Chapter 1

Introduction and Review of Literature
1.1 Introduction

Dr. Hutchison’s words “Let us not make the cure of the disease more unbearable than the disease itself” had been one of the most enduring and inspiring lines in medicine. In spite of having being spoken almost a century ago, these profound words still find a haunting echo in today’s context, particularly when it comes to diagnosing and treating cancer. Keeping in agreement with its Latin meaning – a ‘crab’, the disease is too hard to crack just like the Crustacean. This is clearly reflected in the statistics that indicates the disease as one of the principal causes of deaths worldwide, including India, where it accounts for approximately 6% of total human deaths (Dikshit et al., 2012; GLOBOCAN, 2008a).

Cancer of the cervix is one of the most prevalent forms of cancer worldwide, the major burden of the disease being felt in developing countries like India. Although, early detection and advancement in diagnostic and treatment modalities have led to improved disease management and increased survival of patients in developed countries, in India, cervical carcinoma still continues to be the most common cancer among women and accounts for the maximum cancer deaths each year (GLOBOCAN, 2008b; IARC, 2009). Persistent infections with high-risk (HR) Human Papillomaviruses, such as HPV 16, 18, 31, 33 and 45 have been identified as a major, although not sufficient etiological factor in the development of the disease. The relationship between exposure to HPV and the disease might vary with the infecting HPV type, viral integration status and viral load; these factors could have profound implications on patient prognosis (Josefsson et al., 2000; Woodman et al., 2007).

Till date 120 official HPV types have been identified. However, the prevalence and distribution of HPV types in cervical neoplasias vary with geographic regions and by grade of disease (Clifford et al., 2005; Insinga et al., 2008; Smith et al., 2007).
Therefore, studies involving detection and genotyping of the virus in a population have become extremely important. At least 50% of sexually active men and women get HPV at some point in their lives, but in most cases the infection regresses spontaneously and only in about 10% cases it persists and progresses to high-grade cervical intraepithelial neoplasia. This generally occurs through integration of the HPV genome into the host chromosome with associated loss or disruption of E2 (Wentzensen et al., 2004). In the absence of E2 the expression of E6 and E7 increases eventually leading to immortalization and transformation of the cells (Romanczuk and Howley, 1992).

Besides infection by HPV, genetic instability has been reported as a significant event in disease pathogenesis. Genetic instability may range from point mutations, copy number changes, chromosomal rearrangements to widespread aneuploidy. Identification of such alterations through high throughput techniques such as Next Generation Sequencing would not only provide a detailed picture of the genomic landscape of the disease but would also be significant for understanding disease prognosis and response to therapy.

Studies addressing all the major aspects involved in pathogenesis of cervical cancer such as incidence of HPV infection and contribution of its different genotypes; physical state and site of viral integration combined with viral load in cervical lesions as well as genetic alterations would help in better understanding the disease and be a step forward towards identifying biomarkers and newer treatment modalities for management and cure of cervical cancer.
1.2 Review of Literature

With about 134000 new cases and 72800 deaths annually, cervical cancer ranks as the 1st most frequent cancer among women in India. Characterized by abnormal bleeding, pelvic pain and unusual heavy discharge, the disease develops in the tissues of the cervix, a part connecting the upper body of the uterus to the vagina. It comprises of endocervix or the upper part which is close to the uterus and covered by glandular cells; and the ectocervix, the lower part which is close to the vagina and covered by squamous cells. The two regions of the cervix meet at the ‘transformation zone’ (Fig. 1.1). It is this region where most cervical cancers begin to develop.

Based on the histology, carcinomas of the cervix may be classified into the following groups, with each group further having several morphological variants:

- **Squamous carcinomas** - Carcinomas arising from the ectocervical epithelium and characterized by the highest incidence rate, 85-90%.
- **Adenocarcinomas** - Carcinomas arising from endocervical columnar/glandular epithelium and constituting the remaining 10-15%
• **Adenosquamous carcinomas or mixed carcinomas** - Cancers with features of both squamous cell carcinomas and adenocarcinomas, constitute a small percentage (<1%) of the disease burden

### 1.2.1 Incidence of cervical cancer

Cervical cancer is the third most common cancer in women worldwide, and the seventh overall. Majority of this global burden is felt in low & middle income developing countries (WHO, 2009) and in low socio-economic groups within countries (Kurkure and Yeole, 2006).

![Fig. 1.2: Estimated cervical cancer cases in developed and developing countries. The incidence of cervical cancer in developing countries is approximately 72% higher than that of developed countries. (Globocan, 2008).](image)

Five year survival rates of less than 21% are reported from developing countries whereas rates as high as 70% are reported from developed regions like United States. India has a disproportionately high burden of cervical cancer, with the highest rate of incidence and mortality among Indian women (GLOBOCAN, 2008b; IARC, 2009).

![Fig. 1.3: Incidence and Mortality of cervical cancer in India. Cancer of the cervix has the highest incidence and mortality rate in Indian women (Globocan, 2008).](image)
The incidence of the disease begins to rise at ages 30–39 and peaks in the fifth or sixth decade of life (WCR, 2008). While the cumulative risk of incurring the disease before the age 64 is 1.19% for a female if world population is considered, it is 2.10% for an Indian female, making them highly susceptible group. The age-adjusted incidence is highest in Chennai and lowest in Thiruvanathapuram. There is also a high incidence belt in the north eastern districts of Tamil Nadu, as well as in two districts in the North-Eastern region of the country. Also, compared to the world and Southeastern Asia, the age specific mortality from the disease is highest in India, the five year overall survival rate being only 48% (Sankaranarayanan et al., 1998).

With the given rate of incidence and mortality, the estimated number of new cervical cancer cases and deaths in India is projected to increase by a large fraction by 2025 (WHO, 2009).

![Fig. 1.4: Estimated number of cervical cancer cases and deaths in India projected in 2025. The estimated number of cervical cancer cases and deaths is expected to increase by 86% and 87% in the age group of 65+ by 2025 (IARC Globocan, 2008).](image_url)

### 1.2.2 Natural History of Cervical Cancer

Cancer of the cervix marks its beginning with the development of pre-cancerous, benign lesions. WHO classification describes the first stage of development as mild dysplasia, which can then progress to moderate dysplasia, severe dysplasia, and carcinoma in situ (CIS) or invasive cervical cancer, with increasing degrees of severity. Mild dysplasia usually regresses spontaneously without treatment. However, in a small percentage of
women it progress to more severe forms. Women with moderate to severe dysplasia are at high risk of developing invasive cancer, although the progression may take several years. There are two other systems of classification—

**The Cervical Intraepithelial Neoplasia (CIN) system** (Buckley et al., 1982) is based on the degree of involvement of epithelial thickness by the atypical cells. According to this system, mild to moderate dysplasia are classified as CIN1 and involves lower one third of the mucosa, intermediate dysplasia involving two thirds of the mucosa as CIN2, and severe dysplasia and carcinoma in situ are together classified as CIN3, where the whole epithelial layer is replaced but with no disruption of the basal membrane.

