Chapter 5: Viral and Host Factors in Human Papillomavirus Persistence and Progression

Sophia S. Wang, Allan Hildesheim

Understanding the interdependent roles that host and viral factors play in cervical cancer pathogenesis is important for distinguishing women at the highest risk of human papillomavirus (HPV) persistence and progression to cervical cancer. Ongoing research on viral factors such as viral variants is providing important clues regarding HPV oncogenesis; the comprehensive characterization of the HPV genome and the function of viral genes by HPV type and variant will further this understanding. Although the biologic importance of viral integration and viral load measurements in cervical neoplasia is still being debated, available data are difficult to interpret because of methodologic limitations; to sufficiently address the importance of these events will require further methods validation and subsequent application in epidemiologic studies. Continued and expanded investigation of host immune responses—humoral, cellular, and innate immunity—should specifically address the outcomes of HPV persistence and progression to cervical cancer. Molecularly based assays paired with functional assays will be integral toward the identification and validation of key immune pathways and genes specifically relevant to cervical cancer pathogenesis. Novel technologies such as gene expression microarrays will further allow comprehensive identification of relevant genes that are important at various stages of cervical pathogenesis. The study of viral and host factors will undoubtedly lead to markers that may hold diagnostic and/or prognostic value; the clinical validity and utility of these molecular events will, therefore, need to be carefully assessed before implementation in a population setting.

In this chapter, both the viral and host factors that may play a role in the development of cervical cancer will be discussed, with a specific focus on factors that are important in the persistence and the progression of human papillomavirus (HPV) infections. Discussion of viral factors will include viral variants, viral load, and viral integration; discussion on host factors will include the host immune response, susceptibility genes, and identification of molecular events throughout cervical pathogenesis. For each of the viral and host factors discussed, a very brief summary of the current status of the data will be provided. An emphasis will be placed on issues relevant to defining future steps in each respective area, to identifying gaps in our present understanding, and to posing unanswered questions that might be addressed in future studies.

Viral Factors

Viral Variants

Although there are more than 100 HPV types identified, studies on viral variants have mainly focused on those for HPV type 16 (HPV16). Several studies (1) have now documented an association between HPV16 variants and the development of cervical cancer, with non-European variants being associated with excess risk of cervical cancer. It is now known that negative-association studies for HPV16 variants and cervical cancer were largely conducted in populations where the lower-risk European variant of HPV16 is predominant, whereas positive-association studies were more often conducted in populations where both European and non-European variants are observed. The very limited data available for HPV types other than HPV16 suggest that non-European variants of HPV types 18 and 58 are also associated with the increased risk of cervical cancer (2–5). Because of the relative rarity of these other types, a large multicenter effort and/or pooling across studies may be required to achieve statistical power. However, since viral variants have co-evolved with different ethnic groups in different geographic regions of the world (6–8), the possibility of population stratification (i.e., confounding by ethnicity) will need to be carefully considered in the planning efforts for such future studies. In this regard, highly admixed populations in which ethnic background is not associated with differential screening (and, therefore, not associated with the risk of cervical cancer development) would be ideal for future studies of this topic.

Data are also lacking on the patterns of variability in the regions of the HPV genome other than the upstream regulatory region (URR) and E6 regions (e.g., L1, L2, and E7). Future studies should consider whole-genome sequencing for select subsets of HPV-positive specimens to better define the extended haplotypes (e.g., regions on the genome that demonstrate non-random association). This information would assist laboratory colleagues in targeting viral genomic regions for in vitro functional studies. These functional studies are essential for providing biologic plausibility to associations that are observed in epidemiologic studies.

