# Transcriptome Sequencing Demonstrates that Human Papillomavirus Is Not Active in Cutaneous Squamous Cell Carcinoma

Sarah T. Arron[1,5], J. Graham Ruby[2,3,5], Eric Dybbro[1], Don Ganem[3,4] and Joseph L. DeRisi[2,3]

β-Human papillomavirus (β-HPV) DNA is present in some cutaneous squamous cell carcinomas (cuSCCs), but no mechanism of carcinogenesis has been determined. We used ultra-high-throughput sequencing of the cancer transcriptome to assess whether papillomavirus transcripts are present in these cancers. In all, 67 cuSCC samples were assayed for β-HPV DNA by PCR, and viral loads were measured with type-specific quantitative PCR. A total of 31 SCCs were selected for whole transcriptome sequencing. Transcriptome libraries were prepared in parallel from the HPV18-positive HeLa cervical cancer cell line and HPV16-positive primary cervical and periungual SCCs. Of the tumors, 30% (20/67) were positive for β-HPV DNA, but there was no difference in β-HPV viral load between tumor and normal tissue ($P=0.310$). Immunosuppression and age were significantly associated with higher viral load ($P=0.016$ for immunosuppression; $P=0.0004$ for age). Transcriptome sequencing failed to identify papillomavirus expression in any of the skin tumors. In contrast, HPV16 and HPV18 mRNA transcripts were readily identified in primary cervical and periungual cancers and HeLa cells. These data demonstrate that papillomavirus mRNA expression is not a factor in the maintenance of cuSCCs.

*Journal of Investigative Dermatology* advance online publication, 14 April 2011; doi:10.1038/jid.2011.91

## INTRODUCTION

Although 12% of all human cancers are now known to be caused by viruses (Parkin, 2006; Zur Hausen, 2009), the mere presence of viral DNA in a tumor does not necessarily indicate causality. Multiple lines of evidence suggest a viral etiology for cutaneous squamous cell carcinoma (cuSCC). In immunosuppressed solid organ-transplant recipients, the incidence of cuSCC is 65- to 250-fold higher than in the general population (Hartevelt *et al.*, 1990; Jensen *et al.*, 1999; Lindelof *et al.*, 2000); incidence ratios of this magnitude are commonly seen in other viral cancers, including human herpesvirus-8-mediated Kaposi's sarcoma and hepatitis B virus-associated hepatocellular carcinoma (Vajdic *et al.*, 2006). A second line of evidence supporting viral etiology is the behavior of the keratoacanthoma (KA) subtype of cuSCC. KA can spontaneously regress, and has been suggested to lie along a spectrum of carcinogenesis between hyperplastic viral verrucae and neoplastic SCCs (LeBoit, 2002).

Previous studies have selectively focused on human papillomavirus (HPV) as a potential etiologic agent in cuSCC. Investigators have hypothesized an analogy between cuSCC and cervical SCC, as the latter has been firmly associated with high-risk α-genus HPV (α-HPV) infection, including HPV16 and HPV18 (IARC, 2007; Bouvard *et al.*, 2009). However, different HPV types have site-specific tropism for mucosal or cutaneous epithelium; the high-risk mucosal α-HPVs are not found in cuSCCs, with the exception of genital and periungual tumors (Moy *et al.*, 1989; Alam *et al.*, 2003; Dubina and Goldenberg, 2009). Thus, many studies focus on detection of the cutaneous β-genus HPV types (β-HPV) in cuSCCs (Shamanin *et al.*, 1994, 1996; Surentheran *et al.*, 1998; Berkhout *et al.*, 2000; Harwood *et al.*, 2000; Forslund *et al.*, 2003b; Asgari *et al.*, 2008).

The association of β-HPV with cuSCC is clearly defined for a specific group of patients with epidermodysplasia verruciformis, an autosomal recessive genodermatosis associated with susceptibility to β-HPV. Patients with epidermodysplasia verruciformis develop widespread viral warts and β-HPV5- and β-HPV8-mediated SCCs (Harwood *et al.*, 1999). However, the β-HPV types have not been firmly associated with cuSCC in the general population (IARC, 2007). β-HPV DNA is detected in 27–54% of SCCs from immunocompetent patients and 55–84% of SCCs from immunosuppressed patients (Shamanin *et al.*, 1994, 1996; Berkhout *et al.*,

[1]*Department of Dermatology, University of California, San Francisco, San Francisco, California, USA;* [2]*Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, California, USA;* [3]*Howard Hughes Medical Institute, University of California, San Francisco, San Francisco, California, USA and* [4]*Department of Medicine, University of California, San Francisco, San Francisco, California, USA*

[5]*These authors contributed equally to this work.*

*Correspondence: Joseph L. DeRisi, Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, California 94143, USA. E-mail: joe@derisilab.ucsf.edu*

*Abbreviations: BGM, background model; cuSCC, cutaneous squamous cell carcinoma; HPV, human papillomavirus; KA, keratoacanthoma; mRNA-seq, mRNA-transcriptome sequencing; SCC, squamous cell carcinoma*

*Received 16 December 2010; revised 14 February 2011; accepted 25 February 2011*

1995, 2000; Harwood *et al.*, 2000; Forslund *et al.*, 2003a, 2007; Asgari *et al.*, 2008). Indeed, in other studies, β-HPV has been detected with comparable frequency in normal skin, eyebrow hairs, and premalignant actinic keratoses (Boxman *et al.*, 1997; Antonsson *et al.*, 2000; Forslund *et al.*, 2003b; Hazard *et al.*, 2007; de Koning *et al.*, 2007, 2009; Asgari *et al.*, 2008). Other studies have reported an association between antibody responses to β-HPV and the development of cuSCC, particularly for patients with antibodies to multiple HPV types (Karagas *et al.*, 2006; Bouwes Bavinck *et al.*, 2010).

High-risk α-HPV, which are present in over 95% of cervical SCCs, often integrate into the human genome and express viral proteins that interfere with normal cell cycle control (reviewed in zur Hausen, 1996; Bosch *et al.*, 2002). The E6 and E7 proteins of the high-risk α-HPV types interfere with the tumor-suppressor activities of cellular p53 and retinoblastoma protein to drive carcinogenesis (Dyson *et al.*, 1989; Scheffner *et al.*, 1990; Werness *et al.*, 1990). Ongoing expression of E6 and E7 is required for both induction and maintenance of carcinogenesis. By analogy, β-HPV would be expected to utilize the same mechanism of carcinogenesis, but studies using *in situ* hybridization or reverse transcriptase-PCR to detect HPV mRNA in cuSCC have detected viral transcripts only sporadically and at low levels in occasional tumors, with many other tumors testing negative (Purdie *et al.*, 2005; Dang *et al.*, 2006). Nevertheless, many authors continue to point to β-HPV as a possible etiologic agent in these tumors.

