Serum, Cell-Free HPV-Human DNA Junction Detection and HPV Typing for Predicting and Monitoring Cervical Cancer Recurrence

Anne Van Arsdale[1,2], Olga Mescheryakova[1], Sonia Gallego[1], Elaine C. Maggi[1], Bryan Harmon[3], Dennis Y.S. Kuo[2], Koenraad Van Doorslaer[4], Mark H. Einstein[2], Brian J. Haas[5], Cristina Montagna[1,6]*, and Jack Lenz[1]*

[1]Department of Genetics, Albert Einstein College of Medicine, Bronx, NY, 10461, USA,
[2]Department of Obstetrics Gynecology and Women's Health, Albert Einstein College of Medicine, Bronx, NY, 10461, USA,
[3]Department of Pathology, Albert Einstein College of Medicine, Bronx, NY 10461, USA,
[4]School of Animal and Comparative Biomedical Sciences, College of Agriculture and Life Sciences BIO5 Institute, University of Arizona, Tucson, AZ, 85721, USA,
[5]Broad Institute, Cambridge, MA 02142, USA,
[6]Rutgers Cancer Institute, 195 Little Albany St., New Brunswick, NJ, 08901, USA

*communicating authors

Keywords – cervical cancer, human papilloma virus, HPV, cell-free DNA, cfDNA, DNA integration, cancer recurrence.

Abstract
Almost all cervical cancers are caused by human papillomaviruses (HPVs). In most cases, HPV DNA is integrated into the human genome. We found that tumor-specific, HPV-human DNA junctions are detectable in serum cell-free DNA of cervical cancer patients at initial treatment. Retrospective analysis revealed these junctions were more frequently detectable in women in whom the cancer later recurred. We also found that cervical cancers caused by HPV types outside of phylogenetic clade α9 had a higher recurrence frequency than those caused by α9 types in both our study and The Cancer Genome Atlas cervical cancer database, despite the higher prevalence of α9 types including HPV16 in cervical cancer. Thus, HPV-human DNA junction detection in serum cell-free DNA and HPV type determination in tumor tissue may help predict recurrence risk. Screening serum cell-free DNA for junctions may also offer an unambiguous, non-invasive means to monitor absence of recurrence following treatment.

Main:
The vast majority of cervical cancers are caused by human papillomaviruses (HPVs), and they remain a significant cause of cancer deaths. Despite the availability of HPV vaccines and effective procedures for detecting and treating pre-invasive cervical lesions, they rank as the fourth most common cancer in women globally[1]. Early-stage invasive cervical cancers (FIGO IA1-IB2) are typically treated with surgical excision as a standard of care. However, for patients with advanced stages (FIGO IB3-IVB), surgery is not beneficial, and first line therapy entails cisplatin-radiosensitization combined with ionizing radiation plus brachytherapy, with recurrence risk increasing with stage[2]. In the United States, an estimated 13,820 new cases of cervical carcinoma are projected for 2024, resulting in 4,360 deaths. The 5-year overall survival (OS) rate across all stages is 66%, with rates dropping to 58% for loco-regional spread and 17% for distant metastasis[3]. An enhanced ability to stratify women during initial care by identifying those at highest risk of treatment failure may improve patient care.

HPV genomes replicate as circular, extra-chromosomal, double-stranded DNAs of about 8 kbp. However, in the vast majority of cervical cancers, HPV DNA integrates into human DNA, either within a chromosome or as an HPV-human heterocatemeric, extra-chromosomal DNA[4,5,6]. Each junction between HPV and human DNAs is an unambiguous, patient tumor-specific identifier of clonally expanded tumor cells derived from the cell that initially underwent the integration event.