**The Bethesda system** simplifies it further, by classifying CIN1 as Low Grade Squamous Intraepithelial Lesion (LSIL), and both CIN2 and CIN3 as High Grade Intraepithelial Lesion (HSIL).

The invasive form of carcinoma is further classified into various stages, as per the International Federation of Gynaecology and Obstetrics (FIGO) (Sankaranarayanan and Wesley, 2003). This classification is often used to decide the treatment options.

![FIGO staging of cervical cancer. The FIGO stages I-IV of cervical carcinoma has been depicted](temanggunggaul12.blogspot.com; Sankaranarayanan and Wesley, 2003).
### FIGO staging of cervical cancer

**Stage I:** Carcinoma strictly confined to the cervix.
- **Stage IA:** Invasive cancer identified only microscopically. Invasion is limited to measured stromal invasion with a maximum depth of 5 mm and no wider than 7 mm.
- **Stage IA1:** Measured invasion of the stroma no greater than 3 mm in depth and no wider than 7 mm diameter.
- **Stage IA2:** Measured invasion of stroma greater than 3 mm but no greater than 5 mm in depth and no wider than 7 mm in diameter.
- **Stage IB:** Clinical lesions confined to the cervix or preclinical lesions greater than Stage IA.
  - **Stage IB1:** Clinical lesions no greater than 4 cm in size.
  - **Stage IB2:** Clinical lesions greater than 4 cm in size.

**Stage II:** Carcinoma extends beyond the cervix but does not extend into the pelvic wall. The carcinoma involves the vagina, but not as far as the lower third.
- **Stage IIA:** No obvious parametrial involvement. Involvement of up to the upper two thirds of the vagina.
- **Stage IIB:** Obvious parametrial involvement, but not into the pelvic sidewall.

**Stage III:** Carcinoma that has extended into the pelvic sidewall. On rectal examination, there is no cancer-free space between the tumour and the pelvic sidewall. The tumour involves the lower third of the vagina.
- **Stage IIIA:** No extension into the pelvic sidewall but involvement of the lower third of the vagina.
- **Stage IIIB:** Extension into the pelvic sidewall or hydronephrosis or non-functioning kidney.

**Stage IV:** Carcinoma that has extended beyond the true pelvis or has clinically involved the mucosa of the bladder and/or rectum.
- **Stage IVA:** Spread of the tumour into adjacent pelvic organs.
- **Stage IVB:** Spread to distant organs.

### 1.2.3 Etiology of cervical cancer

Infection by HR-HPV has been identified as the key etiological factor for cervical carcinogenesis. The infecting HPV type, viral integration and viral load – all these have profound implication on the disease. Besides HPV infection, the disease is also associated with a genetic component. Genetic variations ranging from point mutations, single nucleotide polymorphisms and chromosomal abnormalities have been identified to contribute to disease development. Apart from these, several exogenous cofactors, such as hormonal contraceptives, smoking, parity, and co-infection with other sexually transmitted agents also play an important role. Each of these major factors is described:
1.2.3.1 Human papillomavirus (HPV) Infection

The association between HPV infection and carcinoma of the cervix has its origin in the works of Prof. zur Hausen in the early 1980s. Since then, several studies have established that persistent infection with HR-HPV is the major risk factor for the development of high-grade precancerous and cervical carcinoma (de Villiers et al., 1981; Gissmann and zur Hausen, 1980; Nobbenhuis et al., 1999). More than 30 to 40 types of HPV are typically transmitted through sexual contact and at least 50% of sexually active men and women get HPV at some point in their lives. Most HPV infections in young females are transitory and have little long-term significance. However, in 5-10% cases the infection persists and can progress to invasive cervical cancer. Besides being a key etiological factor for cervical cancer, there is growing evidence for the central role of HPV in oral carcinoma as well as cancers of other anogenital sites (Fig. 1.6).

![Image of HPV contribution in several HPV related cancer](image)

**Fig. 1.6: Frequency of HPV contribution in several HPV related cancer.** The annual number of cases worldwide for each of these is depicted. The fraction of cancers estimated to be induced by HPV is shown in red (Wikipedia).

1.2.3.1.1 HPV Genome

Papillomaviruses belong to the Papovaviridae family, characterized by a small non-enveloped DNA genome with a virion size of ~55 nm in diameter. The virus consists of a double stranded DNA genome (~7800-7900 bp) and an icosahedral capsid of 72 capsomers, which contain at least two capsid proteins, L1 and L2. The HPV genome can
be divided into three regions - the noncoding long control region (LCR) or the upper regulatory region (URR), and the protein encoding early (E) and late (L) gene region (Fig. 1.7). The three regions in all papillomaviruses are separated by two polyadenylation (pA) sites: polyA Signal 1 and polyA Signal 2.

![Fig. 1.7: Organization of the HPV genome. The circular form (a) and the simplified linearised form (b) are depicted. The ‘E’ and ‘L’ regions represent the early and late gene regions (Wikipedia and Expert Reviews in Molecular Medicine).](image)

Most of the cis-responsive elements that influence viral transcription and replication are located in LCR, a 400 to 1,000 bp genomic segment. The LCR contains a number of recognizable short motifs that show high conservation across the papillomavirus family. It regulates the production of viral proteins and particles by controlling viral transcription from the early and late regions. This region also contains overlapping binding sites for many different transcriptional activators and repressors such as AP1, glucocorticoid receptor, progesterone receptor NFI TEF-1 Oct-1, etc. which in turn can modulate viral gene expression. Immediately downstream the of the LCR, lies the early region that contains six open reading frames, E1, E2, E4, E5, E6 and E7 and is involved in viral replication and oncogenesis. This is followed by the late region genes. The major function of each of these is summarized in Table 1.1.
### Table 1.1: Function of the HPV genes

