



PHS PUBLIC ACCESS

Author manuscript

Appl Immunohistochem Mol Morphol. Author manuscript; available in PMC 2017 August 01.

Published in final edited form as:

Appl Immunohistochem Mol Morphol. 2016 August ; 24(7): 490–495. doi:10.1097/PAI.0000000000000215.

Identification of Human Papillomavirus Infection in Cancer Tissue by Targeted Next Generation Sequencing

Nathan D. Montgomery, MD, PhD¹, Joel S. Parker, PhD², David A. Eberhard, MD, PhD^{1,2}, Nirali M. Patel, MD^{1,2}, Karen E. Weck, MD^{1,2}, Norman E. Sharpless, MD^{2,3}, Zhiyuan Hu, PhD², D. Neil Hayes, MD^{2,3,*}, and Margaret L. Gulley, MD^{1,2,*}

¹Department of Pathology and Laboratory Medicine, The University of North Carolina at Chapel Hill

²Lineberger Comprehensive Cancer Center, The University of North Carolina at Chapel Hill

³Division of Medical Oncology, Department of Medicine; The University of North Carolina at Chapel Hill

Abstract

Human papillomaviruses (HPV) are oncogenic DNA viruses implicated in squamous cell carcinomas of several anatomic sites, as well as endocervical adenocarcinomas. Identification of HPV is an actionable finding in some carcinomas, potentially influencing tumor classification, prognosis, and management. We incorporated capture probes for oncogenic HPV strains 16 and 18 into a broader next-generation sequencing (NGS) panel designed to identify actionable mutations in solid malignancies. A total of 21 head and neck, genitourinary and gynecological squamous cell carcinomas and endocervical adenocarcinomas were sequenced as part of the UNCSeq project. Using p16 immunohistochemical results as the gold standard, we set a cutoff for proportion of aligned HPV reads that maximized performance of our NGS assay (92% sensitive, 100% specific for HPV). These results suggest that sequencing of oncogenic pathogens can be incorporated into targeted NGS panels, extending the clinical utility of genomic assays.

Keywords

Human Papillomavirus; Next Generation Sequencing; cervical carcinoma; oropharyngeal carcinoma

INTRODUCTION

Human papillomavirus (HPV) is a double-stranded DNA virus implicated in pathogenesis of squamous cell carcinoma at a variety of anatomic sites, as well as endocervical adenocarcinomas and benign warts of skin and mucous membranes.^{1, 2} Worldwide each year, more than 600,000 newly diagnosed cancers, representing nearly 5% of all new

Corresponding Author: Margaret L. Gulley, MD, Department of Pathology & Laboratory Medicine and Lineberger Comprehensive Cancer Center, 913 Brinkhous-Bullitt Building, University of North Carolina, Chapel Hill, NC 27599-7525, Telephone: (919) 843-4595, margaret_gulley@med.unc.edu.

*These authors contributed equally to this work.

malignancies, are attributed to HPV infection.² While cervical cancers and other gynecologic tract malignancies are responsible for much of the global burden of HPV-related cancer, in recent years, HPV-related carcinomas of other mucosal sites have been increasingly recognized as a major public health burden. For instance, in the United States, oropharyngeal squamous cell carcinomas are now more commonly associated with HPV infection than with smoking, reflecting a major epidemiologic transition in this disease.³ HPV infection is also implicated in a subset of squamous cell carcinomas arising in perineal skin, anus, and genitourinary sites.^{2, 4}

More than 100 HPV strains are defined.⁵ Sexually-transmitted HPV strains are classified as high-risk (strains 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 73, and 82) or low-risk based on their observed oncogenicity in cervical tissue.⁶ These high risk strains may be found in dysplasia or in invasive malignancy at other sites, although certain viral strains are more frequently implicated at certain anatomic sites and in selected geographic populations. Identification of HPV in carcinoma is not only pathophysiologically relevant but also may influence management decisions.⁷ For instance, HPV-associated oropharyngeal squamous cell carcinomas are exquisitely radiation sensitive.⁸ As such, patients may be candidates for lower dose radiation therapy protocols with reduced treatment-related morbidity.⁹ For this reason, and because results are also prognostic, practice guidelines now recommend that HPV status be determined on all newly diagnosed oropharyngeal carcinomas.¹⁰

In clinical laboratories, the methods generally utilized to detect HPV in biopsy and surgical resection specimens include p16 immunohistochemistry (IHC), HPV in situ hybridization (ISH), and molecular methods such as PCR.^{10, 11} CDKN2A (p16) is a cyclin-dependent kinase inhibitor which is overexpressed in HPV-associated malignancies due to inhibition of RB1 by virally-encoded E7 protein.^{11, 12} Immunohistochemical detection of host p16 expression provides a robust, albeit indirect, method for localizing HPV infection in malignant cells¹³⁻¹⁵. However, HPV-independent p16 upregulation can occur, meaning that p16 overexpression is not entirely specific for HPV infection.^{16, 17} Moreover, p16 IHC cannot distinguish viral strains. While direct detection of high-risk HPV nucleic acid by ISH could theoretically overcome some of these limitations, available HPV ISH assays are hampered by inferior test performance and high cost relative to p16 IHC.^{17, 18} Meanwhile alternative molecular assays such as PCR are limited by inability to localize infection to malignant cells.¹¹

With the development of massively parallel genome sequencing technology, it is possible to comprehensively characterize the mutational spectrum of tumors.¹⁹ Numerous clinical laboratories now offer next-generation sequencing (NGS) panels to interrogate somatically mutated genes implicated in tumorigenesis or predicting response to therapy. We explored incorporating viral capture probes into such a panel to provide a new method for detecting and subtyping oncogenic viruses like HPV.