Viral Load

The potential utility of viral load measurements as an etiologic risk factor and as a diagnostic tool for cervical cancer continues to be debated. Cross-sectional epidemiologic studies (9–11) have demonstrated an association between increasing HPV viral load and the risk of cervical cancer. However, the longitudinal data evaluating patterns of viral load over time and the subsequent risk of progression of HPV infection to cervical intraepithelial neoplasms (CINs) 2 and 3 and cancer are insufficient (12,13). Therefore, although HPV viral load appears to be correlated to cervical cancer, its predictive utility for identifying...
progression to cancer from HPV infection has yet to be determined. Much of the uncertainty pertaining to viral load is due to the varying sampling techniques of specimens from which the viral load is tested, the varying methods used to measure the HPV viral load (e.g., kinetic polymerase chain reaction [PCR] versus Hybrid Capture II [HCII]), and the differences in cytohistologic classification in different studies. These discrepancies make the interpretation of existing data and the comparisons across studies difficult. For example, much debate ensues regarding the accuracy of the viral load measurements by HCII, particularly for multiple infections, because HCII is not an HPV type-specific assay. In addition, there exists the need to clarify whether the viral load measurements are the result of a few cells with large numbers of virions or large numbers of cells with few virions each (e.g., 1000 copies of virus in one cell or one viral copy each in 1000 cells). Interpretation of viral load may also prove to be challenging because of the transient nature of HPV infections; specifically, viral load estimates may be similar for long-term active HPV infections and for recent infections, making it difficult to delineate between the two.

Future studies regarding HPV viral load should incorporate recently developed state-of-the-art type-specific viral load measures such as kinetic PCR (14) and employ a careful diagnostic work-up of participants to avoid outcome misclassification. Without standardized specimen collection and testing methods, it is unlikely that more sophisticated statistical tests or studies will contribute substantially to the current data. In addition, because most studies have scientifically focused on HPV16 or overall (i.e., non-type-specific) viral load measures, extremely little is known about the relationship of viral load to types other than HPV16 and cervical neoplasia. Future studies should, therefore, be prospective, attempt to measure the viral load at multiple points over time, and expand beyond HPV16. Important scientific questions to be addressed include the determinants of high viral load, the interdependence of viral load measures and the appearance of low-grade morphologic abnormalities, the differences in viral load patterns seen for different HPV types, and the time-dependent association between the HPV viral load and the risk of cancer progression. Potential clinical applications for viral load should also be investigated further, such as whether viral load might be an indicator of a missed lesion among cytologically normal women.

**HPV Integration**

Although HPV is typically found in the episomal form in cervical lesions, viral integration has been reported to be associated with oncogenesis, and HPV integration into the host cell genome has been reported to occur in cervical cancer. However, it remains unclear whether HPV integration is random and biologically conveys no further risk downstream or whether the event confers a cellular growth advantage. Although HPV integration appears to be rare in HPV-infected cells, its occurrence may also be an irreversible event that initiates a chain of events including the impairment of tumor suppressor genes (e.g., p53 and Rb), subsequent genomic instability and cell immortalization (15). Accurate measurement of integration in association with progression to cancer and HPV types may, therefore, be warranted.

During integration, although E1 and/or E2 are frequently disrupted, the E6 and E7 viral oncoproteins are retained. HPV, therefore, appears to be not only stable during integration but also able to express viral oncogenes, plausibly contributing to cervical pathogenesis. The frequency of HPV integration appears to increase with the degree of disease severity, thus potentially correlating with progression to cervical cancer (16). However, the biologic importance of integration is still debated, with much of this uncertainty due to prior assay limitations. Until recently, methodologic difficulties existed in measuring HPV integration because the sites of integration are highly variable and because episomal and integrated forms of HPV often coexist. Currently, the DNA-based detection of integrated papillomavirus sequences by ligation-mediated PCR assay (DIPS–PCR) (17) and the amplification of papillomavirus oncogene transcript (APOT) test (16) appear to be promising for distinguishing the episomal from the integrated HPV. In addition, the potential to determine the site of integration (e.g., via fluorescent in situ hybridization) (18) may be a promising avenue of research. However, additional laboratory-based studies are needed to demonstrate the robustness of current methods and to identify the consistency of findings in cell lines and clinical specimens (e.g., reproducibility and validity) before these novel methods are applied to epidemiologic studies.