The goal of our study was to assess whether β-HPV is capable of causing cuSCC through expression of viral oncogenes using ultra-high-throughput sequencing of the SCC transcriptome. This comprehensive, unbiased analysis of total tumor mRNA expression revealed no HPV transcriptional activity, an observation that was further supported by the absence of a substantial viral load in the tumors. These two observations contradict the hypothesis that transcription of viral oncogenes is required for tumor maintenance.

## RESULTS

### Patient characteristics

We enrolled 38 patients, including 27 males and 11 females, ranging in age from 41 to 95 years (Table 1). Of these patients, 17 were immunocompetent and 21 were immunosuppressed because of solid organ transplantation, hematologic malignancy, HIV, or medication for Wegener's granulomatosis. A total of 89 tissue samples were collected from these patients, including 71 SCCs (23 KA subtype) and 18 normal skin samples. Four tumor samples did not yield enough tissue for DNA extraction but RNA was obtained for transcriptome analysis. Two α-HPV16-mediated primary tumors were obtained for comparison: a periungual SCC from a 53-year-old immunocompetent man and a stage I nonkeratinizing cervical SCC from a 35-year-old woman. The α-HPV18-mediated HeLa cervical cancer cell line was used as an additional control.

**Table 1. Demographics**

| | Total | Male | Female |
|---|---|---|---|
| *Patients* | 38 | 27 | 11 |
| Immunocompetent | 17 | 10 | 7 |
| Immunosuppressed | 21 | 17 | 4 |
| Solid organ transplant | 12 | 9 | 3 |
| Hematologic malignancy | 6 | 5 | 1 |
| HIV | 2 | 2 | 0 |
| Wegener's granulomatosis | 1 | 1 | 0 |
| Age range (years) | 41–95 | 41–95 | 52–95 |
| *SCC samples (KA type)* | 71 (23) | 45 (8) | 26 (15) |
| Immunocompetent | 30 (16) | 14 (6) | 16 (10) |
| Immunosuppressed | 41 (7) | 31 (2) | 10 (5) |
| Solid organ transplant | 28 | 19 | 9 |
| Hematologic malignancy | 10 | 9 | 1 |
| HIV | 2 | 2 | 0 |
| Wegener's granulomatosis | 1 | 1 | 0 |
| *Normal skin* | 18 | 14 | 4 |
| Immunocompetent | 8 | 6 | 2 |
| Immunosuppressed | 10 | 8 | 2 |
| Solid organ transplant | 8 | 6 | 2 |
| Hematologic malignancy | 2 | 2 | 0 |

Abbreviations: KA, keratoacanthoma; SCC, squamous cell carcinoma.

### Low viral load of β-HPV in normal skin and cuSCC

In all, 85 DNA samples were assayed for the presence of β-globin DNA by PCR and all demonstrated sufficient quantity and integrity for β-HPV typing. Of 67 cuSCC tumors, 20 (30%) were positive for HPV DNA by PCR, 18 of which were confirmed by sequencing. Of the 18 normal skin samples, 5 (28%) were also HPV positive by PCR and sequence confirmation (Table 2). In all, 11 HPV types and 14 incompletely sequenced fragment types were detected, with no single type predominating. Multivariate regression modeling demonstrated no difference in β-HPV carriage between tumor and normal tissue when controlling for age, sex, and immunosuppression as well as clustering for multiple samples from the same patient ($P=0.693$). Immunosuppression and older age were significantly associated with β-HPV carriage ($P=0.018$ for both). PCR and sequencing confirmed the presence of α-HPV16 in the primary cervical and periungual SCCs and α-HPV18 in HeLa cells. In addition, the periungual SCC was found to contain β-HPV8 and FA51.2 DNA.

Viral loads were determined for up to 3 HPV types in 23 of the 24 HPV PCR-positive and sequence-confirmed samples. Replicate assays were performed for HPV18 in the HeLa cervical cancer cell line, HPV16 in the cervical SCC, and HPV16 and HPV8 in the periungual SCC (Table 2 and Figure 1a). With the exception of four samples (one normal skin and three tumors), all viral loads were below 1 HPV copy per cell. In contrast, the cervical SCC contained 2.4 HPV16 copies per

