or TD at positions 10 and 14.

Of the 252 study participants, 134 (53%) had normal results for anal smears prior to the first detection of HPV-16, whereas 42 (17%) demonstrated some degree of anal lesions prior to detection of HPV-16. The remaining 76 participants (30%) had both HPV-16 and some degree of anal lesions at their first visit, and for these participants it is unknown whether HPV infection preceded the abnormality. For the 134 participants who had normal smear results prior to HPV detection, we assessed whether E6 variation was predictive of later development of abnormality. For the others, we also addressed whether any E6 variations would predict progression from a lower to a higher level of abnormality and whether participants who ultimately developed LSILs or HSILs harbored different variants from those who never demonstrated a grade of lesion higher than atypia. Because of the influence of cohort membership and HIV status on both anal lesion and the distribution of E6 variants, all analyses controlled for these 2 factors either by incorporating them as predictor variables or by stratification.

Of the 134 participants, 95 (71%) who had normal baseline anal smears and RQ or GQ at positions 10 and 14.

### Table 1. Distribution of human papillomavirus type 16 E6 lineages by sex and racial/ethnic group of study participants.

| Demographic group | European classes, no. (% of total)* | Non-European classes, no. (% of total)* | Mixed or unknown classes, no. | Total no. of samples (%)
|------------------|-----------------------------------|----------------------------------------|-----------------------------|-----------------------------
| Men              | 189 (93)                          | 14 (7)                                 | 8                           | 211 (100)
| White, non-Hispanic | 172 (93)                        | 12 (7)                                 | 8                           | 192 (91)
| African American | 2 (100)                           | 0                                      | 0                           | 2 (1)
| Hispanic         | 9 (90)                             | 1 (10)                                 | 0                           | 10 (5)
| Other/unknown    | 6 (86)                             | 1 (14)                                 | 0                           | 7 (3)
| Women            | 24 (62)                            | 15 (38)                                | 2                           | 41 (100)
| White, non-Hispanic | 12 (67)                          | 6 (33)                                 | 0                           | 18 (44)
| African American | 10 (56)                            | 8 (44)                                 | 2                           | 20 (49)
| Hispanic         | 1 (100)                            | 0                                      | 0                           | 1 (2)
| Other/unknown    | 1 (50)                             | 1 (50)                                 | 0                           | 2 (5)
| Total            | 213 (88)                           | 29 (12)                                | 10                          | 252 (100)

*Percentage determination excluded the no. of mixed/unknown samples.
E6 variant was not predictive of anal abnormality. Among 29 HIV-positive women, those infected with a TD variant were significantly less likely to develop abnormal lesions than those infected with the RQ or GQ variants (OR, 0.15; 95% confidence interval [CI], 0.02–0.99) after controlling for the number of visits and L or V at position 83.

We then examined whether any E6 variations would predict progression from a lower to a higher level of abnormality and whether participants who ultimately developed LSILs or HSILs harbored different variants from those who never demonstrated a grade of lesion higher than atypia. Increase in grade of anal lesions was not predicted by any of the E6 variants nor by HIV status. However, whether a study participant ever had anal lesions was associated with both of these factors. Table 3 presents ORs for the association of E6 variations with the development of HSILs. The odds of developing HSILs were elevated for the amino acid combination 10G/14Q, corresponding to the variant G131, when compared with the prototype (OR, 3.4; 95% CI, 1.1–10). In contrast, the amino acid combination 10T/14D/83L, found in some of the Af1 variants, was negatively associated with progression of anal lesions, compared with the prototype among the HIV-positive women in our sample, although the CIs overlapped unity. The length of follow-up time was not significantly different for RQ, GQ, and TD variants, nor were there differences in baseline anal lesions or time from baseline to HSILs among carriers of these variants. Infection with HPV types other than HPV-16 was not a significant predictor of anal lesions. There was no significant difference in risk between leucine and valine at position 83 once the polymorphisms at positions 10 and 14 and the sex and HIV status of the participants were taken into account.

### Table 2. Distribution of human papillomavirus type 16 E6 amino acid polymorphisms at positions 10, 14, and 83 in human immunodeficiency virus (HIV)-infected men and women.