Given the limitations of prior assays from which available data are based, there exists sufficient uncertainty to warrant exploration with more recent assays. Applying validated methods of identifying HPV integration to epidemiologic studies could address numerous unanswered questions. These questions include determining how often integrated forms of HPV are found in precursor lesions of varying severity, determining whether lesions with integrated HPV possess higher risks of progression, determining whether patterns exist with respect to the site of integration (i.e., specific sequence patterns, integration in regulatory regions, and fragile sites as reported) in cervical cancers and precursor lesions, and determining whether rates and sites of integration vary by HPV type, by HPV variant, or by histology (squamous cell carcinoma versus adenocarcinoma).

Finally, another viral event of increasing interest and worthy of mention includes the assessment of epigenetic events in the HPV genome. Epigenetic events are those that alter gene expression (e.g., phenotype) without a change in the DNA sequence (e.g., genotype); these events include hypermethylation or hypomethylation of viral oncogenes (e.g., the addition or the removal of a methyl group) and the potential implications for suppression or activation of viral oncogenic expression, respectively. Continued investigation of the direct implications of epigenetic events on viral gene expression is warranted.

**Host Factors**

**Immune Response**

**Humoral immunity.** To date, epidemiologic studies have found a positive association between the detection of HPV antibodies (mainly HPV16 antibodies) and the risk of cervical neoplasia, in line with the thinking that HPV antibody detection is a marker of current and/or past exposure to HPV (19–21). Antibodies against HPV have been shown to be largely type specific. Although these antibodies, particularly those that target the proteins comprising the virion capsid (L1 and L2 proteins), might be effective at preventing infection, it is commonly accepted that antibodies are not important effectors of regression of established HPV infections and related cervical lesions (22).
Less clear is whether antibodies against one HPV type protect against subsequent reinfection with the same or another closely related type and, if so, whether this protection is related to specific antibody subsets (e.g., immunoglobulin G [IgG], IgG subclasses, and immunoglobulin A). Epidemiologic studies that have begun to address this question have focused largely on antibodies against L1 virus-like particles. Future studies could benefit by expanding the scope of inquiry to include antibodies against antigens other than L1 (e.g., L2, E2, and E7) and by evaluating not only the presence and the level of detectable antibodies but also the functionality of detectable antibodies (e.g., measures of binding kinetics, such as affinity and avidity).

Cellular immunity. In contrast to antibodies, T-cell responses to HPV have not been demonstrated to be type specific, and these cellular immune responses are likely to be an important effector mechanism for the clearance of established infections (23,24). Therefore, T-cell responses generated after infection might play a role in the protection against the progression of infection and early lesions. Although several studies (23,24), mostly cross-sectional, have been conducted to identify HPV-specific T-cell markers of protection, these markers remain elusive. This situation is due, at least in part, to the difficulties in reproducibly performing functional T-cell assays (i.e., misclassification) and to the inherently weak T-cell systemic responses seen with mucosal infections by HPV.

Studies, to date, have focused largely on T-cell responses to the oncogenic proteins of HPV16 (E6 and E7) and have evaluated the systemic levels of response. Future efforts should begin to evaluate responses against other HPV proteins, particularly those that are expressed consistently in the early stages of the natural history of cervical cancer (e.g., E2) (25). Attempts should be made to examine mucosal responses, in addition to systemic responses to HPV infection. Studies should also continue to examine viral mechanisms of immune evasion that would render an otherwise effective T-cell response unable to mediate resolution of established infections and/or lesions, such as previous work on the human leukocyte antigen (HLA) class I down-regulation and the zeta-chain expression in T lymphocytes.