UNCseq is a research protocol at the University of North Carolina at Chapel Hill (UNC) which aims to identify actionable somatic genomic variants in cancer tissue. Capture probes used in this study include conserved sequences of several viruses, including HPV strains 16 and 18. As of January 2014, 523 tumors had been sequenced, including a total of 21 head

and neck, genitourinary and gynecological squamous cell carcinomas and endocervical adenocarcinomas. Herein, we report test performance characteristics of our NGS panel for HPV detection and demonstrate that incorporating viral capture probes into such panels offers a reliable method for identifying and subtyping HPV in tumor tissues.

MATERIALS AND METHODS

Patients and samples

The UNCseq project is UNC's major research study to establish clinical-grade oncology services using NGS technology. The protocol involves sequencing exons of a custom list of 247 human genes and 10 pathogen genome segments in fixed or frozen cancer tissue and matched germline DNA from consenting local patients. All studies were done with approval of our Institutional Review Board.

As of January 2014, 523 tumors had been sequenced. In each case, non-malignant tissue from the same subject was also sequenced in order to identify somatic changes. Gynecologic, head and neck, and genitourinary squamous cell carcinomas and endocervical adenocarcinomas (N = 21) are the focus of the current study, as these are the malignancies in which HPV has been previously implicated for clinical decision-making.

DNA Isolation, Library Preparation, and Sequencing

A pathologist examined H&E stained slides from each case to confirm the diagnosis and to estimate malignant cell proportion. DNA was extracted from fresh frozen tissue samples (16 cases) or from FFPE tissue sections (5 cases) using either the Gentra Puregene Tissue Kit (QIAGEN, Valencia CA) or the Maxwell 16 DNA purification platform (Promega Corp, Madison, WI), and then fragmented by sonication. Resultant DNA fragments were then subjected to repair and end-polishing (blunt-end or A-overhang), before ligation of custom, single-end adapters. Agilent SureSelect Target Enrichment System was used for baited capture of pertinent human and viral gene segments (Agilent Technologies, Santa Clara, CA). Capture probes were designed to target exons from 247 host genes, as well as ten cancer-related pathogens (HPV strain 16, HPV strain 18, Epstein-Barr virus, human herpesvirus 8, BK virus, JC virus, human T-lymphotropic virus, Merkel cell polyomavirus, and *Helicobacter pylori*). Relevant to the current work, capture probes incorporate viral genome sequence for HPV strains 16 and 18 (GenBank IDs NC_001526 and NC_001357, respectively).

During DNA isolation and library preparation, DNA concentration was measured by fluorometry, DNA quality was assessed using the Agilent 2100 Bioanalyzer high sensitivity DNA assay, and DNA size was determined by Experion automated electrophoresis system (BioRad, Hercules CA) or by Agilent TapeStation. Sequencing was then performed using a HiSeq2000 sequencer (Illumina Inc, San Diego CA).

NGS Data Analysis

Raw sequence reads were analyzed using the CASAVA v.1.8 package (Illumina) to generate FASTQ files, a file format which combines nucleotide sequence data and quality indicators.

Using the Burrows-Wheeler Aligner software (BWA v0.6.2) developed to map sequence data, reads were then aligned to the reference human genome (build 37, GRCh37) and to selected genomic sequences for the ten pathogens listed above, all of which were combined into a single concatenated reference sequence. To decrease problems of misalignment caused by insertions/deletions, local realignment and base quality score recalibration were performed using the Genome Analysis Toolkit (GATK v2.6) and the GATK resource bundle (v2.5)[17]. Filtering was performed by imposing a minimum Phred quality score of read mapping (MapQ). Reads with low mapping quality (MapQ < 5) were removed.

HPV Analysis

The number of aligned HPV (strain 16 or 18) reads and total aligned sequence reads were recorded for each case. The average number of independent reads aligned at particular genomic regions (sequencing depth) varied from sample to sample. In order to adjust for variations in sequencing depth, the number of aligned HPV reads was then calculated per one million total sequence reads aligned to either the human or pathogen genomes. This calculation provides a value reflecting the proportion of all sequence reads that derive from HPV for each tumor.

p16 IHC

For cases in which p16 IHC was performed as part of clinical care, slides and relevant controls were reviewed. For all other cases, a block of tumor was recut and immunohistochemical stains were performed in the UNC Translational Pathology Laboratory using a mouse monoclonal anti-p16 antibody (clone E6H4, Ventana, Tucson, AZ) and appropriate positive and negative controls. Signal localization to at least 70% of malignant cells was interpreted as a positive result.