<table>
<thead>
<tr>
<th>Amino acid position, identity</th>
<th>Variant class(es) represented</th>
<th>Total</th>
<th>HIV+</th>
<th>HIV-</th>
<th>HIV+</th>
<th>HIV-</th>
</tr>
</thead>
<tbody>
<tr>
<td>10, 14, and 83, respectively</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RQL</td>
<td>E, As</td>
<td>104</td>
<td>58</td>
<td>30</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>RQV</td>
<td>E</td>
<td>77</td>
<td>56</td>
<td>16</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>GQL</td>
<td>E</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>GQV</td>
<td>E</td>
<td>11</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>TDL</td>
<td>Af1</td>
<td>12</td>
<td>1</td>
<td>2</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>GDV</td>
<td>Af1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>IDL</td>
<td>Af2</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>IDV</td>
<td>Af2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RDL</td>
<td>Af2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>RHV</td>
<td>NA1, AA</td>
<td>6</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Subtotal</td>
<td></td>
<td>224</td>
<td>132</td>
<td>55</td>
<td>35</td>
<td>2</td>
</tr>
<tr>
<td>Mixtures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RQL + RQV</td>
<td>E, As</td>
<td>12</td>
<td>11</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RQL + GQL</td>
<td>E, As</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RQL + GQV</td>
<td>E, As</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RQV + GQV</td>
<td>E</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TDL + IDL + RDL</td>
<td>Af</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>RHL + RHV + RQL</td>
<td>E, AA</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>RQV + TDL + possible others</td>
<td>E, Af1, indeterminate</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Undetermined variants and/or</td>
<td>combinations</td>
<td>Indeterminate</td>
<td>6</td>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Subtotal</td>
<td></td>
<td>28</td>
<td>22</td>
<td>2</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>252</td>
<td>155</td>
<td>58</td>
<td>39</td>
<td>2</td>
</tr>
</tbody>
</table>

**NOTE.** Polymorphisms are summarized over time for each participant. Amino acid identities that changed over time are reported as mixtures.

*AA, Asian-American; Af1, African-1; Af2, African-2; As, Asian; E, European; NA1, North American-1.

### Table 3. Association of human papillomavirus type 16 E6 variants with detection of anal high-grade squamous intraepithelial lesions (HSILs).

<table>
<thead>
<tr>
<th>E6 variant</th>
<th>Amino acids at positions 10 and 14, respectively</th>
<th>Highest grade of anal lesion</th>
<th>OR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>P350</td>
<td>RQ</td>
<td>HSIL</td>
<td>1.0</td>
<td>—</td>
</tr>
<tr>
<td>G131</td>
<td>GQ</td>
<td>HSIL</td>
<td>3.4</td>
<td>1.1–10</td>
</tr>
<tr>
<td>Af1a</td>
<td>TD</td>
<td>HSIL</td>
<td>0.23</td>
<td>0.03–2.0</td>
</tr>
<tr>
<td>Other E</td>
<td>RQ</td>
<td>HSIL</td>
<td>2.3</td>
<td>0.80–6.7</td>
</tr>
<tr>
<td>Other non-E</td>
<td>RH, GD, ID, RD</td>
<td>HSIL</td>
<td>2.3</td>
<td>0.72–7.2</td>
</tr>
<tr>
<td>Mix or unknown</td>
<td></td>
<td>HSIL</td>
<td>1.8</td>
<td>0.55–5.9</td>
</tr>
</tbody>
</table>

**NOTE.** Odds ratio (OR) estimates and 95% confidence intervals (CIs) were obtained from logistic regression and incorporated human immunodeficiency virus infection status and cohort membership as concurrent predictors.
We performed additional analyses of our data to compare results with other researchers. A Cox regression analysis revealed no difference in incidence of anal HSILs for participants infected with prototype-like (PL) variants than for those carrying non-prototype–like (NPL) types, as defined by Xi et al. [32, 35], after controlling for HIV status, CD4 cell count, baseline disease, and presence of other HPV types. A similar analysis in HIV-positive women revealed an elevated incidence ($P = .05$) of anal HSILs among participants infected with the PL strains, after controlling for age, race, number of sex partners, history of anal sex with men, and number of visits. This increased risk was highest for the G131 strain. When contrasted with NPL types, the relative hazard for incidence of anal HSILs among carriers of G131 in our cohort was 6.4 (95% CI, 1.2–33), whereas that for PL types other than G131 was 3.1 (95% CI, 0.75–13).

We performed a second set of analyses that included only participants with abnormal cytologic test results in an effort to replicate the analysis of Etherington et al. [33, table 1]. Considering all biopsies obtained at or after the first abnormal cytologic test results, the OR for HSILs among participants infected with G131 was 7.2 (95% CI, 0.89–58) relative to participants infected with the prototype strain.