## Table 2. HPV viral loads

| Patient | Sex | Age | Immunosuppression | Sample code | Histology | Site | DNA HPV types | Viral load | | | | Transcriptome sequenced |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | HPV type | HPV copies | Input cells | HPV copies/cell | |
| PC-021 | M | 89 | No | STA01-046 | KA | Scalp | FA7, FA127, FA37 | FA7 | ND | 3,921 | ND | |
| | | | | | | | | FA37 | ND | 3,921 | ND | |
| | | | | STA01-051 | KA | Scalp | HPV21, FA127, FA37 | HPV21 | ND | 13,952 | ND | |
| | | | | | | | | FA37 | 100 | 13,952 | 0.007 | |
| | | | | STA01-053 | Normal skin | Leg | HPV80, FA7 | HPV80 | ND | 6,434 | ND | |
| | | | | | | | | FA7 | 28 | 6,434 | 0.004 | |
| PC-031 | M | 95 | No | STA01-130 | SCC | Scalp | — | | | | | |
| | | | | STA01-131 | Normal skin | Scalp | HPV5, HPV8 | HPV5 | 220,872 | 2,313 | 95.501 | |
| | | | | | | | | HPV8 | ND | 2,313 | ND | |
| PC-046 | M | 84 | No | STA01-068 | SCC | Hand | HPV75, FA108 | HPV75 | 149 | 14,607 | 0.010 | |
| | | | | STA01-079 | Normal skin | Leg | HPV75 | HPV75 | 152 | 4,651 | 0.033 | |
| PC-003 | M | 68 | Heart transplant | STA01-034 | SCC | Forehead | HPV5, HPV49, FAIMVS14 | HPV5 | 696 | 2,611 | 0.267 | |
| | | | | STA01-122 | Normal skin | Cheek | — | | | | | Yes |
| PC-015 | F | 53 | Heart–lung transplant | STA01-018 | SCC | Arm | HPV9 | HPV9 | 984 | 28,332 | 0.028 | Yes |
| | | | | STA01-076 | Normal skin | Arm | HPV9 | HPV9 | 57 | 33,490 | 0.001 | Yes |
| | | | | STA01-077 | KA | Arm | HPV9 | HPV9 | ND | 7,264 | ND | Yes |
| | | | | STA01-094 | Normal skin | Forehead | — | | | | | Yes |
| | | | | STA01-095 | SCC | Forehead | — | | | | | Yes |
| PC-007 | M | 57 | Lung transplant | STA01-074 | SCC | Scalp | HPV20, HPV21 | HPV20 | 1,256 | 51,462 | 0.024 | Yes |
| | | | | | | | | HPV21 | 10,407 | 51,462 | 0.202 | |
| | | | | STA01-078 | Normal skin | Scalp | HPV21, HPV20 | HPV20 | ND | 47,880 | ND | |
| | | | | | | | | HPV21 | ND | 47,880 | ND | |
| PC-041 | M | 60 | Lung transplant | STA01-029 | SCC | Cheek | HPV17 | HPV17 | 26,132 | 45,136 | 0.579 | |
| | | | | STA01-030 | SCC | Cheek | Sequence not obtained | | | | | Yes |
| PC-058 | M | 51 | Lung transplant | STA01-010 | SCC | Lip | HPV5, HPV80, FA14 | HPV5 | ND | 6,693 | ND | Yes |
| | | | | | | | | HPV80 | ND | 6,693 | ND | |
| | | | | | | | | FA14 | 492 | 6,693 | 0.074 | |
| | | | | STA01-045 | SCC | Cheek | FANIMVS11.4, FAIMVS11.3 | NA | | | | |
| | | | | STA01-059 | SCC | Cheek | — | | | | | |
| | | | | STA01-132 | SCC | Scalp | — | | | | | |
| | | | | STA01-133 | Normal skin | Cheek | — | | | | | |

Table 2 continued on the following page

## Table 2. Continued

| Patient | Sex | Age | Immunosuppression | Sample code | Histology | Site | DNA HPV types | Viral load | | | | Transcriptome sequenced |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | HPV type | HPV copies | Input cells | HPV copies/cell | |
| PC-030 | F | 73 | Renal transplant | STA01-031 | KA | Leg | FA14, FA16.3 | FA14 | 28,006 | 5,098 | 5.494 | |
| | | | | | | | | FA16.3 | 1,068 | 5,098 | 0.209 | |
| | | | | STA01-032 | KA | Leg | FA16.3, FA75 | FA16.3 | ND | 17,901 | ND | |
| | | | | STA01-090 | KA | Leg | FA14, HPV96, FA140.2, FA16.3 | FA14 | 17,216 | 4,712 | 3.654 | |
| | | | | | | | | FA16.3 | ND | 4,712 | ND | |
| | | | | | | | | FA140.2 | ND | 4,712 | ND | |
| | | | | STA01-091 | KA | Leg | FA16.3, FA140.2, FA14 | FA14 | 1,808 | 3,372 | 0.536 | Yes |
| | | | | | | | | FA16.3 | 44,140 | 3,372 | 13.089 | |
| | | | | | | | | FA140.2 | 110 | 3,372 | 0.033 | |
| PC-054 | M | 55 | Renal/pancreas transplant | STA01-065 | SCC | Arm | — | | | | | |
| | | | | STA01-066 | SCC | Arm | HPV8 | HPV8 | 4,462 | 25,593 | 0.174 | Yes |
| PC-011 | M | 64 | CLL/SLL | STA01-064 | SCC | Forehead | HPV19, HPV49, FA123 | HPV19 | 1,132 | 26,360 | 0.043 | Yes |
| | | | | | | | | FA123 | 4,717 | 26,360 | 0.179 | |
| | | | | STA01-071 | SCC | Forehead | HPV19, FA33 | HPV19 | 38 | 3,883 | 0.010 | |
| PC-053 | M | 53 | HIV | STA01-099 | SCC | Forehead | HPV5, FA16 | HPV5 | 1,126 | 32,837 | 0.034 | Yes |
| | | | | | | | | | | | | |
| *Control viral loads* | | | | | | | | | | | | |
| PC-043 | M | 53 | No | STA01-035 | SCC | Hand | HPV16 | HPV16 | 570,172 | 12,156 | 46.904 | Yes |
| | | | | | | | HPV8, FA51.2 | HPV8 | 3,685 | 12,157 | 0.303 | Yes |
| HeLa | F | | | HeLa | Cell line | Cervix | HPV18 | HPV18 | 8,496,079 | 1,350,389 | 6.291 | Yes |
| ILS-19363 | F | 35 | No | ILS-19363-1 | SCC | Cervix | HPV16 | HPV16 | 3,403,823 | 1,403,307 | 2.426 | Yes |

Abbreviations: ALL, acute lymphoblastic leukemia; CLL/SLL, chronic lymphocytic leukemia/small lymphocytic lymphoma; F, female; GVHD, graft-vs.-host disease; HPV, human papillomavirus; KA, keratoacanthoma; M, male; MDS, myelodysplastic syndrome; NA, no assay could be designed; ND, not detected; SCC, squamous cell carcinoma; SCT, stem cell transplant.
Patients with HPV PCR-positive samples are shown.