<table>
<thead>
<tr>
<th>Viral region</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1</td>
<td>Functions in origin recognition and exhibit both ATPase and 3’-5’ helicase activities. Works in conjunction with E2 protein and unwinds the DNA strands for viral replication; also linked to genome maintenance.</td>
</tr>
<tr>
<td>E2</td>
<td>Regulates viral replication and transcription by enabling E1 protein to bind to the viral origin of replication in the LCR; encodes a LCR-binding protein that regulates transcription of the early region and facilitates the correct segregation of genomes during cell division. High levels of E2 also have the ability to repress E6 and E7.</td>
</tr>
<tr>
<td>E4</td>
<td>Although E4 is part of the early region, it is expressed later in the virus life cycle, when complete virions are being assembled. The exact role played by E4 is not clear, but it is believed to interact with cytokeratin in the epithelial cells.</td>
</tr>
<tr>
<td>E5</td>
<td>Interacts with cell membrane growth factors such as EGF and PDGF and increases cellular proliferation and DNA synthesis. Also associated with an increase in mitogen-activated protein kinase activity and down regulation of major histocompatibility complex (MHC) class I molecules.</td>
</tr>
<tr>
<td>E6</td>
<td>Viral oncogene. Transforming potential is attributed to its ability to bind to p53 targeting its rapid degradation via ubiquitin ligase; and induction of telomerase activity.</td>
</tr>
<tr>
<td>E7</td>
<td>Viral oncogene. Has the ability to bind to hypophosphorylated form of Rb thereby promoting S phase entry; also interacts with inhibitors of cyclin dependent kinases</td>
</tr>
<tr>
<td>L1</td>
<td>Encodes the major capsid protein; functions in self-assembling into virus-like particles.</td>
</tr>
<tr>
<td>L2</td>
<td>Encodes the minor portions of the capsid and is expressed before L1 to allow for proper construction of the capsid. Has the ability to improve infectivity and packaging.</td>
</tr>
</tbody>
</table>

#### 1.2.3.1.2 Viral life cycle

HPVs are host-specific and show distinct tropism for squamous epithelial cells. The virus enters the epithelial basal layer through mild abrasion or microtrauma and infects the dividing basal cells. The nature of the receptor that facilitates viral entry is still a subject of debate, but it is generally agreed that heparan sulphate proteoglycans might function in the initial binding and/or virus uptake (Joyce et al., 1999; Shafti-Keramat et al., 2003). Besides heparan sulphate proteoglycans, efficient HPV infection is also believed to require secondary receptors such as α6 integrin, laminin 332 and syndecans (Abban and Meneses, 2010; Culp et al., 2006). After being taken into the cell, papillomavirus particles disassemble in late endosomes and/or lysosomes, eventually
transferring the viral DNA to the nucleus with the help of the minor capsid protein L2 (Day et al., 2004). Following infection, the early HPV genes E1, E2, E4, E5, E6 and E7 are expressed, and the virus establishes itself as a stable episome, maintaining its genome as a low copy number, in the basal cells of the epithelium (Fig. 1.8). This stage is mediated by the viral early genes E1 and E2. (Francis et al., 2000). In the suprabasal layers of the epithelium, the virus switches to a rolling circle mode of DNA replication and begins its productive stage (Flores and Lambert, 1997). This involves amplification of the viral genome to higher copy number by the E4 and E5 protein (Wilson et al., 2005), expression of the late genes encoding L1 and L2 that form the capsid proteins in upper epithelial layers and production of viral progeny (Fig. 1.8). In cervical lesions, the viral genome generally gets integrated with associated loss or disruption of E2 (Durst et al., 1985; Jeon et al., 1995). In absence of functional E2, the transcriptional control on the viral oncogenes, E6 and E7 is lost, thereby increasing their expression (Jeon and Lambert, 1995) and in turn proliferation of the suprabasal epithelial cells (Fig. 1.8).

**Fig. 1.8: HPV life cycle.** HPV infects the basal layer of epithelium. In the suprabasal layers E6 and E7 drive continued proliferation as cells differentiate and other early genes increase HPV gene expression and genome amplification. In the epithelial granular layers, the late L1 and L2 genes are expressed, forming infectious virions. With increase in disease progression, proliferation of the suprabasal epithelial cells increases mediated by increased expression of E6 and E7 (Doorbar, 2006).
1.2.3.1.3 Molecular Pathophysiology of HPV infection- Role of E6 and E7

The viral oncogenes, E6 and E7 interfere with the normal cell cycle machinery through biochemical interactions with several important cellular molecules.

Viral E6 binds E6-associated protein (E6-AP), an endogenous E3 ubiquitin ligase in the epithelial cells. This complex has the ability to bind the critical cell cycle protein p53 thereby targeting it for degradation by the 26S proteasomal pathway (Scheffner et al., 1993). In absence of p53, epithelial cells fail to recognize senescence or apoptosis signals and continue to divide throughout the stratified squamous epithelial layers. Besides, E6 has also been associated with degradation of the pro-apoptotic protein Bak and proteins containing PDZ domain such as human Scribble, MUPP - 1, and MAGI - 1, 2, and 3, as well as induction of the telomerase activity in the cervical cancer cells (Klingelhutz et al., 1996) (Fig. 1.9).

![Fig. 1.9: Role of E6 in pathophysiology of HPV infection.](image)

**Fig 1.9: Role of E6 in pathophysiology of HPV infection.** E6 targets p53, Bak and PDZ domain-containing proteins, for degradation. It also activates telomerase via induction of expression of the catalytic subunit, hTERT (modified from Jo et al., 2005).

E7 has been reported to function in tandem with E6 to drive continuous division of differentiated epithelial cells. It has the ability to bind to members of the ‘pocket protein’ family such as retinoblastoma protein (Rb) and target it for ubiquitin-mediated
degradation by the proteasome (Boyer et al., 1996). Degradation of pRb releases E2F, a transcriptional activator from the pRB/E2F complex, which then activates expression from S-phase promoters resulting in DNA replication and cellular division (Fig. 1.10). Besides, E7 can also associate and interfere with the activity of several endogenous molecules such as cyclin dependent kinase inhibitors p21cip1 and p27kip1, histone deacetylases and members from the AP1 family of transcription factors. HR- E7 can also activate the proto-oncogene DEK, which may be critical in HPV mediated malignant progression (Wise-Draper et al., 2005). E6 and E7 have also been reported to cooperatively disturb chromosome duplication and segregation during mitosis, thereby inducing severe chromosomal instabilities (Duensing and Munger, 2001).

![Fig. 1.10: Role of E7 in pathophysiology of HPV infection](image)

**Fig. 1.10: Role of E7 in pathophysiology of HPV infection** E7 targets retinoblastoma protein, for proteasomal degradation, releasing E2F which promotes S-phase entry of the cell cycle (modified from Jo et al., 2005).

The relationship between exposure to HPV and the disease might vary with the infecting HPV type, viral integration status and viral load (Josefsson et al., 2000; Woodman et al., 2007). Each of these parameters might have profound implications on patient prognosis and are described in greater detail.
1.2.3.1.4 HPV types

Currently, 120 different HPV types are officially recognized (Bernard et al., 2010). These can be classified into five evolutionary groups – Alpha, Beta, Gamma, Mu and Nu, based on whether L1 nucleotide sequence of one is at least 10% dissimilar from that of any other papillomavirus type. HPV types that infect the cervix belong to the group Alpha with over 60 members. The Alpha group also include cutaneous viruses such as HPV2, which cause common warts, and are very rarely associated with cancers. Those belonging to the Beta, Gamma, and Mu and Nu groups are primarily related to cutaneous infections (Fig. 1.11).