Therefore, we investigated whether HPV-human DNA junctions could be detected circulating in serum cell-free DNA (cfDNA), and whether they could be used to assess residual disease following initial cervical cancer treatment. HPV-human DNA junctions were retrospectively assessed using archived samples from 16 cervical cancer patients (12 squamous carcinomas, 4 adenocarcinomas) with available frozen tissue and at least one serum sample obtained during initial care. (**Figure 1A, 1B & Supplementary Tables 1 & 2**). Eight patients had locally advanced cervical cancer (FIGO IB3-IIIC1r), while eight had surgically resectable, early-stage cancers (FIGO IB1-IB2). HPV-human DNA junctions were identified using DNA isolated from tissue biopsies or surgical excisions. Hybridization capture combined with Illumina sequencing (HC+SEQ) was performed for full genomes of 225 HPV types followed by computational analysis using CTAT-

Virus Integration Finder[7] to identify HPV-human DNA junctions (**Supplementary Figure 1A, 1B & 1C**). This approach also enabled the unambiguous identification of HPV types in each tumor including lower prevalence types HPV70 and HPV73 in one patient each, plus individual, tumor-specific differences compared to viral type reference sequences at single nucleotide resolution (**Figure 1A, 1B & Supplementary Table 2**). Junction-PCRs using tumor-specific primers, one targeting HPV and the other targeting human genome sequences, confirmed the HPV-human junctions in the cognate tumor tissue DNAs (**Supplementary Table 3**). Five of the sixteen patients had tumor recurrences, all occurring at distant metastatic sites outside standard-of-care radiation fields or excised organs. Biopsy samples were available from all five recurrences, and junction-PCR analysis confirmed identical junctions to those found in the cognate primary tumor (**Supplementary Figure 2)**, thus confirming clonal origin from the primary tumors.

The usefulness of junction detection in serum cfDNAs was evaluated using samples obtained during initial care from fifteen of the patients and at six months post-treatment from seven patients (**Figure 1A**). Serum cfDNA concentrations were significantly higher in women diagnosed with locally advanced cancers stage compared to those with early-stage cancers (Wilcoxon rank sum test, $p<0.05$) (**Figure 1C**), but no significant difference was observed between those with or without a recurrence (**Figure 1D**). HPV-human DNA junctions in cfDNAs were assessed by HC+SEQ directly on cfDNA and junction PCR methods, which yielded concordant results except for Patient 2 (Methods). Junctions were detected in cfDNA from seven patients, including all cases of recurrences except no baseline sample was available for Patient 4 (**Figure 1A**).

The significance of junction detection for predicting recurrence at baseline and six months was assessed by Kaplan-Meier analyses of recurrence-free survival, with patients left-censored at the date of last-follow-up **(Figures 1E & 1F**). Stratified log-rank tests adjusted for stage at diagnosis showed that detection at 6 months was statistically significant and prognostic for recurrence ($p=0.035$). Detection at initial treatment was of borderline statistical significance ($p=0.06$). Due to the small sample size, accurate determination of effect size of detection at either time point with appropriate 95% confidence intervals could not be accomplished using either univariable or multivariable Cox regression modeling. Nonetheless, junction detection in serum cfDNA at 0 and 6 months showed promise as a predictor of tumor recurrence.

HPV DNA detection alone in cfDNA of cervical cancer patients has demonstrated value for helping predict tumor recurrence risk.[8] However, while circulating tumor DNA (ctDNA) is the predominant component of cfDNA in cancer patients, HPV DNA from non-tumor tissues might also be present in cfDNA. Therefore, HPV type-specific E7 DNA detection in cfDNA was also evaluated (**Figures 1B**). Among available samples, 14/15 (93.3%) baseline serum cfDNAs and 6/7 post-treatment sera, including all five from patients with recurrences, tested positive for HPV type-specific E7 DNA by HC+SEQ and E7 gene PCR. Kaplan-Meier analysis at baseline and six months post-treatment using stratified log-rank tests for HPV E7 positivity, adjusted for stage, did not significantly associate with recurrence at either time point ($p=0.52$ and $p=0.31$, respectively) (**Figure 1G and 1H**). Due to the small sample size, accurate effect size

determination with 95% confidence intervals using univariable or multivariable Cox regression modeling was not feasible.