Innate immunity. The first line of defense at the mucosal surface against infection is the innate immune response. Natural killer (NK) cells induce apoptosis or programmed cell death in virally infected cells and in tumor cells. Although altered HLA class I expression in cervical cancers has long been recognized, HLA class I allele findings in cervical cancer epidemiologic studies also now suggest evidence of NK cell involvement (26). This hypothesis is indirectly supported by a growing body of evidence regarding the role of innate immunity via NK cells in other immune-related diseases, such as the human immunodeficiency virus/acquired immune deficiency syndrome (27).

To measure the importance of NK cells requires the typing of killer immunoglobulin-like receptors (KIRs). There are inhibitory as well as activating KIRs, enabling NK cells to distinguish normal from virally infected or tumor cells. Because KIRs bind and recognize specific HLA class I alleles (28), the KIR–HLA complex is believed to regulate NK cell-mediated innate immunity. Normal cells that express abundant HLA class I molecules are thought to engage KIRs and to inhibit NK cell activation; in virally infected or tumor cells, however, HLA is down-regulated, leading to the activation of NK cells and cytolysis. How this relates to the known specificity of HLA–KIR binding has yet to be determined; however, associations observed between HLA class I alleles and the reduced risk of cervical cancer would make biologic sense if the appropriate KIRs were found in association with their respective class I allele. This finding might indicate specific receptor–ligand combinations that are more effective in the innate response against HPV infection and cervical cancer. To determine the role that NK cells play in HPV infection and in cervical cancer, the genotyping of activating and inhibitory KIRs in large epidemiologic studies that have conducted HLA class I allele genotyping may, therefore, be promising. Identifying an optimal marker for reliably measuring NK cell activity would further complement KIR genotyping efforts.

Genetic Susceptibility

In addition to the functional assessment of differences in the immune response to HPV infection, natural polymorphisms or genetic variations between individuals in immune-related genes might be of interest. Specifically, two avenues of investigation for identifying genetic susceptibility factors that should be considered are 1) identifying susceptibility factors for HPV persistence and 2) identifying susceptibility factors for progression to carcinoma.

To date, HLAs have been the most extensively studied immune-related genes. HLA molecules present foreign antigens to T lymphocytes and thus play a major role in the regulation of immune function. Results from available studies suggest a consistent protective role for HLA DRB1*1301 (1). This protection has been observed in multiple studies conducted in diverse populations. The type specificity of this finding is suggested by the fact that results are often strengthened by restriction to HPV16. In contrast to this protective allele, less consistency has been observed for HLA alleles that are hypothesized to increase the risk of cervical neoplasia. This result suggests the possibility that epidemiologic studies are better able to identify protective HLA alleles than risk alleles. Biologically, this possibility might be explained if a single allele suffices to confer protection (via effective binding and presentation of HPV to the immune system), whereas several HLA risk alleles are required before increases in risk can be detected (i.e., multiple HLA alleles that are unable to effectively present HPV to the immune system might be necessary to confer risk).

Because HLA genes are but a few of the many polymorphic genes involved in the immune response, future studies might benefit from an expansion to other immune-related genes. Given the complexity of the immune system, future efforts will likely require the use of high-throughput technology to assess the polymorphisms of numerous genes simultaneously, permitting the comprehensive assessment of immune gene pathways rather than a limited number of candidate genes. Such an effort would require a large study size for optimal statistical power, particularly for identifying alleles with modest frequency and hypothesized to confer modest risk.

With respect to HLA, future studies would benefit from approaches that move beyond individual allele–disease associations and begin to investigate biologically relevant groupings, such as those based on epitope motifs and the resulting binding sites. Also, a sufficient number of HLA studies have been conducted to date to warrant attempts to pool data across the studies, an effort that, in fact, is already under way. As mentioned in the “Viral Variants” section in this chapter, concerns of population
stratification need to be carefully considered in planning such efforts.