HPV ISH

ISH for detection of HPV strains 16 and 18 was performed using the Pathogene(R) HPV Type 16/18 probe (Enzo Life Sciences, Farmingdale, NY) on a Leica Bond-III instrument according to the manufacturer's instructions with appropriate positive and negative controls. Signal localization to malignant cells was interpreted as a positive result.

Statistics

After the number of HPV reads per one million total aligned reads was calculated for all cases, mean values for p16-positive and p16-negative samples were compared by an unpaired t-test using GraphPad Prism (GraphPad Software, Inc., La Jolla, CA).

RESULTS

Characterization of p16 status by immunohistochemistry

A total of 21 head and neck, genitourinary or gynecological squamous cell carcinomas or endocervical adenocarcinomas were examined. By IHC, 13 were p16-positive, and 8 were p16-negative. Differences in anatomic site of origin were noted between p16-positive and p16-negative cases, consistent with the epidemiology of these diseases. Specifically, 11 of

12 gynecologic malignancies, 2 of 6 head and neck malignancies, and 0 of 3 genitourinary malignancies were p16-positive (Table 1). For these 21 neoplasms, the percent of malignant cells input into the sequencing assay ranged from 20 to 90% (mean 51%).

Determination of HPV status by NGS

Given variation in sequencing depth between cases (range = $8.5 - 43 \times 10^6$ total aligned sequence reads), the number of HPV reads was normalized by calculating the number of aligned HPV reads per 10^6 total aligned sequence reads. Observed normalized values ranged from 0 to 39,968 HPV reads per million aligned reads. When cases were stratified by IHC results, the mean number of corrected HPV reads was significantly different between p16-positive cases and p16-negative cases (mean 13,231 versus 0.1 HPV reads per million reads, $p < 0.0001$) (Figure 1). In fact, six of eight p16-negative cases had no aligned HPV 16 or 18 reads, and the highest corrected value of HPV-aligned reads among p16-negative cases was just 0.7 HPV reads/ 10^6 total aligned reads. By comparison, the lowest corrected value of HPV-aligned reads among p16-positive cases was 3.6, and twelve of thirteen p16-positive cases had corrected values greater than 200 HPV reads/ 10^6 total aligned reads. Using a conservative cutoff of 100 HPV reads per 1 million total aligned reads (see Discussion), the sensitivity and specificity of the NGS assay are 92% and 100%, respectively.

HPV status by *in situ* hybridization

As part of routine clinical care, high risk HPV ISH was performed on five tumors. Although p16 and NGS results were concordant in all cases, one p16-positive/NGS-positive case was negative by ISH (Table 2), suggesting that the ISH result was false-negative for HPV detection (and is further discussed below).

Strain specificity explored by NGS

Distinct HPV16 and HPV18 capture probes are utilized in our platform, and we explored the proportion of reads assigned to each reference HPV genome (strain types 16 and 18). In ten of thirteen p16-positive cases, all HPV reads aligned to a single strain (Table 3). In the remaining three cases, at least 90-fold differences in HPV16 *versus* HPV18 aligned reads were noted. Although we have not excluded coinfection, cross-reactive hybridization, or misalignment as causes for ambiguous strain typing, these findings suggest the potential for strain assignment based on NGS data analysis.

HPV in unexpected tumor types

As part of the broader UNCSeq project, we examined tumors considered unlikely to be HPV-associated based on anatomic site and histopathologic spectrum, and some of these tumors had low levels of aligned HPV reads (Figure 2). In this broader group of tumors, the neoplasm with the highest HPV level (9.9 HPV reads/ 10^6 total aligned reads) was an ovarian carcinoma that was demonstrated to be positive for HPV by ISH targeting strains 16/18.

DISCUSSION

Here, we report accurate detection of HPV as part of a broader next generation sequencing panel intended to detect actionable somatic variants. This work confirms previous work by

Conway et al²⁰, who used similar methods to demonstrate the utility of NGS methods for HPV identification in head and neck cancer. We now extend the utility of NGS to detect HPV in cervical and genitourinary cancers, and we demonstrate that sequencing-based methods can perform robustly when incorporated into a targeted somatic mutation test panel.

In our study, p16 immunohistochemistry was used as a gold standard to identify HPV - related tumors. This method is widely employed for surrogate HPV identification in anatomic and clinical pathology laboratories, and numerous prior studies have demonstrated that p16 IHC is generally equivalent to or superior to competing microscope-based methods, such as high risk HPV ISH.^{17, 18} High risk HPV ISH was performed on five cases in this study as part of routine clinical care. Interestingly, although p16 IHC and sequencing results were concordant in all cases, one p16-positive/NGS-positive case was negative by ISH (Table 2). Although a small sample size, these results suggest that NGS-based HPV detection may perform favorably in comparison to other available methods.

