CHAPTER 1

General Introduction
CERVICAL CANCER

Epidemiology
Cervical cancer is the fourth most common cancer type in women worldwide. The majority of cervical cancers, 85% of the estimated 528,000 new cases and 87% of 266,000 deaths in 2012, occurs in the less developed countries. With these numbers, it is the most common cancer in women in Eastern and Middle Africa. The substantially lower incidence in developed countries results from the introduction of organised cervical screening programmes, which lead to earlier detection and earlier treatment of cervical precancer and cancer. In The Netherlands, around 750 new cases of cervical cancer and 200 deaths are reported each year. The highest incidence is found in women aged 40-44 years.

The uterine cervix and cervical cancer
The cervix consists of the ectocervix, which is covered by multilayered squamous epithelial cells, and the endocervix covered with a single layer of glandular columnar cells. The border between these two cell types is called the squamo-columnar junction (SCJ). During puberty, columnar cells are replaced by squamous epithelial cells (squamous metaplasia) due to hormonal changes, which results in a shift of the SCJ from the ectocervix to the endocervix. The region from the original location of the SCJ and the new SCJ is called the transformation zone (TZ) (Fig. 1). Cervical cancer arises at the TZ following a persistent infection with a high-risk type of the human papillomavirus (hrHPV).

There are different histological subtypes of cervical cancer, with squamous cell carcinoma (SCC) comprising most cases (76%), followed by adenocarcinoma (AdCA; 11%). The remaining small percentage of cervical cancer cases consist of rare histotypes, i.e. neuro-endocrine and clear-cell carcinomas.

HRHPV-INDUCED CERVICAL CARCINOGENESIS

Human Papillomavirus
A persistent infection with hrHPV is the causative agent of nearly all cervical cancers. HRHPV DNA is detected in virtually all SCCs and in the majority (94-100%) of AdCAs. HPVs are member of the Papillomaviridae family and are non-enveloped DNA viruses. Each virus contains double-stranded circular DNA of approximately 7900 base pairs. The viral genome consists of 3 genomic regions: 1. The non-coding long control region (LCR); 2. Up to 7 open reading frames (ORFs) of the early region (E1, E2, E4, E5, E6, E7 and E8) encoding proteins for viral replication and 3. The late region containing 2 ORFs (L1 and L2) encoding viral capsid proteins.

To date, more than 180 types of HPV have been identified based on their DNA sequence and approximately 50 of them target mucosal epithelium. The HPV productive lifecycle is controlled by the differentiation programme of the host cells with a tight crosstalk between viral replication and host differentiation programmes. The virus targets undifferentiated epithelial cells in the basal layer of the epithelium, while viral progeny are produced in terminally differentiated cells in the surface layer of the epithelium. In such productive infections, E6 and E7 contribute to the viral life cycle by activating the cellular DNA replication machinery to allow viral genome amplification in differentiated cells. Only in so-called transforming infections that are associated with cancer development, hrHPV E6 and E7 function as oncoproteins. Their expression is deregulated and very high in proliferating basal cells. Both viral oncoproteins are essential for the development and maintenance of a malignancy by targeting various critical cellular processes in the host cell. This results in genetic and epigenetic host cell changes, leading to genomic instability, activation of oncogenes and inactivation of tumour suppressor genes.

To date, 12 hrHPV types are associated with cervical cancer by the International Agency for Research on Cancer (IARC class 1; HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58 and 59). These 12 high-risk types differ in their oncogenic potential. HPV16 (61%) and HPV18 (10%) cause in total 71% of cervical cancer, while other hrHPV types are less frequently detected. Another 8 HPV types are classified as probably (IARC class 2A; HPV68) or possibly (IARC class 2B, HPV26, 53, 66, 67, 70, 73 and 82) high-risk, due to rare detection of these types in cervical cancer.