It is probable that, for cervical cancer, genetic involvement will consist of large numbers of alleles conferring modest levels of risks. Although these genes are difficult to study, distinct populations fortunately exist that may be informative for identifying such susceptibility genes. These include 1) women participating in vaccine trials to assess the heterogeneity of immune response to vaccination, 2) young women with rapid-onset disease to assess early exposures and genetic factors, 3) women with a high number of sexual partners who do not develop cervical neoplasia (e.g., prostitutes with repeated exposure to HPV infection but who have developed a natural immunity to HPV persistence), and 4) older women with persistent oncogenic HPV who do not develop cervical cancer. While several groups are attempting to identify specific genetic polymorphisms associated with cervical cancer [e.g., MTHFR (29), WAF1 (30), and IL-10 (31)], given the large effort and sample size requirement previously stated, we believe that the targeted study of these distinct populations could be particularly useful.

Finally, studies focusing on inherited susceptibility within families are of scientific interest. Although multiplex families, in which more than one woman in a family has cervical cancer, are rare, there is evidence that having a sister or a mother with cervical cancer increases a woman’s risk of cervical cancer twofold (32) and that heritability might explain some of the variation in cervical cancer risk (33). As alluded to earlier, identifying young women with rapid-onset disease (as opposed to the decade or more time to cancer onset after infection) may be a worthy area of pursuit, particularly given that early age of onset is a well-established risk factor for familial cancers. Although familial clustering where diseases or exposures are more common within families has not been observed for HPV16 positivity (34), identifying familial clustering for HPV persistence and/or progression to cancer among infected individuals may be worthy of future efforts. Also, familial studies on specific HPV and/or wart-related genetic diseases or highly immunocompromised populations, such as epidermodysplasia verruciformis (EV) and warts, hypogammaglobulinemia, infections, and myelokathexis (WHIM syndrome), can serve as important models for investigating the host genetics related to HPV infection and neoplasia. In the example of EV and HPV, the frequent malignant conversion seen at young ages indicates the importance of its use as a model for understanding the genetics of HPV-associated cancer (35).

Pathogenesis

It is accepted that genetic events, whether induced by or independent of HPV infection, are required for progression to cervical cancer. However, the specific events and the sequence of molecular events have not yet been determined for cervical cancer pathogenesis. Therefore, identifying the necessary changes during the natural history of cervical cancer, with a key focus on progression from low-grade squamous intraepithelial lesions/persistent HPV to high-grade squamous intraepithelial lesions/CIN 3, is a crucial area of research.

There is a substantial body of literature regarding chromosomal abnormalities in cervical cancer. There are numerous reports regarding the loss of heterozygosity (LOH) in cervical cancers, where the deletion of one of two alleles occurs; to identify LOH, many researchers have used comparative genomic hybridization analyses, where genetic changes on the chromosomal level such as LOH can be visualized (36,37). Although chromosomal events have been consistently identified, such as LOHs at 3p, chromosomes 6, 11, 13, 16, 17, and 19, and chromosomal gains at 3q, identifying the target genes (oncogenes/tumor suppressor genes) affected in these areas should now be the focus of current research (e.g., 3p LOH and the validation of the candidate tumor suppressor gene FHIT [fragile histidine triad] identified at chromosome 3p). Similarly, the inactivation of p53 and Rb gene products by E6 and E7 proteins has been well-studied, but the role of additional target genes (e.g., c-MYC, RAS, and telomerase/hTERT) needs to be clarified. The clarification of these processes will require careful assessment of HPV oncogenes and their role in generating cellular instabilities (e.g., genetic events or “hits”) required for carcinogenesis.