Given that there was a five-fold difference in HPV reads between the lowest p16-positive case and the highest p16-negative case, a cut-off of 2 HPV reads/10⁶ would result in excellent NGS assay performance with 100% accuracy (100% sensitivity and 100% specificity) in this small study. However, as part of the broader UNCSeq project, low HPV read counts (<10 per million aligned reads) were occasionally found in tumors considered unlikely to be HPV-associated by virtue of their anatomic site and histopathologic appearance (Figure 2), perhaps reflecting occasional true positive infection of tumor or background normal tissue, cross-hybridization of capture probes, or misalignment of sequences having homology to the HPV reference sequences. Applying a more conservative cut-off of 100 HPV reads per 10⁶ total aligned reads, the assay exhibits 92% sensitivity and 100% specificity. If these cutoffs are confirmed in larger studies, it may be reasonable to omit confirmatory testing in tumors having high (>100) or low (<1) HPV read counts per million aligned reads. On the other hand, for cases with intermediate numbers of HPV reads (e.g. 1 to 100 reads per 10⁶ reads), p16 IHC may be appropriate as a confirmatory test to distinguish infected from uninfected tumors. A suggested algorithm for tiered testing is shown in Figure 3.

There are several advantages to an NGS assay for viral detection. First, although p16 IHC is much less expensive than NGS, there is minimal additional cost required to incorporate viral capture probes and accompanying bioinformatic analysis into a somatic mutation panel that is otherwise used as part of routine clinical care. NGS findings may limit the number of IHC, ISH, or other ancillary test methods that are required. Second, sequencing-based methods might simultaneously identify and subtype relevant oncogenic viruses. Although we have not confirmed the accuracy of strain assignments made with our assay, the tendency for sequence reads to align to either HPV16 or HPV18 is promising in this regard. Third, targeted therapy for HPV infection and its associated biochemical pathways is being explored,²¹ making it all the more important to accurately identify HPV-infected tumors. Finally, capture probes and reference sequences for additional strains can be incorporated into NGS-based assays. The adaptability of NGS technology has several potential advantages for patient care and for research. For instance, shifts in prevalence of HPV and in

viral strain types in response to HPV vaccination^{22, 23} could reflect evolving epidemiology of pertinent malignancies.

Before HPV detection by next-generation sequencing is incorporated into routine clinical care, additional study is warranted. Future studies should evaluate lesions infected by non-oncogenic HPV strains, such as the low risk strains 6 and 11, to evaluate possible cross reactivity/misalignment with their high risk counterparts. Although we did not explore the minimum tumor content required for accurate HPV result interpretation, it is likely that this variable impacts quantitative results for HPV, and future studies should explore the minimum tumor content required for accurate HPV detection. Similarly, HPV DNA copy number can vary from 1 to over 100 viral genomes per tumor cell, and it may be useful to clarify how such variation influences assay performance. Other relevant variables include measures of DNA quality and quantity, and the depth of coverage at pertinent loci. Such quality indicators and limits of acceptability need to be established.

This study extends the potential utility of NGS sequencing panels in routine pathologic evaluation of tumor specimens. Our success in identifying HPV in routinely collected tumor samples has implications for the pathogenesis, prognosis, and treatment of several types of malignancy known to be associated with oncogenic viruses. Emerging genomic studies suggest that pathogens may be implicated in more tumors than was historically appreciated.²⁴ Extension to other oncogenic viruses, such as EBV, HHV8, HTLV1, HBV, and Merkel cell polyomavirus, should be explored. Evaluation of premalignant lesions is also feasible.²⁵

ACKNOWLEDGEMENTS

We wish to acknowledge the UNC Translational Pathology Laboratory, the Pre-clinical Genomic Pathology Laboratory, and the Rapid Adoption Molecular Laboratory for supporting this project.

Conflicts of Interest and Sources of Funding: Dr. Gulley has collaborated with Illumina on separate studies. This research was supported by the University of North Carolina Department of Pathology and Laboratory Medicine, the University Cancer Research Fund, an award for Clinical Translational Science (NIH UL1 TR001111), and the Network Group Integrated Translational Science Center (NCI U10 CA181009)

REFERENCES

1. McLaughlin-Drubin ME, Meyers J, Munger K. Cancer associated human papillomaviruses. *Curr Opin Virol.* 2012; 2:459–466. [PubMed: 22658985]
2. Forman D, de Martel C, Lacey CJ, et al. Global burden of human papillomavirus and related diseases. *Vaccine.* 2012; 30(Suppl 5):F12–F23. [PubMed: 23199955]
3. Chaturvedi AK, Engels EA, Pfeiffer RM, et al. Human papillomavirus and rising oropharyngeal cancer incidence in the united states. *J Clin Oncol.* 2011; 29:4294–4301. [PubMed: 21969503]
4. de Martel C, Ferlay J, Franceschi S, et al. Global burden of cancers attributable to infections in 2008: A review and synthetic analysis. *Lancet Oncol.* 2012; 13:607–615. [PubMed: 22575588]
5. de Villiers EM, Fauquet C, Broker TR, Bernard HU, zur Hausen H. Classification of papillomaviruses. *Virology.* 2004; 324:17–27. [PubMed: 15183049]
6. Munoz N, Bosch FX, de Sanjose S, et al. Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N Engl J Med.* 2003; 348:518–527. [PubMed: 12571259]
7. Ang KK, Harris J, Wheeler R, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. *N Engl J Med.* 2010; 363:24–35. [PubMed: 20530316]