To test if cfDNA E7 matched the HPV variant integrated in the cognate tumor, we compared cfDNA-tumor HC+SEQ pairs at single nucleotide resolution. Integrated HPV DNAs always include the URR-E6-E7 segment, they usually comprise only a subgenomic segment of the viral genome (**Supplementary Figure 1**), and they generally contain some nucleotide differences versus a reference sequence. All pairs matched, indicating that the HPV DNA in serum cfDNA derived from the respective tumors, except Patient 16's cfDNA contained both HPV16 and HPV18, while the excised tumor had only HPV 16 DNA (confirmed by E7 PCR, **Supplementary Figure 3**). Neither HPV genome in Patient 16 cfDNA had junctions with human DNA. Moreover, the HPV16 DNA in Patient 16 cfDNA had different SNPs versus the reference sequence, and it included sequences spanning a deleted 1kb segment in the tumor-integrated HPV16 (**Supplementary Figure 4**). Thus, the two HPVs detected in Patient 16 cfDNA during initial care likely did not arise from the tumor-integrated HPV16 DNA in this patient, showing that the source of HPV DNA in cfDNA can be of uncertain tissue origin in some patients. HPV-human DNA junctions in serum cfDNA should thus provide a more rigorous, tumor-specific means for detecting the presence of tumor cells in some patients than detecting just HPV DNA.

Unexpectedly, all five recurrences in this cohort involved HPV types from phylogenetic clade α7 (HPV18, HPV39, HPV45, HPV70) or α5 (HPV69), with none involving clade α9 types (HPV16, HPV35) (**Figure 1A**). Patients with cervical cancers containing clade α7 HPV types have inferior survival versus those with clade α9 [9,10]. Therefore, HPV clade types were also assessed as an independent prognostic factor for recurrence using Kaplan-Meier analysis of clinical information and tumor tissue sequences from the larger cervical cancer dataset in The Cancer Genome Atlas (TCGA) (**Figure 2**). HPV types were determined using CTAT-Virus Integration Finder[7] analysis of TCGA sequencing data, revealing 214 tumors (69.5%) with clade α9 types, 83 (26.9%) with other clade types (α5, α6, α7, or α11) , and 11 (3.6%) being HPV-negative**.** Clinicopathologic variables known to be *a priori* prognostic factors were extracted from cBioPortal[11 12], including age, histology, and stage at diagnosis, with N=297 patients having complete data. Comparisons of HPV clade α9 with non-clade-α9 (α5/α6/α7/α11) showed no significant baseline differences in these variables (**Supplementary Table 4**). However, patients with non-α9 clades exhibited inferior recurrence-free survival (RFS) compared to those with HPV clade-α9 tumors based on a log-rank test stratified for histology and stage (p=0.03). A multivariable Cox regression model with age (years), histology (squamous/adenocarcinoma) and stage (I, II, III, IV) was performed to account for all *a priori* variables, and women with HPV clades other than α9 showed inferior RFS in comparison to HPV clade-α9 tumors (Hazard Ratio (HR)=2.48, 95% CI: 1.34-4.49, p=0.004) after adjusting for all variables (**Figure 2**). The effect of HPV clade on RFS was second only to stage IV disease (HR: 4.40, 95% CI: 1.89-10.18, p=0.001) in this adjusted analysis (**Supplementary Table 5**). Identification of non-α9 clade HPV types (at least α5, α6, α7, and α11) as a recurrence risk factor in the TCGA/cBioPortal data[11 12] and the small retrospective study presented here establish that HPV clade type specifically affects cervical cancer recurrence following treatment as well as survival, and may reflect the same underlying

molecular genetic properties of these HPV types. These effects contrast with the high prevalence of α9 HPV16 (~50%) in cervical cancer and the correlation between HPV16 and progression of preinvasive, cervical intraepithelial neoplasia[13 14]. They also highlight a potential role for precise HPV typing in cervical cancer management using HPV-type-multiplexed, hybridization capture plus sequencing methods like HC+SEQ to facilitate clade assignment. These approaches are technically comparable to clinically-utilized assays like exon capture, and inclusion of all known HPV types also enables broad diagnostic specificity plus tumorigenicity assessment of rare HPV types.

The results presented here show that HPV-human junction detection in cfDNA and HPV type determination in tumor tissue at the time of initial cervical cancer treatment might be useful to help predict the likelihood of cervical cancer recurrence. Detection of HPV DNA alone in cfDNA is useful for predicting relapse.[8 15] However, the detection of HPV's different than the one integrated in the cfDNA of Patient 16 here (**Supplementary Figure 4**) illustrates the potential for occasional misleading results, and HPV DNA has been reported in blood of healthy donors[16], which is also consistent with the possibility that it may occasionally arise from non-tumor tissue sources. Our results suggest that HPV-human DNA junctions could be tested in other types of HPV-induced cancers. HPV-human DNA junctions in cfDNA offer the advantage of unequivocal origin from the cancer cell clones in patients. Long-term prospective studies are required to confirm the findings here, and the potential for improved recurrence risk prediction emphasizes the need to develop improved upfront therapies for women at risk for cervical cancer relapse. Nonetheless, at a minimum, junction detection in cfDNA potentially offers a non-invasive means for rigorously monitoring disease recurrence in patients following cervical cancer treatment. In summary, HPV-human DNA junction detection in cfDNA and HPV type determination at the time of initial treatment offer promise for predicting and monitoring cervical cancer recurrence.