![HPV cladogram](image)

Fig. 1.11: HPV cladogram. Different HPV types are classified based on their sequence similarities. The genus alpha papillomaviruses primarily infect mucosal epithelium, while genus beta, gamma, mu and nu papillomaviruses primarily infect the skin (Doorbar, 2006).

Twenty four of the most common members of the Alpha family of papillomaviruses can further be divided into three categories high risk, putative high-risk types and low risk or cutaneous (Munoz et al., 2003; Schmitt et al., 2006) that include:

- 15 high-risk (HR) types - 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 73, 82
- 3 putative high-risk types - 26, 53, and 66
- 6 low-risk (LR) types - 6, 11, 42, 43, 44 and 70
However, the classification of HPV types into high-risk, putative high-risk and low-risk groups is often confusing especially for the weakly carcinogenic and rare HPV types, and most importantly due to the co-infection of multiple HPV types within the cervical epithelium (Schiffman et al., 2009; Wentzensen et al., 2009).

**Methods for HPV detection and genotyping**

Detection of HPV infection is quiet challenging since they cannot be cultured in conventional cell cultures, and serological assays for the detection of anti-HPV antibodies have limited accuracy. Moreover, classical direct virological diagnostic techniques, such as electron microscopy and immunohistochemistry, lack sensitivity as well as specificity for routine detection. Therefore, almost all HPV detection techniques that are presently being used, rely on the detection of HPV nucleic acids in a specimen (Poljak and Kocjan, 2010). The most common method for detection as well as HPV genotyping till date is PCR followed by sequencing. The primers used for the PCR can be either type specific primers (Baay et al., 1996) or universal primers that include degenerate MY09/11 and its modified version PGMY 09/11 (Gravitt et al., 2000; Manos et al., 1989), CP65/70, and internal primers, CP66/69 (Berkhout et al., 1995), consensus GP 5/6 and the modified version GP5+/6+ (Jacobs et al., 1997; van den Brule et al., 1990), OBI/II (Jenkins et al., 1991), CPI/CPII (Tieben et al., 1993), SPF1/2 (Kleter et al., 1998), and FAP primers specific for the cutaneous HPV types (Forslund et al., 1999). As opposed to type specific primers, general or consensus PCR primers have the ability to detect a broad spectrum of HPV genotypes. These general primers, also known as universal primers are mostly designed in the L1 region of the viral genome (except for CPI/II primers that are based in the E1 region) which is well-conserved across all HPVs.
Some of the commonly used general primers for detection of mucosal HPV infections are depicted in Fig 1.12.

![Diagrammatic illustration of the position of the different general primer sets on the HPV genome.](image)

**Fig. 1.12: Diagrammatic illustration of the position of the different general primer sets on the HPV genome.** The circular HPV DNA genome, ~8 kb in size is divided into early (E) and late (L) genes. The general primers are usually designed in the L1 region as this is well conserved across the HPV types. The positions of the amplification targets of these primer sets along with the expected amplicon sizes are depicted (modified from Kleter et al., 1999).

However, with advancement in technologies, multiplex detection and genotyping methods have become popular. Most of these technologies are based on PCR coupled with other high-throughput methods for detection as well as genotyping of multiple HPV types. Some of the important multiplex detection methods available for the alpha-HPVs are summarized in Table 1.2.
### Table 1.2: Important currently available methods for multiplex detection of HPVs

<table>
<thead>
<tr>
<th>Assay</th>
<th>Underlying principle</th>
<th>HPV types detected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hybrid Capture 2 (hc2) HPV DNA Test (Digene Corporation, USA)</td>
<td>Nucleic acid hybridization assay. Most frequently used HPV diagnostic assay. Does not allow exact determination of HPV type</td>
<td>13 HR-HPVs (cocktail B): 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59 and 68 and five low-risk HPVs (cocktail A): 6, 11, 42, 43 and 44.</td>
</tr>
<tr>
<td>Cervista HPV HR Test (Third Wave Technologies USA)</td>
<td>PCR amplification coupled to fluorescence detection. Does not determine the exact HPV types.</td>
<td>14 HPVs using 3 oligonucleotide probe sets: A5/A6 (51, 56 and 66), A7 (18, 39, 45, 59 and 68) and A9 (16, 31, 33, 35, 52, 58).</td>
</tr>
<tr>
<td>The Amplicor HPV Test (Roche Molecular Systems, USA)</td>
<td>PCR amplification and detection on microwell plates. Does not determine the exact HPV types.</td>
<td>same 13 HPV types as hc2</td>
</tr>
<tr>
<td>CareHPV Test (Qiagen)</td>
<td>Based on hc2 technology, allows quick detection (~3 h)</td>
<td>13 HPV types included in the original hc2 plus HPV66,</td>
</tr>
<tr>
<td>Cervista HPV 16/18 Test (Hologic)</td>
<td>Real-time PCR. Currently the only FDA approved HPV genotyping assay.</td>
<td>HPV16 and HPV18</td>
</tr>
<tr>
<td>digene HPV Genotyping RH Test RUO</td>
<td>PCR amplification combined with reverse line-blot hybridization.</td>
<td>18 HPV types: 6, 18, 26, 31, 33, 35, 39, 45, 51–53, 56, 58, 59, 66, 68, 70, 73 and 74.</td>
</tr>
<tr>
<td>The PapilloCheck HPV-Screening Test (Greiner Bio-One GmbH, Germany)</td>
<td>PCR combined with microarray.</td>
<td>Identifies 24 alpha-HPV types: 6, 11, 16, 18, 26, 31, 33, 35, 39, 40, 42–45, 51–53, -56, 58, 59, 66, 68, 70, 73 and 82.</td>
</tr>
<tr>
<td>The Multiplex HPV Genotyping Kit (Progen/Multimetrix, Germany)</td>
<td>PCR combined with bead-based xMAP technology</td>
<td>Detection and identification of 24 HPV types: 6, 11, 16, 18, 26, 31, 33, 35, 39, 42–45, 51–53, 56, 58, 59, -66, 68, 70, 73 and 82.</td>
</tr>
<tr>
<td>PreTect HPV-Proofer (NorChip, Norway)</td>
<td>Based on detection of viral E6/E7 mRNA transcripts by nucleic acid sequence-based amplification (NASBA).</td>
<td>5 most common HR-HPV types: 16, 18, 31, 33 and 45.</td>
</tr>
<tr>
<td>INFORM HPV (Ventana, USA)</td>
<td>In situ hybridization (ISH)</td>
<td>12 HPVs: 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58 and 66.</td>
</tr>
</tbody>
</table>

Using one or combination of these methods, a huge number of studies have been carried out across the globe that shed considerable light on the association of HPV infection...
with cervical cancer as well as contribution of different genotypes to disease pathogenesis.