8. Kimple RJ, Smith MA, Blitzer GC, et al. Enhanced radiation sensitivity in HPV-positive head and neck cancer. *Cancer Res.* 2013; 73:4791–4800. [PubMed: 23749640]
9. Tribius S, Ihloff AS, Rieckmann T, Petersen C, Hoffmann M. Impact of HPV status on treatment of squamous cell cancer of the oropharynx: What we know and what we need to know. *Cancer Lett.* 2011; 304:71–79. [PubMed: 21376458]
10. National comprehensive cancer network. Head and neck cancers. Version 2.2013. [Accessed July 14, 2014] Available at: <http://oralcancerfoundation.org/treatment/pdf/head-and-neckpdf>.
11. El-Naggar AK, Westra WH. p16 expression as a surrogate marker for HPV-related oropharyngeal carcinoma: A guide for interpretative relevance and consistency. *Head Neck.* 2012; 34:459–461. [PubMed: 22180304]
12. Romagosa C, Simonetti S, Lopez-Vicente L, et al. p16(Ink4a) overexpression in cancer: A tumor suppressor gene associated with senescence and high-grade tumors. *Oncogene.* 2011; 30:2087–2097. [PubMed: 21297668]
13. Hoffmann M, Ihloff AS, Gorogh T, et al. p16(INK4a) overexpression predicts translational active human papillomavirus infection in tonsillar cancer. *Int J Cancer.* 2010; 127:1595–1602. [PubMed: 20091864]
14. Sano T, Oyama T, Kashiwabara K, Fukuda T, Nakajima T. Expression status of p16 protein is associated with human papillomavirus oncogenic potential in cervical and genital lesions. *Am J Pathol.* 1998; 153:1741–1748. [PubMed: 9846965]
15. Klusmann JP, Gultekin E, Weissenborn SJ, et al. Expression of p16 protein identifies a distinct entity of tonsillar carcinomas associated with human papillomavirus. *Am J Pathol.* 2003; 162:747–753. [PubMed: 12598309]
16. Smeets SJ, Hesselink AT, Speel EJ, et al. A novel algorithm for reliable detection of human papillomavirus in paraffin embedded head and neck cancer specimen. *Int J Cancer.* 2007; 121:2465–2472. [PubMed: 17680565]
17. Singhi AD, Westra WH. Comparison of human papillomavirus in situ hybridization and p16 immunohistochemistry in the detection of human papillomavirus-associated head and neck cancer based on a prospective clinical experience. *Cancer.* 2010; 116:2166–2173. [PubMed: 20186832]
18. Schlecht NF, Brandwein-Gensler M, Nuovo GJ, et al. A comparison of clinically utilized human papillomavirus detection methods in head and neck cancer. *Mod Pathol.* 2011; 24:1295–1305. [PubMed: 21572401]
19. Cottrell CE, Al-Kateb H, Bredemeyer AJ, et al. Validation of a next-generation sequencing assay for clinical molecular oncology. *J Mol Diagn.* 2014; 16:89–105. [PubMed: 24211365]
20. Conway C, Chalkley R, High A, et al. Next-generation sequencing for simultaneous determination of human papillomavirus load, subtype, and associated genomic copy number changes in tumors. *J Mol Diagn.* 2012; 14:104–111. [PubMed: 22240447]
21. Rizzo G, Black M, Mymryk J, et al. Defining the genomic landscape of head and neck cancers through next-generation sequencing. *Oral Dis.* 2014 Apr 12. Epub ahead of print.
22. FUTURE II Study Group. Quadrivalent vaccine against human papillomavirus to prevent high-grade cervical lesions. *N Engl J Med.* 2007; 356:1915–1927. [PubMed: 17494925]
23. Paavonen J, Naud P, Salmeron J, et al. Efficacy of human papillomavirus (HPV)-16/18 AS04-adjuvanted vaccine against cervical infection and precancer caused by oncogenic HPV types (PATRICIA): Final analysis of a double-blind, randomised study in young women. *Lancet.* 2009; 374:301–314. [PubMed: 19586656]
24. Cimino PJ, Zhao G, Wang D, et al. Detection of viral pathogens in high grade gliomas from unmapped next-generation sequencing data. *Exp Mol Pathol.* 2014; 96(3):310–315. [PubMed: 24704430]
25. Yi X, Zou J, Xu J, et al. Development and validation of a new HPV genotyping assay based on next-generation sequencing. *Am J Clin Pathol.* 2014; 141(6):796–804. [PubMed: 24838323]