Figure 1

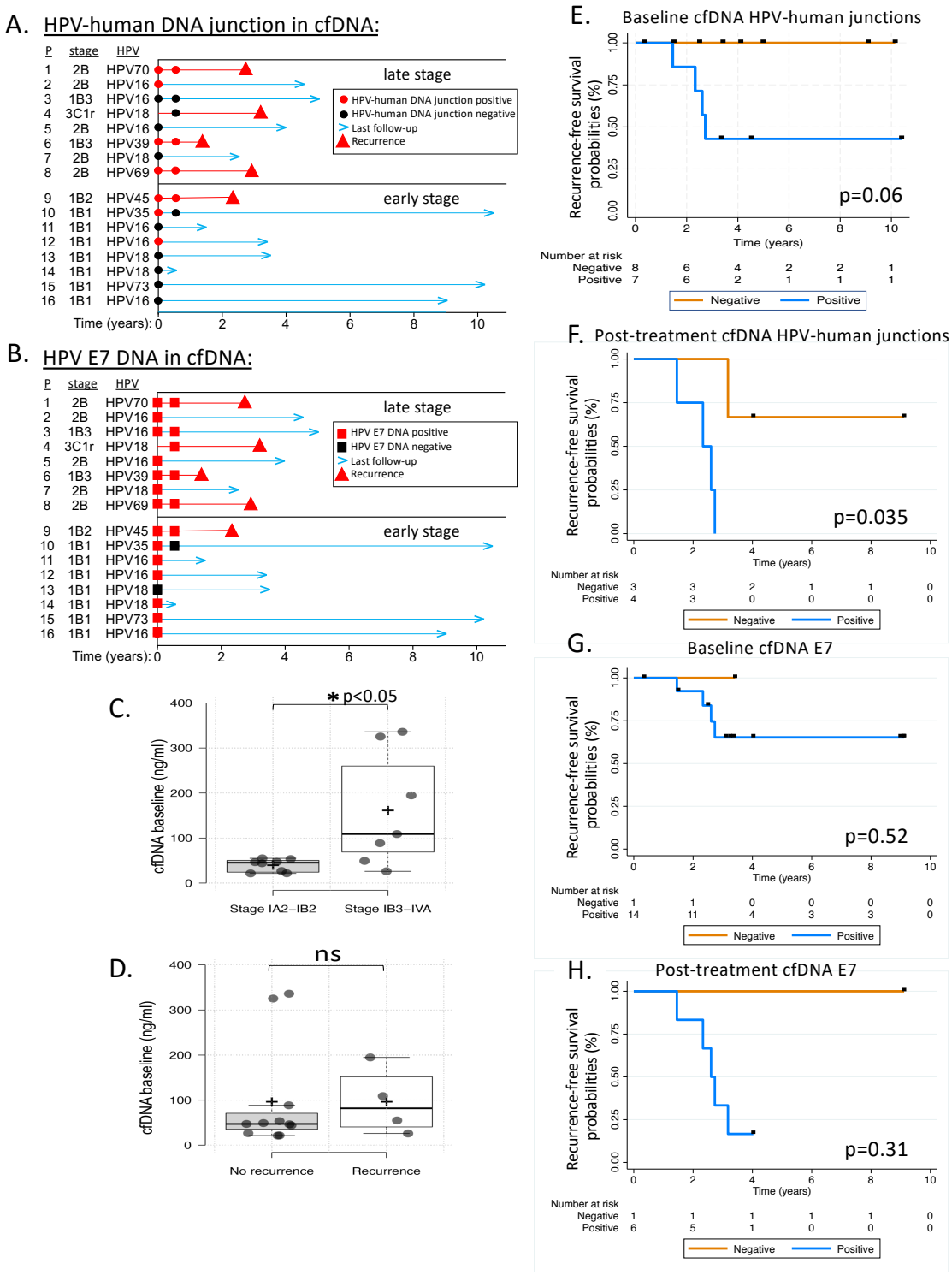
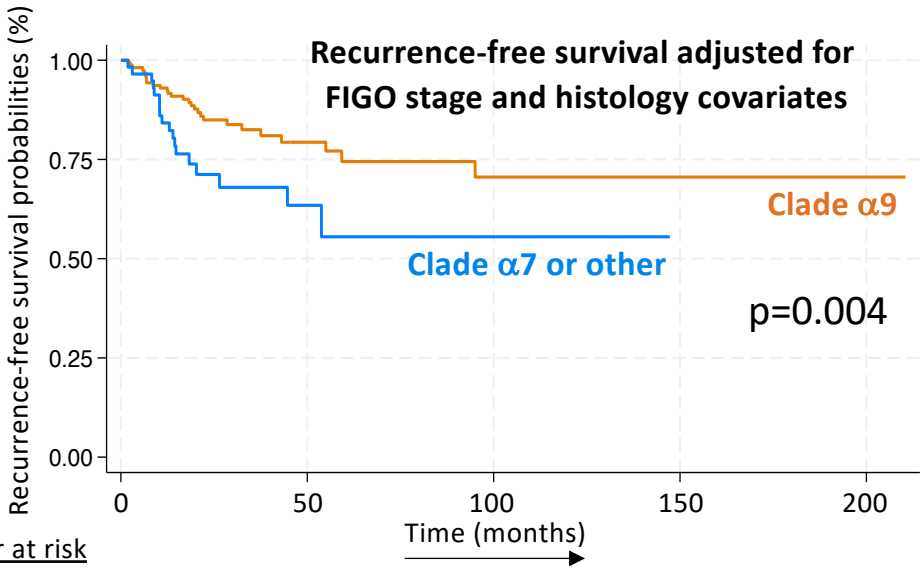