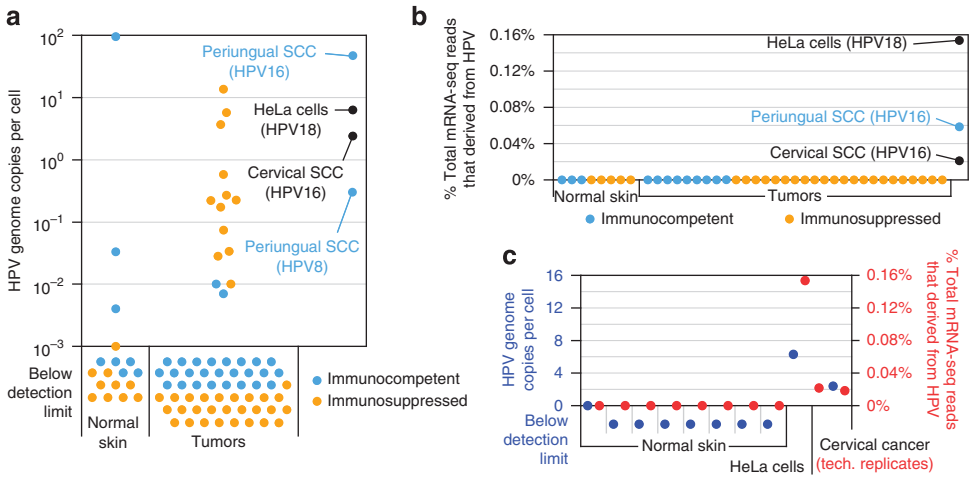


**Figure 1. Comparison of human papillomavirus (HPV) DNA viral load and abundance of HPV-derived transcripts for established HPV-driven cancers versus normal skin and cutaneous squamous cell carcinomas (cuSCCs).** (**a**) HPV DNA viral loads determined by type-specific quantitative PCR. For each sample with multiple-type infection, the sum of the type-specific viral loads is shown (for details, see Table 2.) (**b**) Abundance of HPV-derived transcripts determined by mRNA-transcriptome sequencing (mRNA-seq). Reads were filtered to remove host-derived and low-complexity sequence before viral mapping (see Materials and Methods), and HPV counts are each presented as a percentage of their total data set. The most frequently matched HPV type for the HeLa, periungual SCC, and cervical SCC samples is indicated. Cervical cancer is presented as the union of two technical replicate data sets. (**c**) Congruence between HPV genomic load (blue; presented as in **a**) and abundance of HPV-derived transcripts (red; presented as in **b**) among those control samples from which both types of data were collected.

cell and the HeLa cell line contained 6.3 copies per cell, consistent with viral integration. The periungual SCC contained 46.9 α-HPV16 copies per cell but only 0.3 β-HPV8 copies per cell.

Random-effects interval regression modeling demonstrated no difference in β-HPV viral load between tumor and normal tissue when controlling for age, sex, and immunosuppression as well as clustering for multiple samples from the same patient ($P=0.310$). Immunosuppression and age were significantly associated with higher viral load ($P=0.016$ for immunosuppression; $P=0.0004$ for age).

### No HPV transcripts observed in cuSCC

The potential oncogenicity of HPV viruses in our samples was assessed in terms of viral gene expression by mRNA-transcriptome sequencing (mRNA-seq). A total of 31 cuSCC tumors (10 KA type) were assayed by high-throughput mRNA-seq, including 10 β-HPV DNA-positive samples with viral loads ranging from undetectable to 13.7 copies per cell. Parallel libraries were prepared from eight patient-paired normal skin samples as well as the aforementioned periungual SCC tumor, cervical SCC, and HeLa cell samples. Paired-end read counts per library ranged from 1.5 million to 10.6 million reads (corresponding to sequence from 740,000 to 5.3 million complementary DNA amplicons), with a median count of 3.5 million reads (1.8 million amplicons; Supplementary Table S2 online). After removing sequences with low sequence complexity or with high-quality matches to the human genome or transcriptome, reads were queried against a database of all fully sequenced viral genomes from RefSeq (see Materials and Methods). Abundant HPV-matching reads were detected in the HeLa cell-derived and cervical SCC-derived data sets (0.15 and 0.02% of the total reads, respectively Figure 1b). In both, the HPV subtype identified was an α-HPV subtype known to drive tumorigenesis in that sample type (HPV18 and HPV16, respectively) (IARC, 2007; Bouvard *et al.*, 2009). Read frequencies from two technical replicate libraries prepared from cervical SCCs were nearly identical (0.022 and 0.018%; Figure 1c, Supplementary Table S1 online).

Abundant HPV16 reads were also detected in the transcriptome of the periungual SCCs (0.06% of total reads; Figure 1b), with no reads mapping preferentially to HPV8 or any other HPV subtype. No potentially HPV-derived reads were detected in 30 of the cuSCC tumors and 7 of the paired normal skin samples (Figure 1b). One normal skin sample that was HPV negative by genomic PCR contained two HPV-matching reads (0.00008% of the total reads), and one skin SCC sample contained five HPV-derived reads (0.0001% of the total read set; this sample had no DNA for genomic PCR).

No libraries contained a higher frequency of total viral-matching reads than the lowest HPV-derived read count among the HPV-positive control samples (0.018%), and only five had viral read frequencies within an order of magnitude of that value (Supplementary File S2 online). These five samples contained a mixture of phage sequence (likely deriving from bacteria on the skin), human klassevirus 1 (a candidate etiological agent for diarrhea that was isolated from human stool in our lab and is therefore a likely cross-lab contaminant; Greninger *et al.*, 2009), and Moloney murine leukemia virus. A reverse transcriptase deriving from the latter was used for library construction (see Materials and Methods), making it a likely reagent contaminant. Although Moloney murine leukemia virus and related viruses do have oncogenic capabilities (Cuypers *et al.*, 1984), the sample in which it was detected at the highest frequency was a normal

skin control, not a tumor (sample ID STA01-122, Dataset C, barcode CGT, Supplementary Table S2 and Supplementary File S2 online).

In these 39 cases of normal skin and cuSCCs, the frequencies of HPV reads were orders of magnitude less than those observed for any of the *bona fide* HPV-driven tumors. In control samples for which both mRNA-seq and viral genome load data were obtained, the results obtained by these two metrics agreed with one another and with previous descriptions of the role of HPV as an oncogenic virus; genomes and transcripts were both abundant in HeLa cells and cervical SCCs and both absent in healthy skin (Figure 1c). The similarity of HPV genome and transcriptome quantitation between HeLa cells and cervical SCCs versus periungual SCCs supported the role for HPV in periungual SCC tumorigenesis, whereas the quantitative similarity between the cuSCC versus healthy skin samples implied no role for HPV transcription in the maintenance of those tumors.

## DISCUSSION

Previous studies have proposed β-HPV as a potential causative agent in cuSCC, citing the presence of viral DNA in tumor tissue, but these have not definitively proved an epidemiologic association or evaluated any particular mechanism of transformation. We used transcriptome sequencing to test the hypothesis that HPV is required for the maintenance of cuSCCs through expression of viral oncoproteins. Transcriptome sequencing revealed a complete absence of HPV mRNA in these tumors, similar to paired normal skin. This stood in stark contrast to the abundant HPV messages detected in cervical SCC and its derivative HeLa cell line. Our results, in fact, contradict the hypothesis that expression of viral oncogenes is required for maintenance of cuSCCs.