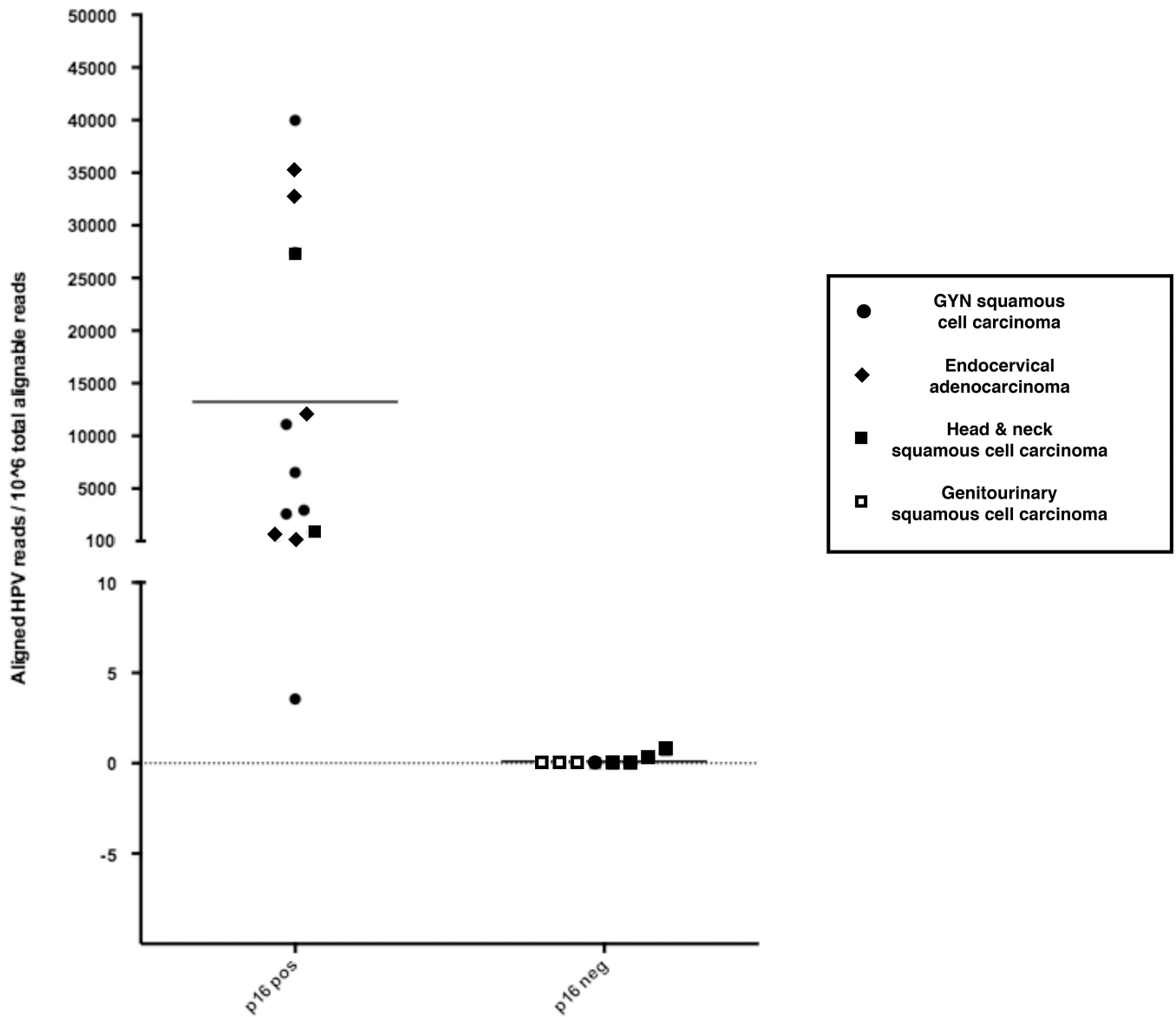


Figure 1. HPV sequence reads in p16-positive and -negative tumors

Vertical scatter plot demonstrating number of HPV reads per one million total aligned reads in p16-positive (n=13) and p16-negative (n=8) tumors. Solid circles: gynecological (GYN) squamous cell carcinoma; solid diamonds: endocervical adenocarcinoma; solid squares: head and neck squamous cell carcinoma; open circles: genitourinary squamous cell carcinomas. Mean values are indicated by horizontal lines (p16-positive = 13,231; p16-negative = 0.1; $p < 0.0001$). Only one tumor fell in the grey zone between 1 and 100 HPV reads per million aligned reads.