**Incidence and prevalence of different HPV types in cervical cancer**

The prevalence and distribution of HPV types in cervical neoplasias vary with geographic regions and by grade of disease (Clifford et al., 2005; Insinga et al., 2008; Smith et al., 2007). Also, the magnitude of risk of cancer following infection is virus type specific (Chan et al., 1995; Munoz and Bosch, 1997; Munoz et al., 2003). Some of the recent major studies that highlight the global as well as Indian scenario of incidence of HPV infection and the prevalence of different HPV genotypes are summarized in Table 1.3.

**Table 1.3: Incidence and prevalence of HPV – Global and Indian scenario**

<table>
<thead>
<tr>
<th>Study</th>
<th>No. of cases</th>
<th>Method of detection</th>
<th>Prevalence (%)</th>
<th>Major HPV types (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Varela et al., 2011)</td>
<td>30,848</td>
<td>Meta-analysis of 243 studies published from 1990 to 2010.</td>
<td>89.9</td>
<td>16 (57), 18 (16), 58 (4.7), 33 (4.6), 45 (4.5), 31 (3.8), 52 (3.4), 35 (1.7), 59 &amp; 39 (1.3), 51 (1)</td>
</tr>
<tr>
<td>(de Sanjose et al., 2010)</td>
<td>10,575</td>
<td>PCR (SPF10 primers) and enzyme immunoassay Reverse hybridisation line probe assay (LiPA25)13</td>
<td>85</td>
<td>16 (61), 18 (10), 31 &amp; 33 (4), 45 (6), 52 (3)</td>
</tr>
<tr>
<td>(Ciapponi et al., 2011)</td>
<td>5540</td>
<td>Meta-analysis of 62 studies of invasive cervical carcinoma from Latin America and the Caribbean</td>
<td>89</td>
<td>16 (53.2), 18 (13.2), 31 (7.5), 58 (3), 33 (4.3), 45 (4.6), 52 (3.2).</td>
</tr>
<tr>
<td>(Deodhar et al., 2012)</td>
<td>113</td>
<td>Multiplex E7 PCR/APEX assay</td>
<td>92</td>
<td>16 (76), 18 (10.6), 59 (4), 33, 45 (2.7), 58 (4.4), 59 (3.5)</td>
</tr>
<tr>
<td>(Bhatla et al., 2008)</td>
<td>558</td>
<td>Meta-analysis of 9 studies from India</td>
<td>94.6</td>
<td>16 (63.3), 18 (15.6), 45 (6), 33 (5.4), 35 (5), 58</td>
</tr>
</tbody>
</table>
The above reports clearly bring out the fact that the incidence of HPV infection in carcinoma of cervix is more in the Indian subcontinent as compared to other countries. However, HPV16 and 18 are most the most common types universally and hence clinically more important.
Rather than simply detecting the presence of virus, studies involving further characterization of the infection status such as viral load and its physical status would be much more informative.

1.2.3.1.5 Viral load

A high viral load has been found to be directly associated with the persistence of HPV infection (Dalstein et al., 2003; Ho et al., 1998). Also, there are reports stating direct correlation of viral titre with disease stage, being gradually increased from mild dysplasia to cervical cancer (Abba et al., 2003; Hernandez-Hernandez et al., 2003; Josefsson et al., 2000; Swan et al., 1999; van Duin et al., 2002). However, most of these findings are in context of HPV16 and it has been observed that the viral load of other HPVs such as HPV18, 31 and 45 does not increase with increasing disease severity (Gravitt et al., 2003; Ho et al., 2005; Swan et al., 1999), indicating a type specific relation between copy number and disease prognosis. Further, in the study by Singh et al. in a North Indian cohort, although significant increase in viral loads for HPV16 and 18 was observed from controls through SILs to tumours, but no significant differences was detected between different stages of cancer (Singh et al., 2009). In this context it has also been argued that it is indeed the E6 and E7 expression and not the viral load that is associated with disease prognosis (de Boer et al., 2007). An altogether different finding was presented by Kim et al. wherein they showed that patients with lower HPV viral load showed worse disease free survival (Kim et al., 2009) after radiation therapy. These conflicting results might be due to the variation in sampling techniques and different methods used to calculate viral load. Nonetheless, the relationship between viral load and disease is far more complex than was initially thought. The picture is further complicated by another important characteristic of the virus – viral integration.
While with increasing disease severity, the integrated forms of the virus become dominated, integration itself is followed by a decrease in viral load. In most of the studies that have measured viral load, integration status is not defined. Considering these facts, it becomes apparent that studies involving robust quantitation of type specific viral load in samples with known physical status of the virus could provide useful insights into the pathophysiology of HPV infections and their relationship to disease.

1.2.3.1.6 Viral integration

An important event in cervical carcinogenesis is the integration of the HPV into the host genome. Viral integration has been looked upon as a significant event in progression of cervical cancer from precancerous lesions to invasive carcinoma (Durst et al., 1985; Kalantari et al., 2001; Klaes et al., 1999). The frequency of viral integration into the host genome in cervical carcinomas has been reported to be as high as 100% in HPV18 positive tumours (Corden et al., 1999; Cullen et al., 1991; Pirami et al., 1997) and up to 80% in HPV16 positive tumours (Cullen et al., 1991; Melsheimer et al., 2004; Pirami et al., 1997). However, recent reports have confirmed presence of only episomal form of the virus in advanced cervical squamous cell carcinomas, thereby establishing that integration might not be absolutely mandatory for the process of carcinogenesis (Gray et al., 2010; Vinokurova et al., 2008).

Studies have also been undertaken to study the association between physical state of the virus and disease prognosis. The reports are quiet conflicting; while some report that the integration event is associated with a decreased disease free survival (Kalantari et al., 1998; Vernon et al., 1997) there are others according to which physical state of the virus does not correlate with disease free survival (Holm et al., 2008; Nambaru et al.,
2009). Whatever the case may be, the event of viral integration by itself is a significant episode in cervical carcinogenesis.

**Methods of detecting HR-HPV integration**

Several strategies have been used to study viral integration. Existing methods can be divided into two broad categories (Pett and Coleman, 2007):

a) Detection of viral integrants those are transcriptionally active by techniques such as Amplification of Papillomavirus Oncogene Transcripts assay (APOT) (Klaes et al., 1999), based on 3’ RACE-PCR and RNA ISH (Van Tine et al., 2004)

b) Detection of integrated viral DNA irrespective of its transcriptional status by Southern blotting (Cullen et al., 1991), Quantitative real-time PCR (Peitsaro et al., 2002), Ligation-mediated PCR (Luft et al., 2001), Restriction site-PCR (Thorland et al., 2000) and DNA ISH (Adler et al., 1997; Evans and Cooper, 2004)

**Significance of viral integration**

The event of integration is characterized by deletion of viral genes essential for synthesis of an infectious virion, and hence is not a normal part of the HR-HPV life cycle. It generally occurs in the E2 region and results in complete or partial disruption of the open reading frame (ORF) for E2 (Woodman et al., 2007). However, the viral oncogenes E6 and E7 oncogenes together with the viral upstream regulatory region are always retained. The integrant derived transcripts usually comprise of viral sequences at their 5’-ends and cellular sequences at their 3’-ends (Type A transcript). In certain rare cases, however, the viral genome either gets disrupted within the E4 region resulting in viral-cellular fusion transcripts that comprises of E6-E7-E1 sequences at their 5’-ends followed by E4 sequences and cellular sequences at their 3’-ends (Type B transcript); or
directly read through from viral to cellular sequences within the E1 gene (Type C transcript) (Fig. 1.13). As opposed to this, transcripts derived from the episomal form mostly comprise of the E1-splice donor signal spliced to the E4-splice acceptor site and terminated at the viral polyadenylation site (Fig. 1.13).