Periungual SCC represents a special site on the cutaneous epithelium. These tumors are associated with high-risk α-HPV (Moy *et al.*, 1989; Alam *et al.*, 2003; Kreuter *et al.*, 2009), which has been reported as episomal, and in a single case, integrated (Sanchez-Lanier *et al.*, 1994; Theunis *et al.*, 1999). In our control periungual SCC, β-HPV8 and FA51.2 DNA were detected along with α-HPV16. This tumor contained 46.9 α-HPV16 copies per cell but only 0.3 β-HPV8 copies per cell. HPV16 mRNA reads represented 0.06% of the transcriptome, but no β-HPV mRNA reads were detected. Taken together, this control specimen supports our impression of α-HPV as the driver and β-HPV as a mere passenger in periungual SCC.

As in previous studies (Shamanin *et al.*, 1994, 1996; Berkhout *et al.*, 1995, 2000; Boxman *et al.*, 1997; Antonsson *et al.*, 2000; Harwood *et al.*, 2000; Forslund *et al.*, 2003a, b, 2007; Hazard *et al.*, 2007; de Koning *et al.*, 2007, 2009; Asgari *et al.*, 2008), β-HPV DNA was detectable by nested PCR in 30% of SCCs, but was also found in a comparable proportion (28%) of normal skin samples. Moreover, we found extremely low viral loads in tumors that were positive for the viral genome. In all but three tumor samples, the viral load was $<1$ copy per cell. Importantly, the three contradictory samples all came from a single renal transplant recipient with multiple KAs of the lower leg and may reflect a unique feature of that case. Use of PCR and sequencing allowed identification of a broad range of HPV types, although the detection of the multiplicity of infection may have been limited by the number of clones sequenced. Alternate methods for β-HPV detection such as line blots and microarrays allow simultaneous detection of types but are limited in the types detected. Although DNA from other HPV types may have been present in these samples, this does not alter the conclusion of this study. The low copy number of β-HPV DNA, combined with the absence of virally derived oncogenic messages, strongly suggested that β-HPV transcription is not required for tumor maintenance.

Our data were consistent with previous evidence that β-HPV merely colonizes the skin. Immunosuppression and older age were associated with a higher prevalence and viral load of β-HPV, consistent with previous studies (Boxman *et al.*, 2001; Struijk *et al.*, 2003; de Koning *et al.*, 2009). These phenomena likely reflected the role of the immune system and age-related immune senescence in controlling epidermal colonization with HPV rather than explaining the increased incidence of SCC in organ-transplant recipients and older patients. The prevalence of HPV DNA in tape-stripped biopsies is far lower than that on the surface, further supporting a passenger role (Forslund *et al.*, 2004). Support for β-HPV as a passenger also comes from a study of tumors from patients with xeroderma pigmentosum, in which prevalence of viral carriage increases with age and is very low in tumors from children (Luron *et al.*, 2007). A reversed relationship in which SCC somehow results in the presence or increase of β-HPV DNA or antibodies is possible, although further investigation would be required to substantiate this.

Insertional mutagenesis is another mechanism of viral oncogenesis; this mechanism has been described for oncoretroviruses but not for DNA viruses. High-risk α-HPV types can integrate into the host genome but require continual expression of the viral E6 and E7 proteins for their oncogenic activities (Dyson *et al.*, 1989; Scheffner *et al.*, 1990; Werness *et al.*, 1990). The recently described Merkel cell polyomavirus, another small DNA oncovirus, also integrates into the host genome (Feng *et al.*, 2008), but continued expression of the Merkel cell polyomavirus truncated large T antigen is similarly required for carcinogenesis (Houben *et al.*, 2010). In contrast, there are no reports of β-HPV integration into the genome of cuSCC. The low viral loads of β-HPV in cuSCC reported here further indicate that even if β-HPV had integrated, at most only a small proportion of the genomes within any tumor could contain integrated β-HPV, casting doubt upon integration as a carcinogenic mechanism.

It has also been suggested that β-HPV might play a role in induction but not maintenance of cuSCC (based on higher viral load of HPV in precancerous actinic keratoses vs. primary SCC, metastatic tumor, or perilesional skin; Weissenborn *et al.*, 2005). This may occur by interfering with cellular DNA repair or apoptosis following UV irradiation, creating a pool of genomically unstable cells at risk of oncogenic transformation. Our study was not designed to address this hypothesis. But it should be noted that such a hypothesis

would represent a substantial departure from the role played by α-HPV in mucosal SCC and from the carcinogenic mechanisms known to be employed by other families of DNA tumor viruses in general. Therefore, the most straightforward interpretation of our data is that the sporadic and low-level presence of β-HPV genomic DNA in these tumors, unaccompanied by evidence of active viral gene expression, most likely represents colonization rather than an etiologic association.

## MATERIALS AND METHODS

### Sample collection

All subjects provided informed consent according to the procedures approved by the University of California, San Francisco Committee on Human Research and adherent to the Declaration of Helsinki Principles. Tumor tissue was collected from patients during the course of biopsy or excision. All specimens were held for further processing until final pathology confirmed a diagnosis of cuSCC. Normal tissue was collected when surgical discard was available from postoperative reconstruction. All tissue was snap-frozen and stored on liquid nitrogen until nucleic acid extraction.

A primary cervical cancer specimen containing HPV16 was obtained from a commercial tissue bank (ILSBio, LLC, Chestertown, MD).

### Nucleic acid isolation

Tissue samples were minced, divided, and placed in parallel extraction pathways. DNA was extracted using the QIAamp DNA Mini Kit with RNase A (Qiagen, Valencia, CA) as per the manufacturer's protocol. RNA extractions were carried out using the RNeasy Lipid Tissue Mini Kit with on-column RNase-free DNase I (Qiagen) as per the manufacturer's protocol.