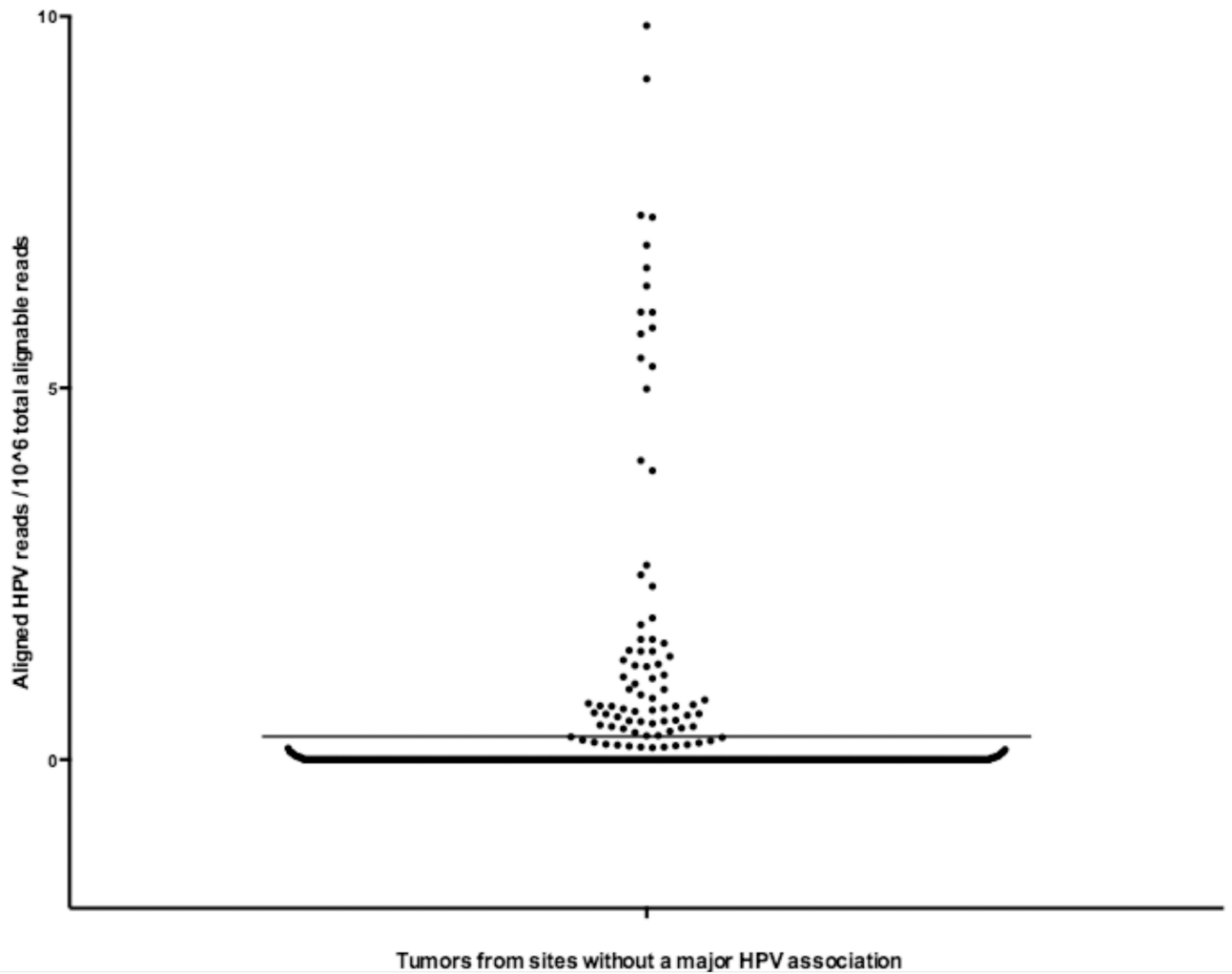


Figure 2. HPV sequence reads in tumors from anatomic sites without a strong HPV association
Vertical scatter plot demonstrates the number of HPV reads per one million aligned reads in 500 tumors unlikely to be HPV-associated based on anatomic site and histopathologic spectrum (i.e. all tumors in the UNCseq project other than the 21 endocervical adenocarcinomas and oropharyngeal, gynecological and genitourinary squamous cell carcinomas included in this study). The mean value (0.31 HPV reads per one million aligned reads) is indicated by the horizontal bar. The highest value (9.9 HPV reads per one million aligned reads) was seen in an ovarian serous carcinoma which was confirmed as HPV-positive by HPV 16/18 in situ hybridization.