![Fig. 1.13: Common types of viral transcripts generated from viral - cellular fusion. Integration may occur in any of the three ways-Type A, B and C; type A being most common (modified from Klaes et al., 1999).](image)

The significance of viral integration in the context of cervical carcinogenesis may be due to the following reasons:

- Loss of viral E2 not only inhibits transcription from the integrated viral promoter (Dowhanick et al., 1995; Hwang et al., 1993), but also releases the transcriptional control on E6 and E7, thereby resulting in their increased expression.
- Since hTERT expression is inhibited by E2 (Lee et al., 2002) and activated by E6 (Veldman et al., 2001), HR-HPV integration activates telomerase and, along with E7, brings about immortalization of epithelial cells (Kiyono et al., 1998).
Disruption of the viral genome results in the failure of early gene transcription from the viral early polyadenylation signal. This leads to the use of host poly(A) signals and generation of stable virus–host fusion transcripts with a longer half life (Couturier et al., 1991; Di Luca et al., 1986; Jeon and Lambert, 1995), which imparts the cells with a selective growth advantage (Jeon et al., 1995).

**Fig. 1.14: Significance of HR-HPV integration in cervical cancer.** Viral integration results in the disruption of E2 ORF. In absence of E2, expression of the oncogenes E6 and E7 increases, imparting a proliferative advantage to the cells. Loss of E2 combined with overexpression of E6 activates telomerase, shifting the cellular balance towards immortality. Increased Levels of these oncogenes can also bring about genomic instability (adapted from Pett and Coleman, 2007).

In addition to the effects mentioned above, integrated viral genes may activate cellular oncogenes or inactivate tumour suppressive genes though insertional mutagenesis, eventually leading to alterations in cellular growth and proliferation. For example, APM-1, a putative tumour suppressor has been reported to get inactivated in a number of cervical carcinoma cells lines as a result of insertional mutagenesis combined with the deletion of its second allele (Reuter et al., 1998). Also, in a number of tumours, viral integration has been observed near the MYC locus (Ferber et al., 2003b; Peter et al., 2006) and into the human telomerase reverse transcriptase (Ferber et al., 2003a). Besides, viral integration can render both viral coding genes as well as the
cellular genes susceptible to epigenetic changes which could regulate their expression. Overall, the process may be looked upon as an event that promotes cellular immortalization, deregulated proliferation, and increased genomic instability, all of which are the cellular hallmarks of cancer.

An additional important and largely overlooked situation arises in case of a cell harbouring mixed forms of the virus, i.e. both integrated and episomal. In such cases, E2 may be available in trans to modulate the expression levels of oncogenic E6 and E7, and that overcoming this inhibition would represent an important event on part of the virus for selection of its integrated form (Arias-Pulido et al., 2006) Therefore, as demonstrated by Pett et al., loss of episomes is as much important as integration of the virus into the host genome for progression of lesions to cervical neoplasia (Pett et al., 2006).

‘Hotspots’ of viral integration- whether a random event

Studies on viral integration have identified integration sites to be distributed throughout the host genome, with no specific preference for any chromosomal loci (Wentzensen et al., 2004). However, according to more recent reports, although the integration event encompasses almost all the chromosomal loci, certain hotspots can be identified such as 3q28, 4q13.3, 8q24.21, 13q22.1, 1p36.23, 1p36.33, 3q26.33, 3q28, 6q22.31, 6q23.3, 8p11.21, 9q22.32, 13q22.1 and 20q11.21 (Das et al., 2012; Kraus et al., 2008; Schmitz et al., 2012). Besides, the virus has been shown to prefer fragile sites, translocation break points and transcriptionally active regions for its integration (Koopman et al., 1999; Thorland et al., 2000; Wentzensen et al., 2004; Ziegert et al., 2003). Fragile sites are specific regions in the chromosomes that nonrandomly undergoes break in response
to certain stress, making them as susceptible targets for foreign DNA integration. These can be divided into two types:

- Rare fragile sites (RFSs) – present in less than 5% of the population and are associated with expanded CCG repeats or expanded AT-rich minisatellite repeats (Richards, 2001).
- Common fragile sites (CFSs) - present in all individuals and the mostly induced in response to treatment with the DNA polymerase inhibitor, aphidicolin (Thorland et al., 2003)

A high correlation has been reported between the fragile sites and the HPV integration sites, owing mainly to their accessibility to the viral DNA (Dall et al., 2008; Kraus et al., 2008; Matovina et al., 2009; Thorland et al., 2003; Thorland et al., 2000; Wentzensen et al., 2002). Apart from the fragile sites, several tumour related genes such as myc, TP63, NR4A2, APM-1, FANCC, TNFAIP2, and hTERT have been shown to be involved in the viral integration process (Wentzensen et al., 2002). Further, a recent study has highlighted an association of site of HPV integration with micro RNAs (Schmitz et al., 2012).

1.2.3.2 Genetic Alterations

Though infection by HPV has been established as a major etiological factor for the genesis of cervical cancer, it may not be sufficient for tumour development. The evidence for this is based on the fact that the disease develops in a small proportion of women who have been infected with HPV and generally arises decades after the initial HPV exposure. Therefore, the possible role of genetic aberrations contributing to malignant transformation and tumour progression cannot be ruled out. These alterations may include chromosomal abnormalities, point mutation (both somatic and germline),
single nucleotide polymorphisms and allelic discrimination that might either be involved directly in causing the disease or increase the susceptibility to developing it. The commonly reported alterations in cervical carcinoma are summarized in Table 1.4 and 1.5.