Figure 1. Patient cohort survival and analyses. Patients 3, 7, 11, and 14 had adenocarcinomas; the others had squamous cancers. HPV-human DNA junction (**A**) and HPV E7 DNA detection (**B**) cfDNA concentrations in early versus late-stage cancers (**C**) and recurrences versus non-recurrences (**D**). Kaplan-Meier analyses of recurrences by junction (**E-F**) or E7 detection in cfDNA (**G-H**). Squares show left-censoring at last patient visit.

Figure 2



Number at risk

Clade α9: 181	41	16	7	2
Clade α7 or other: 66	11	1	0	0

Figure 2. Kaplan-Meier recurrence-free survival plot of 297 TCGA cervical cancer patients

References

1. Sung, H.*, et al. CA Cancer J Clin* **71**, 209-249 (2021).
2. Girda, E.*, et al. Gynecol Oncol* **179**, 115-122 (2023).
3. Siegel, R.L., Giaquinto, A.N. & Jemal, A. *CA Cancer J Clin* **74**, 12-49 (2024).
4. Rossi, N.M.*, et al. Cancer Res* **83**, 1768-1781 (2023).
5. Van Arsdale, A.*, et al. NPJ Genom Med* **9**, 35 (2024).
6. Akagi, K.*, et al. Cancer Discov* **13**, 910-927 (2023).
7. https://github.com/TrinityCTAT/CTAT-VirusIntegrationFinder.
8. Han, K.*, et al. J Clin Oncol* **42**, 431-440 (2024).
9. Gagliardi, A.*, et al. Nat Genet* **52**, 800-810 (2020).
10. Rader, J.S.*, et al. Int J Cancer* **144**, 2206-2214 (2019).
11. Cerami, E.*, et al. Cancer Discov* **2**, 401-404 (2012).
12. Gao, J.*, et al. Sci Signal* **6**, pl1 (2013).
13. Sykes, P.H.*, et al. Am J Obstet Gynecol* **226**, 222 e221-222 e213 (2022).
14. Tainio, K.*, et al. BMJ* **360**, k499 (2018).
15. Leung, E.*, et al. Clin Cancer Res* **27**, 5857-5868 (2021).
16. Vergara, N.*, et al. J Med Virol* **91**, 2009-2015 (2019).