**Discussion**

We have characterized the distribution of HPV-16 E6 variant strains present in 628 HPV-16–positive anal samples from a combined cohort of 41 women and 211 men who have sex with men. The distribution of E6 variants among our samples differed for the 2 cohorts, but it did not differ by race after adjustment for cohort. Only 1 variant was detected in the majority of participants, and samples taken over time indicated that the same variant tended to persist. Persistence of a single variant and lack of association of variant with race of the host agree with earlier reports by Xi et al. [32, 35]. Detection and persistence of an HPV-16 single variant is of interest given the multiple sexual exposures of many of our participants and the frequent detection of multiple HPV types in many specimens. This observation suggests that a dominant variant of an HPV type is established, and, through mechanisms that remain unknown, infection with or replication of the DNA of other variants is precluded.

Seventeen different E6 variants were detected among our anal samples. Efforts by other investigators to identify a variant or variant class that confers higher risk for anogenital neoplasia often have relied upon the phylogenetically based classification system used by Yamada et al. [30] to reduce the number of E6 variants to a manageable number of categories. We chose to group E6 variants on the basis of 3 amino acid polymorphisms in the E6 protein, residues 10, 14, and 83, that accounted for the majority of sequence variation among our samples.

Our analysis of the association of E6 sequence variations with anal neoplasia revealed an elevated risk of anal HSILs associated with substitutions at amino acids 10 and 14 of the HPV-16 E6 protein. Compared with participants infected with the prototype (P-T or P-G) variant, participants infected with a G131 variant showed >3-fold higher odds of developing HSILs after controlling for HIV status. This value is of a similar order of magnitude as the OR computed from the data presented by Etherington et al. [33, table 1] in their study of cervical HSILs among women with cytologic abnormalities.

The methodology used by Etherington et al. [33] differed from ours in some respects. Those investigators used a cross-sectional study design, only enrolled participants with abnormal cytologic test results, and performed biopsies on all participants. In contrast, we used a longitudinal design, and ~16% of our participants with HPV-16 E6 prototype or G131 variants had no abnormal cytologic test results, and 38% of them had no biopsy. We performed a second set of analyses that included only participants with abnormal cytologic test results in an effort to replicate the analysis by Etherington et al. [33]. These data showed an OR of 7.2 (not statistically significant [$P = .06$]) for HSILs among participants infected with G131 relative to participants infected with the prototype strain. Combined with the risk estimates of the other 2 analyses, our results suggest that further investigation of the G131 variant is warranted.

Aside from the G131 variant, we did not confirm the findings of other investigators of an elevated risk for any other HPV-16 E6 strain. Xi et al. [32, 35] classified variants as either prototype (including P-T, P-G, G131, and others) versus nonprototype (including the Af1, Af2, and AA variants). They reported an elevated risk for both anal carcinoma in situ [35] and cervical intraepithelial neoplasia [32] associated with the nonprototype strains. We found no difference in the incidence of anal HSILs for the prototype and the nonprototype strains in our combined cohort when we used a similar Cox regression analysis method. When only HIV-positive women were included in the analysis, we found an elevated risk associated with the PL variants that was most exaggerated for the G131 strain. We were unable to confirm the reports of Londesborough et al. [31] or of Zehbe et al. [34] of an elevated risk for the leucine-to-valine transition at amino acid 83. The absence of elevated risk for this amino acid polymorphism is consistent with the functional similarity between these 2 amino acids.

Conflicting findings among these investigators may be explained in part by confounding factors, some of which are difficult to measure. HPV-16–positive participants enrolled into a cohort study may have been infected for years prior to enrollment, perhaps even for decades. In some cases, chronic HPV infection may be latent, and the participant may not test positive for HPV. In addition, HPV-associated neoplastic transformation is a slow process, and a false-negative finding in the early stages of incident neoplasia may make it difficult to pinpoint the time of transition to a disease state. Thus, the time period during which study participants were observed may have covered the beginning, middle, or end of a biologic process spanning many years, making a clear picture of the causative sequence difficult to obtain. Last, coinfection with other HPV types is common, especially among HIV-positive


23. Halbert CL, Demers GW, Galloway DA. The E7 gene of human papillo-

24. Dyson N, Howley PM, Munger K, Harlow E. The human papilloma virus—

References


4. Koutsoukis LA, Holmes KK, Critchlow CW, et al. A cohort study of the risk of cervical intraepithelial neoplasia grade 2 or 3 in relation to papilloma-


5. Zaki SR, Judd R, Coffield LM, Greer P, Rolston F, Evatt BL. Human papil-