**Table 1.4: Major genetic alterations associated with cervical cancer**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Nature of alteration</th>
<th>Alterations reported</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLA class II antigen</td>
<td>Allelic variation</td>
<td>DQB1<em>03, DRB1</em>1501 and DQB1<em>0602, DRB</em>13 and DQB1*0603 alleles</td>
<td>(Hildesheim et al., 1998; Madeleine et al., 2002).</td>
</tr>
<tr>
<td>LKB1/STK11</td>
<td>Missense and Nonsense point mutation; Deletion Frameshift</td>
<td>S19X,P6Q, E57X, 140611_1420 del CTC TGTCAGGG AAATTCAACTACT H107R, P294fs, F298L, R304W</td>
<td>(Wingo et al., 2009); COSMIC</td>
</tr>
<tr>
<td>CDKN2A</td>
<td>Missense point mutation; Deletion</td>
<td>codon 74 (CTG→ATG); codon 129 (ACC→ATC); del exon 2–intron 2 region, etc.</td>
<td>(Nakashima et al., 1999; Park et al., 1999); COSMIC</td>
</tr>
<tr>
<td>TP53</td>
<td>Point mutation, Deletion; Frameshift Polymorphism</td>
<td>Codon 234 (TAC→TGC), Codon 273 (CGT→TGT); Codon 170 (ACG→CCG); Codon 154 (GGC→AGC); Codon 297 (CAC-CGC); Codon (AGC→ACC); Codon 273GCT→TGT; codon 166 (A insertion, frameshift); codon 175 (CGC→CAC); codon 192 (CAG→TAG); Codon 213 (CGA→CGG); R72P</td>
<td>(Andersson et al., 2006; Busby-earle et al., 1994; Crook et al., 1992; Ikenberg et al., 1995; Kim and Kim, 1995; Liu et al., 1994; Miwa et al., 1995; Storey et al., 1998; Tenti et al., 1998); COSMIC</td>
</tr>
<tr>
<td>PTEN</td>
<td>Nonsense point mutation, Deletion</td>
<td>R233X, 800delA, 415delTATT, 1038del16 bp, etc.</td>
<td>(Harima et al., 2001; Minaguchi et al., 2004; Yaginuma et al., 2000); COSMIC</td>
</tr>
<tr>
<td>CTNNB1</td>
<td>Point mutation, Frameshift</td>
<td>Codon 45 (TCT→TTT); Codon 47 (AGT→GGT); Codon 37 (TCT→ACT); Codon 45 TCT→TT Frameshift; Codon 41 (ACC→AAC); Codon 37 (TCT→TTT).</td>
<td>(Shinohara et al., 2001); COSMIC</td>
</tr>
<tr>
<td>FGFR3</td>
<td>Point mutation</td>
<td>S249C</td>
<td>(Yee et al., 2000); COSMIC</td>
</tr>
<tr>
<td>PI3KA</td>
<td>Missense point mutation</td>
<td>R88Q; E542K; E545K; Q546E; H1047R; M1043I</td>
<td>(Cui et al., 2009; Miyake et al., 2008); COSMIC</td>
</tr>
<tr>
<td>Ras (H-ras, K-ras, N-ras)</td>
<td>Point mutation</td>
<td>Codon 12, 13 and 61</td>
<td>(Grendys et al., 1997; Huang et al., 1996; Kang et al., 2007; Lee et al., 1996; Pappa et al., 2009).</td>
</tr>
<tr>
<td>Nature of alteration</td>
<td>Chromosomal region</td>
<td>Study</td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>---------------------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td><strong>Amplifications</strong></td>
<td>3q21, 3q26–q29, 5p, 7q22, 8q23–q24, 8q24.3, 9p22, 9p23–24, 10q21, 11q13, 11q21, 11q22–23, 12p13, 14q12, 17q12, 17q25, 19q13.1, 20q11.2, 20q13.1</td>
<td>(Maliekal et al., 2003)</td>
<td></td>
</tr>
<tr>
<td><strong>Losses</strong></td>
<td>2q33–q37, 3p12–23, 4p16.3–p16.1.4, 4q, 4q28.3–q32.1, 4q13.3, 4q35.2, 4q28, 6q, 8p23.3, 8p12–21.3, 8q23.2–q23.3, 9p, 11p15.5, 11q13.3, 11q22.3–25, 3q12.11–13q14.3, 13q14.3–q21.33, 13q31.1–q31.3, 17p13.3, 18q11.2–18q23</td>
<td>(A et al., 2004; Acevedo et al., 2002; Bethwaite et al., 1995; Bhattacharya et al., 2004; Choi et al., 2007; Chuaqui et al., 2001; Herrington et al., 2001; Kersemaekers et al., 1998; Manolaraki et al., 2002; Miyai et al., 2004; Narayan et al., 2003; O'Sullivan M et al., 2001; Pulido et al., 2000; Sherwood et al., 2000; Tsuda et al., 2002)</td>
<td></td>
</tr>
</tbody>
</table>

Table 1.5: Major chromosomal aberrations associated with cervical cancer
The mutations involved in cervical cancer, described so far are mostly based on candidate gene approach and till date there are no reports describing the entire genomic landscape of the disease. With advances in technology, it is now possible to carry out such high-throughput studies.

1.2.3.2.1 Next generation sequencing - a new face of cancer research

During the past 5 years, ‘next generation’ sequencing technologies or NGS have surfaced as a promising tool for genomics research. With the advent of NGS, cancer genomics have moved from focused approaches based on single-gene sequencing and arrays to comprehensive genome-wide approaches. Contrary to Sanger, NGS is based on sequential imaging of the stepwise addition of nucleotides, resulting in the generation of hundreds of megabases to gigabases of nucleotide sequence output in a single instrument run, depending on the platform. Some of the popular NGS platforms available today include Roche 454 GS FLX, Illumina Genome Analyzer, Applied Biosystems SOLiD, Helicos BioSciences HeliScope, etc.

NGS technologies broadly include whole genome sequencing, whole exome sequencing and transcriptome sequencing. Of these whole exome sequencing has gained popularity over others in identifying disease mutation. Though exons constitute only about 1% of the genome (37.6 Mb), sequencing them can yield significant information as they have been reported to harbor most variations (Botstein and Risch, 2003). Also most frequent type of disease mutations are those that cause amino acid substitutions resulting from variations in exons. Further, approaches involving targeted sequencing provide increased sequence coverage of a particular region of interest at high throughput and lower cost compared to whole genome sequencing making it highly suitable for understanding the genetic landscape of cancer (Meyerson et al., 2010). Some of the
major recent studies where exome sequencing has successfully been employed for identifying novel mutations are summarized in Table 1.6. In addition to these genetic alterations, studies have identified an epigenetic component in cervical cancer. Promoter CpG island hypermethylation in genes such as p16, DAPK, MGMT, APC, HIC-1, E-cadherin RARβ, FHIT, GSTP1 and hMLH1 have been associated with cervical carcinogenesis (Dong et al., 2001; Virmani et al., 2001).