In addition, recent studies (38,39) have identified the silencing of tumor suppressor genes via promoter hypermethylation in HPV-infected host cells as a frequent human epigenetic event, a parallel but distinct event from viral gene methylation events described previously. Promoter hypermethylation, where a methyl group is attached to the promoter region of a gene resulting in the suppression of gene expression, is now commonly seen in human cancers and transformed cell lines, and its importance in cervical cancer warrants further attention, particularly in identifying key tumor suppressor genes that may be involved in progression. For both somatic and epigenetic changes, selecting specific populations of interest will be important; specifically, because both the somatic and epigenetic events are associated with increasing age, age will need to be carefully considered in the study design and analysis. Finally, recent technologic advances have provided the momentum needed to move forward research aimed at identifying genetic events that are important in cervical cancer pathogenesis. Genetic and protein-expression studies (40,41) using microarray technology should allow the characterization of the tumor microenvironment and the identification of gene families (e.g., immune genes) that are up-regulated or down-regulated at varying stages of cervical carcinogenesis. This novel technology enables the assessment of gene expression in thousands of genes simultaneously. Coupled with laser-capture microdissection, a method for procuring pure cell types from tissues, these technologies should soon allow for the careful examination of gene expression and other molecular events occurring in specific cells that represent the lesion of interest (42); for cervical cancer, this means the potential to carefully examine events specific to cells in the transformation zone, where squamous cell carcinomas of the cervix are known to specifically arise. Microarray analyses can be applied to identify the differences in pathogenesis with regard to both host and viral factors; these include comparing genetic events at different stages of the pathogenic process (e.g., cancer versus CIN 3 versus HPV infection), comparing cancer types (e.g., adenocarcinomas and squamous cell carcinomas), and comparing different HPV types and variants. Specifically, assessing molecular events in tissues of women with cancer compared with those of women diagnosed with CIN 3 will provide useful clues with regard to the necessary events for progression to cancer. Moreover, comparing cancer or CIN 3 with adjacent normal cells from the same women might also provide useful clues regarding genetic events occurring in cervical pathogenesis. Such methods of analyses should prove to be promising in yielding essential data regarding the necessary and/
or sufficient genetic events for cervical neoplasia. Although such methods hold great potential, the use of model systems, such as cell cultures and transgenic mice, is essential for carefully validating these methods before their widespread use in epidemiologic studies. The utility of such data in epidemiologic studies will further require specified parameters, such as specific populations, age definitions, and histologic types.

**Clinical Application of Candidate Markers**

Undoubtedly, identification of viral and host factors that contribute to the delineation of women at the highest risk of HPV persistence and progression to cervical neoplasia will lead to selected candidate markers that can potentially be used for diagnostic or prognostic purposes. The clinical validation and utility of candidate markers must, therefore, be thoroughly investigated. The clinical validity of a marker evaluated by the sensitivity, specificity, and positive and negative predictive value with the intended phenotype or disease outcome (43), such as HPV persistence or cancer, is crucial. These measures of validity will vary by population, because different prevalence of viral variants and host alleles are likely to exist. A subsequent evaluation of clinical utility to determine the benefits and risks of marker implementation within a population, considering cost, access to care, and other factors, is also needed (43). Priorities for evaluating a marker’s clinical utility will vary by the intended use of the marker (e.g., screening, diagnosis, and prognosis) and by population; for early cervical cancer detection in developed countries, an emphasis on marker specificity will likely be placed to prevent the overtreatment of HPV-infected women, whereas for developing countries, a simple and inexpensive test with high levels of sensitivity and specificity might likely be emphasized. Few markers currently warrant such evaluation; one such marker is p16INK4a. As a marker of oncogenic HPV expression, p16INK4a has been shown to be associated with HPV-infected high-grade lesions. To date, the sensitivity and specificity of p16INK4a with CIN 2 and 3 and normal tissue, respectively, appear to be high (44); however, the positive predictive value and sensitivity in the prospective follow-up for relevant outcomes (e.g., prognostic value for CIN 2 and 3) have yet to be determined.

**Conclusion**

The assessment of the aforementioned host and viral factors will hopefully permit us to distinguish between the majority of HPV-infected women who will clear their infection from the minority whose disease will progress to cervical cancer. Gaining a comprehensive understanding of the viral and host factors involved in cervical pathogenesis requires a happy marriage between epidemiology and laboratory colleagues.

**References**