### HPV DNA detection using PCR

A PCR assay for β-globin DNA was performed on each sample to control for DNA integrity and for the presence of adequate quantity of DNA. Next, 5 μl of each DNA sample (30–800 ng) was tested with primers PCO4 and GH20 as described (Bauer *et al.*, 1991). β-HPV PCR was carried out using the nested primer sets FAP59-FAP64 and FAP6085F-FAP6319R (Forslund *et al.*, 2003a, b). For the first round of PCR, 5 μl of each DNA sample was amplified using 2 μM of the FAP59 and FAP64 primers in a 50 μl reaction volume including 1 × Taq Buffer, 2 mM $MgCl_2$, 0.25 mM dNTPs, and 1 U Taq polymerase. The reaction was carried out under the following PCR conditions: 94 °C for 2 minutes followed by 25 cycles of 94 °C for 30 seconds, 50 °C for 1 minute, and 72 °C for 1 minute, with a final extension time of 7 minutes at 72 °C. A 5 μl aliquot of the product was removed for a second round of amplification using the nested FAP 6085F and FAP 6319R primer pair under the same cycling conditions. α-HPV PCR was carried out using the nested primer sets MY09-MY11 and GP5-GP6 as described (Manos *et al.*, 1989; Snijders *et al.*, 2005).

The products were visualized by agarose gel electrophoresis and bands of expected size were isolated using the PureLink Quick Gel Extraction Kit (Invitrogen, Carlsbad, CA) and cloned using the TOPO TA cloning System (Invitrogen). A minimum of 12 colonies were sequenced on the ABI 3130 × l Genetic Analyzer (Applied Biosystems, Carlsbad, CA) in order to detect potential multiple infections.

### Quantification of HPV viral load by real-time quantitative PCR

Quantitative real-time PCR was performed using the Universal Probe Library system (Roche Applied Science, Indianapolis, IN). Primer and probe assay combinations were individually designed for each HPV type and for human β-2-microglobulin using the online universal probe library Design Center software (Roche). For samples with multiple infections, we designed discriminatory assays to measure type-specific viral loads of as many individual types as possible (Supplementary Table S1 online). To generate standard curves, assay-specific PCR amplicons were separated on 4% agarose gel and purified using the PureLink Quick Gel Extraction Kit (Invitrogen). Internal standards were generated using 10-fold dilutions of the gel-purified products ranging from 1,000,000 to 10 input copies.

DNA samples were assayed in 20 μl reactions with a final concentration of 1 × LightCycler 480 Probes Master mix, 400 nM of each primer, and 200 nM of the universal probe library probe. Using the LightCycler 480 (Roche), samples were heated to 95 °C for 10 minutes followed by 45 cycles of 95 °C for 10 seconds, 60 °C for 30 seconds, and 70 °C for 1 second. Data were analyzed with the LightCycler 480 software. DNA samples and standard dilution series were run in duplicate, and total input copy numbers were calculated using the mean crossing point (Cp) values for each sample. Input cellular equivalents were calculated based on two copies of β-2-microglobulin copies per cell, and HPV viral loads were calculated as viral copies per cell.

### Statistical analysis

HPV prevalence was analyzed by logistic regression including age, sex, immunosuppression, tissue sample type, and accounting for clustering of multiple samples within patients. HPV viral loads were analyzed with random-effects interval regression where the HPV copy number was left-censored at the lower limit of detection by quantitative PCR. For both prevalence and viral load, univariate regression was performed before multivariate modeling. Statistical analysis was performed using Stata 11 (StataCorp LP, College Station, TX).

### mRNA-seq library preparation and analysis

Poly(A)+ RNA was isolated from 3 μg of total RNA using the Oligotex Mini kit (Qiagen) according to the manufacturer's instructions. The resulting poly(A) RNA was then amplified using the MessageAmp II amplified RNA Amplification Kit (Ambion, Austin, TX) using an *in vitro* transcription time of 14 hours at 37 °C to generate amplified RNA. Then, 200 ng of amplified RNA was used to generate libraries for transcriptome sequencing using an adaptation of the protocol previously described by Yozwiak *et al.* (2010). In order to multiplex up to 16 samples within 1 sequencing library, amplified RNA samples were randomly primed and reverse transcribed using a primer containing a 14-bp sequence common to the 3′-end of both Illumina (San Diego, CA) adapters, a random monomer followed by a unique 3-bp barcode, and a random hexamer (pr1A_barcode). Second-strand synthesis was primed using pr1A followed by PCR amplification using the 18-bp Illumina/barcode sequence without the hexamer (pr1B_barcode) for 25 cycles. PCR products were purified using DNA Clean and Concentrator columns (Zymo Research, Orange, CA). Next, 200 ng of each individually barcoded sample was mixed together to generate a library of up to 16 samples, with each sample marked by a unique 3-bp barcode. Library purification, size selection, and

amplification proceeded as previously described (Yozwiak *et al*., 2010). Three multiplexed transcriptome libraries were analyzed on three separate paired-end sequencing runs using the Genome Analyzer II (Illumina) and designated Datasets A, B, and C. Barcodes and corresponding samples are listed in Supplementary Table S2 online. Each run generated pairs of 65 nt reads. These data have been submitted to the NCBI Sequence Read Archive under accession number SRA029929.

Read pairs from each library were sorted by 3 nt barcode (nucleotides 2–4 of each read), requiring that at least one of the two reads from each pair contained a perfect match to an input barcode and that the other contained at most one mismatch. This yielded the ''total'' read counts shown in Supplementary Table S2 online. For analysis, we removed from each read the nucleotide preceding the barcode, the barcode itself, the 6 nucleotides deriving from the random hexamer used for priming, and the last nucleotide of the read, yielding 54 nt reads.

Background model (BGM) DNA sequence data sets included the human genome (UCSC build hg18; BGMhg) (Lander *et al*., 2001; Fujita *et al*., 2010), the human mRNA transcriptome (representative H-Invitational transcripts, 43,159 records; BGMht) (Imanishi *et al*., 2004), a collection of sequenced human VDJ recombination products (*Homo sapiens* entries from IMGT release 201028-6 67,611 records; BGMvdj) (Lefranc, 2001), the Illumina paired-end adapter sequences ligated to one another (BGMad), and an *in vitro*-transcribed *Xenopus* EF1α message that contaminated Dataset A and was reconstructed from that data (Supplementary File S1 online; BGMef1a). Matches to BGMhg, BGMht, and BGMef1a of >80% sequence identity across the entire read length were sought using BLAT (-minIdentity = 80 -noTrimA) (Kent, 2002), and matches to those data sets plus BGMvdj and BGMad were sought using Blastn (default settings) (Altschul *et al*., 1990). Matching sequences and their paired ends were filtered from the query pool, leaving the ''host-filtered'' read counts shown in Supplementary Table S2 online. Barcode AGG from Dataset A was excluded from further analysis because of the majority of reads mapping to the BGMef1a contaminating sequence (not shown). Low-complexity sequences were defined as those generating <30 new additions to the string table during the Lempel–Ziv–Welch compression (Welch, 1984) and were removed, leaving the ''complexity-filtered'' read counts shown in Supplementary Table S2 online.