Table 1.6: Mutations in different cancers as identified by whole exome sequencing

<table>
<thead>
<tr>
<th>Type of cancer</th>
<th>Gene harbouring mutation</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myelodysplasia</td>
<td>RNA splicing machinery genes-U2AF35, ZRSR2 and SRSF2</td>
<td>(Yoshida et al., 2011)</td>
</tr>
<tr>
<td>Head and Neck Squamous cell Carcinoma</td>
<td>Notch1</td>
<td>(Agrawal et al., 2011)</td>
</tr>
<tr>
<td>Prostate</td>
<td>No. of mutations in 20 genes including TP53, PDZRN3, SDF4, etc.</td>
<td>(Kumar et al., 2011)</td>
</tr>
<tr>
<td>Renal cell Carcinoma</td>
<td>PBRM1</td>
<td>(Varela et al., 2011)</td>
</tr>
<tr>
<td>Acute Monocytic Leukemia</td>
<td>DNMT3A</td>
<td>(Yan et al., 2011)</td>
</tr>
<tr>
<td>Primary Colon cancer</td>
<td>BMPR1A</td>
<td>(Timmermann et al., 2011)</td>
</tr>
<tr>
<td>Multiple Endocrine Neoplasia Type 2A</td>
<td>RET Germline mutation</td>
<td>(Qi et al., 2011)</td>
</tr>
<tr>
<td>Uveal Melanomas</td>
<td>BAP1</td>
<td>(Harbour et al., 2010)</td>
</tr>
<tr>
<td>Kaposi Sarcoma</td>
<td>STM1</td>
<td>(Byun et al., 2010)</td>
</tr>
</tbody>
</table>

1.2.4 Molecular biomarkers in cervical cancer – from diagnosis to prognosis

‘Cancer can be cured, if detected early’- this paradigm is apt in the context of cervical carcinoma. Although, early detection combined with improved treatment modalities can help managing the disease and control the death toll to a very large extent, the existing screening and treatment modalities face issues with specificity and sensitivity. Hence, in recent times the main focus of cervical cancer research have shifted towards
identification of molecular biomarkers, which, combined with the existing screening and treatment procedures, are expected to improve diagnosis, prognosis, prediction of response or recurrence, and disease monitoring. A number of biomolecules including HPV oncogenes E6 and E7 are effectively being used as diagnostic biomarkers. E6 and E7 interfere with the cell cycle regulators and bring about a change in the expression pattern of a large number of molecules, thereby contributing to neoplastic progression. Some of the important biomarkers of cervical carcinoma include mini chromosome maintenance (Stoeber et al., 2002), cell division cycle protein 6, p16INK4A (Murphy et al., 2003), squamous cell carcinoma antigen (serum markers of squamous cell carcinoma) (Duk et al., 1990; Farghaly, 1992) and cell proliferation markers- PCNA and Ki-67 (Konishi et al., 1991; Mittal et al., 1993).

Besides these diagnostic biomarkers, molecules such as VEGF and EGFR are being used as potential therapeutic targets in cervical carcinoma. VEGF inhibitors including monoclonal antibodies, such as bevacizumab and small-molecule tyrosine kinase inhibitors (TKIs), such as sunitinib, could be successfully used in clinics either in the form of combinatorial therapy or monotherapy for treating locally advanced or recurrent disease (Monk et al., 2009). Similarly treatment regimen involving targeting of EGFR by small molecules such as cetuximab, gefitinib, erlnotinib or anti-EGFR antibody matuzumab, in combination with standard radiotherapy and chemotherapy protocols are in various phases of clinical trial (Goncalves et al., 2008).

The discovery of these and other plentiful molecular biological markers for diagnostics, therapeutics and prognosis has paralleled advances in high-throughput molecular biologic techniques in the genomic, transcriptomic, and proteomic fields. The use of Next generation sequencing technologies in this context has already been described. Some of other high-throughput technique includes Array Comparative
Genomic Hybridization for identification of chromosomal copy number changes, Single Nucleotide Polymorphism (SNP) Profiling, Gene Expression Profiling, Proteomics and Metabolomics. Of all the high-throughput technologies, the field of microarray technology have witnessed the maximum growth. It has been used extensively for classification, subclass identification, identification of prognostic biomarkers and predictive signatures in the field of cancer biology. Identification of HER2/neu (ERBB2) in metastatic breast cancer and EGFR in metastatic colorectal cancer as potential drug targets as well as determination of gene signatures to predict aggressiveness in prostate cancer and survival in colon cancer are some of the recent achievements of microarray. Not only this, two diagnostic arrays, Oncotype DX (Genomic Health, USA) and MammaPrint assay (Agendia, The Netherlands) have even found their way to the clinics for prediction of breast cancer recurrence and response to therapy.

In the field of cervical cancer, a plethora of biomolecules have been identified from a large number of studies that focuses on diagnostic, predictive and prognostic biomarkers. Differential gene expression pattern has been reported between squamous cell carcinoma and adenocarcinoma of the cervix (Chao et al., 2006; Contag et al., 2004). Besides, overexpression of EGFR, ERBB2, CDKN2A, KRAS, MYCN, KIT, TOP2A have been associated with malignancy of cervical epithelium. Apart from this, microarray has also been employed in cervical cancer to predict treatment response to radiotherapy or concurrent chemo-radiotherapy as well as prognosis after therapy. A number of studies have come up in recent time with different gene signatures in important pathways such as MAPK, apoptosis, metastasis, hypoxia, b-catenin, etc. that can predict response to therapy as well as prognosis (Harima et al., 2003; Huang et al., 2011; Huang et al., 2012; Iwakawa et al., 2007; Rajkumar et al., 2009).
Overall, the availability of high-throughput technologies has vastly broadened the potential for biomarker discovery, however, establishing validity and maintaining quality control throughout each phase of biomarker discovery and development remains the primary concern.

**RATIONALE OF THE STUDY**

Considering the overwhelming increase in the rate of cervical cancers, particularly in developing countries, approaches addressing all the major aspects of disease pathogenesis are the need of the hour. HPV has been recognized as the major etiological factor for the disease, however, the risk of disease development, progression and prognosis of the infected individuals depend to a great extent on the infecting viral genotype, physical state of the virus (i.e., integrated or episomal) and the viral load. HPV by itself can also bring about several genetic alterations thereby destabilizing the genome. Identifying the association of these viral factors as well as the genetic alterations in the tumour genome, in context of disease development, is important in understanding cervical carcinogenesis.

Pretreatment cervical cancer biopsies as well as blood samples from a cohort of Indian patients with a good follow up were available in the laboratory. This formed an ideal set to address most of the questions related to the pathogenesis of the disease. We undertook a multifaceted approach to study the contribution of several viral cofactors as well as the genetic alterations that are believed to play a role in the genesis of the disease. Through this study we aim towards identification of a set of genetic biomarkers for genesis, progression and prognosis of the disease.
AIMS AND OBJECTIVES

1. Determination of HPV status in cervical cancer biopsies.

2. Identification of integration site of the virus by Amplification of Papillomavirus Oncogene Transcripts (APOT) assay.

3. Identification of genomic alterations in cervical cancer biopsies by exome sequencing

4. Correlation of the data from the study with clinical data.