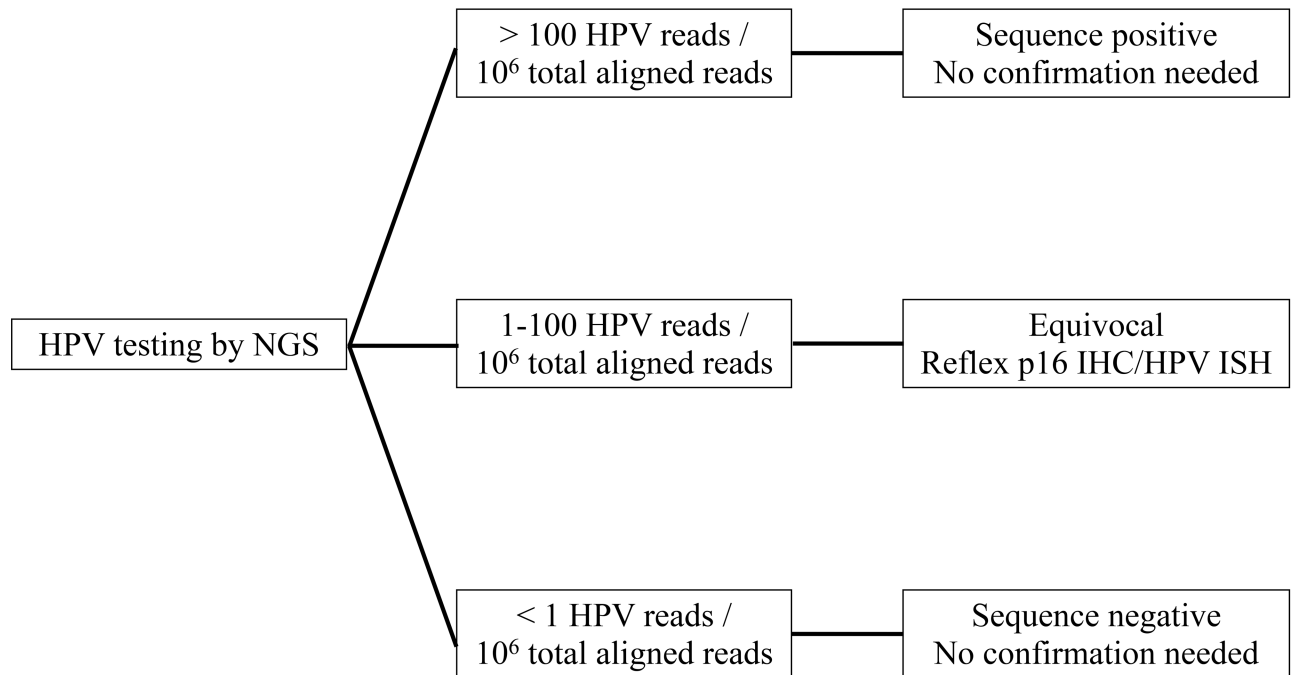


Figure 3. Proposed 2-step algorithm for HPV testing of tumor tissue by next generation sequencing

A proposed algorithm incorporates sequencing data to assess HPV status using capture probes and a reference sequence that is appropriate to actionable HPV strains for a given tumor type. In cervical cancer, such an algorithm should include all 14 high risk strains of HPV, whereas the pertinent strain types are less well defined for oropharyngeal and other genitourinary sites. Based on the scenario and data generated in the current study, sequence-positive tumors (defined by greater than 100 HPV reads per one million total aligned reads) and sequence-negative cases (defined by less than 1 HPV read per one million total aligned reads) may not require histochemical confirmation. On the other hand, tumors with intermediate results between 1 and 100 HPV reads per one million total aligned reads should reflex to p16 IHC or to HPV ISH for purposes localizing potential HPV infection to neoplastic cells. Further work is required to refine such an algorithm to fit the clinical scenario, population, specimen type, and test system.

Table 1

p16 status and tumor type

Tumor type	p16 positive cases (n = 13)	p16 negative cases (n = 8)
Gynecologic squamous cell carcinoma	6	1
Endocervical adenocarcinoma	5	0
Head and neck squamous cell carcinoma	2	4
Genitourinary squamous cell carcinoma	0	3

Histologic type and site of origin for 21 tumors stratified by p16 immunohistochemistry.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Concordance between p16 immunohistochemistry, HPV 16/18 *in situ* hybridization, and HPV 16/18 next generation sequencing results

	NGS positive HPV 16/18 ISH positive	NGS positive HPV 16/18 ISH negative	NGS negative HPV 16/18 ISH negative
p16 IHC positive (n = 4)	3	1	0
p16 IHC negative (n = 1)	0	0	1

p16 immunohistochemistry (IHC), high risk HPV *in situ* hybridization (ISH), and next generation sequencing (NGS) results for five cases included in this study. Cases with greater than 100 HPV reads / 10^6 total aligned sequenced reads were considered NGS positive.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

HPV strain typing by next generation sequencing

Case	HPV16 reads/ 10 ⁶ total aligned reads	HPV18 reads/ 10 ⁶ total aligned reads
Endocervical adenocarcinoma	0	208
Endocervical adenocarcinoma	6.1	540
Cervical squamous cell carcinoma	3.6	0
Head & neck squamous cell carcinoma	821	0
Cervical squamous cell carcinoma	2495	0
Vulvar squamous cell carcinoma	2938	0
Cervical squamous cell carcinoma	6520	0
Vulvar squamous cell carcinoma	11086	0
Endocervical adenocarcinoma	11961	0
Head & neck squamous cell carcinoma	27390	0
Endocervical adenocarcinoma	32804	0.2
Endocervical adenocarcinoma	35165	0
Vulvar squamous cell carcinoma	39968	0.1

Corrected numbers of HPV 16 and HPV18 reads (HPV reads / 10⁶ total aligned reads) for the 13 p16-positive tumors included in this study.