Matches to the remaining reads were sought in a database of all complete viral genome sequences in Genbank (3,525 records; 72 million nucleotides; downloaded on 18 January 2010; GIs listed in Supplementary File S3 online) (Benson *et al*., 2009) using tBlastx (–e 1e–3). Read counts were allocated to the viral genome record with the highest alignment bitscore. In the case of a tie, the read count was initially distributed evenly to all records with equal bitscore matches, and then re-assigned to whichever record(s) had the greatest total read count for the given data set. ''HPV-matching'' read counts for each sample are shown in Supplementary Table S2 online.

## CONFLICT OF INTEREST

The authors state no conflict of interest.

## ACKNOWLEDGMENTS

This work was supported by the NIH/NCRR/OD UCSF-CTSI grant number KL2 RR024130, a Canary Foundation/American Cancer Society Postdoctoral Fellowship for the Early Detection of Cancer, a Dermatology Foundation Career Development Award, and an American Society for Dermatologic Surgery Cutting Edge Research grant to STA. We thank Clement Chu for assistance with deep sequencing, Peter Skewes-Cox for assistance with sequence database collection, Amy J Markowitz for critical reading of this article, and Charles McCulloch for advice on statistical analysis. These mRNA-seq data have been submitted to the NCBI Sequence Read Archive under accession number SRA029929.

## SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at http://www.nature.com/jid

## REFERENCES

Alam M, Caldwell JB, Eliezri YD (2003) Human papillomavirus-associated digital squamous cell carcinoma: literature review and report of 21 new cases. *J Am Acad Dermatol* 48:385–93

Altschul SF, Gish W, Miller W *et al.* (1990) Basic local alignment search tool. *J Mol Biol* 215:403–10

Antonsson A, Forslund O, Ekberg H *et al.* (2000) The ubiquity and impressive genomic diversity of human skin papillomaviruses suggest a commensalic nature of these viruses. *J Virol* 74:11636–41

Asgari MM, Kiviat NB, Critchlow CW *et al.* (2008) Detection of human papillomavirus DNA in cutaneous squamous cell carcinoma among immunocompetent individuals. *J Invest Dermatol* 128:1409–17

Bauer HM, Ting Y, Greer CE *et al.* (1991) Genital human papillomavirus infection in female university students as determined by a PCR-based method. *JAMA* 265:472–7

Benson DA, Karsch-Mizrachi I, Lipman DJ *et al.* (2009) GenBank. *Nucleic Acids Res* 37:D26–31

Berkhout RJ, Bouwes Bavinck JN, ter Schegget J (2000) Persistence of human papillomavirus DNA in benign and (pre)malignant skin lesions from renal transplant recipients. *J Clin Microbiol* 38:2087–96

Berkhout RJ, Tieben LM, Smits HL *et al.* (1995) Nested PCR approach for detection and typing of epidermodysplasia verruciformis-associated human papillomavirus types in cutaneous cancers from renal transplant recipients. *J Clin Microbiol* 33:690–5

Bosch FX, Lorincz A, Munoz N *et al.* (2002) The causal relation between human papillomavirus and cervical cancer. *J Clin Pathol* 55:244–65

Bouvard V, Baan R, Straif K *et al.* (2009) A review of human carcinogens–Part B: biological agents. *Lancet Oncol* 10:321–2

Bouwes Bavinck JN, Neale RE, Abeni D *et al.* (2010) Multicenter study of the association between betapapillomavirus infection and cutaneous squamous cell carcinoma. *Cancer Res* 70:9777–86

Boxman IL, Berkhout RJ, Mulder LH *et al.* (1997) Detection of human papillomavirus DNA in plucked hairs from renal transplant recipients and healthy volunteers. *J Invest Dermatol* 108:712–5

Boxman IL, Russell A, Mulder LH *et al.* (2001) Association between epidermodysplasia verruciformis-associated human papillomavirus DNA in plucked eyebrow hair and solar keratoses. *J Invest Dermatol* 117:1108–12

Cuypers HT, Selten G, Quint W *et al.* (1984) Murine leukemia virus-induced T-cell lymphomagenesis: integration of proviruses in a distinct chromosomal region. *Cell* 37:141–50

Dang C, Koehler A, Forschner T *et al.* (2006) E6/E7 expression of human papillomavirus types in cutaneous squamous cell dysplasia and carcinoma in immunosuppressed organ transplant recipients. *Br J Dermatol* 155:129–36

de Koning MN, Struijk L, Bavinck JN *et al.* (2007) Betapapillomaviruses frequently persist in the skin of healthy individuals. *J Gen Virol* 88:1489–95

de Koning MN, Weissenborn SJ, Abeni D *et al.* (2009) Prevalence and associated factors of betapapillomavirus infections in individuals without cutaneous squamous cell carcinoma. *J Gen Virol* 90:1611–21

Dubina M, Goldenberg G (2009) Viral-associated nonmelanoma skin cancers: a review. *Am J Dermatopathol* 31:561–73

Dyson N, Howley PM, Munger K *et al.* (1989) The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product. *Science* 243:934–7

Feng H, Shuda M, Chang Y *et al.* (2008) Clonal integration of a polyomavirus in human Merkel cell carcinoma. *Science* 319:1096–100

Forslund O, Iftner T, Andersson K *et al.* (2007) Cutaneous human papillomaviruses found in sun-exposed skin: Beta-papillomavirus species 2 predominates in squamous cell carcinoma. *J Infect Dis* 196:876–83

Forslund O, Lindelof B, Hradil E *et al.* (2004) High prevalence of cutaneous human papillomavirus DNA on the top of skin tumors but not in ''Stripped'' biopsies from the same tumors. *J Invest Dermatol* 123:388–94

Forslund O, Ly H, Higgins G (2003a) Improved detection of cutaneous human papillomavirus DNA by single tube nested 'hanging droplet' PCR. *J Virol Methods* 110:129–36

Forslund O, Ly H, Reid C *et al.* (2003b) A broad spectrum of human papillomavirus types is present in the skin of Australian patients with non-melanoma skin cancers and solar keratosis. *Br J Dermatol* 149:64–73

Fujita PA, Rhead B, Zweig AS *et al.* (2010) The UCSC Genome Browser database: update 2011. *Nucleic Acids Res* 39(Database issue):D876–82

Greninger AL, Runckel C, Chiu CY *et al.* (2009) The complete genome of klassevirus – a novel picornavirus in pediatric stool. *Virol J* 6:82

Hartevelt MM, Bavinck JN, Kootte AM *et al.* (1990) Incidence of skin cancer after renal transplantation in The Netherlands. *Transplantation* 49:506–9

Harwood CA, McGregor JM, Proby CM *et al.* (1999) Human papillomavirus and the development of non-melanoma skin cancer. *J Clin Pathol* 52:249–53

Harwood CA, Surentheran T, McGregor JM *et al.* (2000) Human papillomavirus infection and non-melanoma skin cancer in immunosuppressed and immunocompetent individuals. *J Med Virol* 61:289–97

Hazard K, Karlsson A, Andersson K *et al.* (2007) Cutaneous human papillomaviruses persist on healthy skin. *J Invest Dermatol* 127:116–9

Houben R, Shuda M, Weinkam R *et al.* (2010) Merkel cell polyomavirus infected Merkel cell carcinoma cells require expression of viral T antigens. *J Virol* 84:7064–72

IARC (2007) Human papillomaviruses. *IARC Monogr Eval Carcinog Risks Hum* 90:1–636

Imanishi T, Itoh T, Suzuki Y *et al.* (2004) Integrative annotation of 21,037 human genes validated by full-length cDNA clones. *PLoS Biol* 2:e162

Jensen P, Hansen S, Moller B *et al.* (1999) Skin cancer in kidney and heart transplant recipients and different long-term immunosuppressive therapy regimens. *J Am Acad Dermatol* 40:177–86

Karagas MR, Nelson HH, Sehr P *et al.* (2006) Human papillomavirus infection and incidence of squamous cell and basal cell carcinomas of the skin. *J Natl Cancer Inst* 98:389–95

Kent WJ (2002) BLAT–the BLAST-like alignment tool. *Genome Res* 12:656–64

Kreuter A, Gambichler T, Pfister H *et al.* (2009) Diversity of human papillomavirus types in periungual squamous cell carcinoma. *Br J Dermatol* 161:1262–9

Lander ES, Linton LM, Birren B *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature* 409:860–921

LeBoit PE (2002) Can we understand keratoacanthoma? *Am J Dermatopathol* 24:166–8

Lefranc MP (2001) IMGT, the international ImMunoGeneTics database. *Nucleic Acids Res* 29:207–9

Lindelof B, Sigurgeirsson B, Gabel H *et al.* (2000) Incidence of skin cancer in 5356 patients following organ transplantation. *Br J Dermatol* 143:513–9

Luron L, Avril MF, Sarasin A *et al.* (2007) Prevalence of human papillomavirus in skin tumors from repair deficient xeroderma pigmentosum patients. *Cancer Lett* 250:213–9

Manos M, Ting Y, Wright D *et al.* (1989) Use of polymerase chain reaction amplification for the detection of genital human papillomaviruses. In: Furth M, Greaves M (eds) *Cancer Cells 7: Molecular Diagnostics of Human Cancer*. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 209–14

Moy RL, Eliezri YD, Nuovo GJ *et al.* (1989) Human papillomavirus type 16 DNA in periungual squamous cell carcinomas. *JAMA* 261:2669–73

Parkin DM (2006) The global health burden of infection-associated cancers in the year 2002. *Int J Cancer* 118:3030–44

Purdie KJ, Surentheran T, Sterling JC *et al.* (2005) Human papillomavirus gene expression in cutaneous squamous cell carcinomas from immunosuppressed and immunocompetent individuals. *J Invest Dermatol* 125:98–107

Sanchez-Lanier M, Triplett C, Campion M (1994) Possible role for human papillomavirus 16 in squamous cell carcinoma of the finger. *J Med Virol* 44:369–78

Scheffner M, Werness BA, Huibregtse JM *et al.* (1990) The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell* 63:1129–36

Shamanin V, Glover M, Rausch C *et al.* (1994) Specific types of human papillomavirus found in benign proliferations and carcinomas of the skin in immunosuppressed patients. *Cancer Res* 54:4610–3

Shamanin V, zur Hausen H, Lavergne D *et al.* (1996) Human papillomavirus infections in nonmelanoma skin cancers from renal transplant recipients and nonimmunosuppressed patients. *J Natl Cancer Inst* 88:802–11

Snijders PJ, van den Brule AJ, Jacobs MV *et al.* (2005) HPV DNA detection and typing in cervical scrapes. *Methods Mol Med* 119:101–14

Struijk L, Bouwes Bavinck JN, Wanningen P *et al.* (2003) Presence of human papillomavirus DNA in plucked eyebrow hairs is associated with a history of cutaneous squamous cell carcinoma. *J Invest Dermatol* 121:1531–5

Surentheran T, Harwood CA, Spink PJ *et al.* (1998) Detection and typing of human papillomaviruses in mucosal and cutaneous biopsies from immunosuppressed and immunocompetent patients and patients with epidermodysplasia verruciformis: a unified diagnostic approach. *J Clin Pathol* 51:606–10

Theunis A, Andre J, Noel JC (1999) Evaluation of the role of genital human papillomavirus in the pathogenesis of ungual squamous cell carcinoma. *Dermatology* 198:206–8

Vajdic CM, McDonald SP, McCredie MR *et al.* (2006) Cancer incidence before and after kidney transplantation. *JAMA* 296:2823–31

Weissenborn SJ, Nindl I, Purdie K *et al.* (2005) Human papillomavirus-DNA loads in actinic keratoses exceed those in non-melanoma skin cancers. *J Invest Dermatol* 125:93–7

Welch TA (1984) A technique for high-performance data-compression. *Computer* 17:8–19

Werness BA, Levine AJ, Howley PM (1990) Association of human papillomavirus types 16 and 18 E6 proteins with p53. *Science* 248:76–9

Yozwiak NL, Skewes-Cox P, Gordon A *et al.* (2010) Human Enterovirus 109: a novel inter-species recombinant enterovirus discovered in acute pediatric respiratory illness in Nicaragua. *J Virol* 84:9047–58

Zur Hausen H (1996) Papillomavirus infections–a major cause of human cancers. *Biochim Biophys Acta* 1288:F55–78

Zur Hausen H (2009) The search for infectious causes of human cancers: where and why. *Virology* 392:1–10