Concept of hrHPV-mediated cervical carcinogenesis
Approximately 80% of sexually active women will become infected with hrHPV at some point in their life. Although hrHPV infections are common, the majority (80%) are transient...
infections and will be cleared by the immune system within 1-2 years without notice. The remaining infections may lead to cervical intraepithelial neoplasia (CIN), graded 1 to 3 depending on the width of the epithelium covered by atypical cells\textsuperscript{19,20}. According to the original concept of cervical carcinogenesis (Fig. 2), a small fraction (< 20\%) of the hrHPV-infected metaplastic squamous epithelium cells in the TZ will develop into a low grade CIN lesion (CIN1/2) associated with viral production and also referred to as productive CIN. The majority of these CIN lesions will regress to normal cervical epithelium when the infection is cleared by the immune system. Only a minority of hrHPV infections will be persistent and may convert to a transforming infection, resulting in a transforming CIN lesion. These transforming CIN lesions morphologically correspond to high-grade CIN2/3 lesions (further referred to as “CIN2/3”). CIN2/3 represents a heterogeneous disease, and progression to cancer may take 15 to 30 years. Only a subset of CIN2/3 is suggested to have a high short-term progression risk to cancer, which are also referred to as advanced CIN2/3. Conversely, so-called early CIN2/3 have a low short-term progression risk to cancer and may still regress to normal cervical epithelium\textsuperscript{21}.

In 2012, a new concept was proposed in which the majority of cervical cancers arise from a discrete population of cuboidal epithelial cells from embryonic origin localised in the SCJ of the cervix. These cuboidal cells are supposed to be highly susceptible for transforming hrHPV infections, leading directly to transforming CIN lesions (Fig. 2)\textsuperscript{22}. The single-layered cuboidal cells are characterised by the expression of various supposed SCJ-specific biomarkers, such as keratin 7 and AGR2, but some of them have also been found to be related with the differentiation stage of epithelial cells. These biomarkers were also present in the majority of cervical cancers and CIN2/3 lesions, underlining the role of SCJ cells in cervical carcinogenesis\textsuperscript{23}. According to this new concept, infection with hrHPV in other cells of the TZ results in productive CIN lesions, which may progress to CIN2/3\textsuperscript{21}.

Recently, another concept was described, suggesting that the viral load changes in the beginning of the infection enable classification of CIN lesions in hrHPV-driven transforming infections or in transient productive infections\textsuperscript{24}.

Figure 2. hrHPV-mediated cervical carcinogenesis according to the original and new concept. The majority of hrHPV infections in the transformation zone (TZ) are cleared. A small subset of infections may progress to productive infections leading to productive CIN lesions (CIN1/2), or to transforming infections resulting in transforming CIN lesions (CIN2/3). hrHPV infections in squamo-columnar junction (SCJ) cells may directly result in transforming CIN lesions. Transforming CIN lesions can be divided in early CIN2/3 and advanced CIN2/3, with respectively a low and high short-term progression risk to cancer. From the onset of cervical carcinogenesis, it takes 15 to 30 years for development of cervical cancer and additional genetic and epigenetic host cell alterations are necessary. Adapted from Steenbergen et al.\textsuperscript{21}.
MOLECULAR HOST CELL ALTERATIONS

From the onset of a transforming hrHPV infection, it takes another 15 to 30 years before cervical cancer may develop. During this long time window, additional genetic and epigenetic host cell alterations occur (Fig. 2; see Human Papillomavirus). An accumulation of these molecular changes is necessary for progression to cervical cancer. In concordance with this, the rates and amount of molecular aberrations increase with the severity of CIN2/3 lesions. Genetic host cell alterations detected in cervical cancer and CIN2/3 lesions include copy number aberrations (CNA) and DNA mutations. Epigenetic changes include altered microRNA (miRNA) expression and DNA methylation.

Copy number aberrations

In a meta-analysis published by Thomas et al., the most frequently observed altered genomic region in cervical SCC is gain of 3q (55%), followed by loss of 3p (36%), loss of 11q (33%) and gain of 1q (29%). In CIN2/3 lesions, gain of 3q (27%), loss of 3p (8%), loss of 11q (15%) and gain of 1q (19%) are already detected, however at lower rates. These specific CNA are less frequently detected in AdCa. Alternatively, a gain of 17q (36%) is most common in AdCa.

DNA mutations

The presence of DNA mutations in cervical cancer has been analysed in a comprehensive manner in the last few years. Combining these results, the most frequently mutated gene is PIK3CA with frequencies up to 31%, followed by other mutated genes such as FBXW7, EP300 and KRAS. Most mutations observed in PIK3CA are activating mutations in the helical-domain: known hotspot mutations p.E542K and p.E545K. SCC and AdCa display different mutation profiles. Both histotypes show mutations in PIK3CA, whereas KRAS mutations are mainly detected in AdCa and rarely SCC. No comprehensive mutation studies have been performed in CIN2/3.

miRNA expression

miRNAs are small non-coding RNA molecules of 18-25 nucleotides and are expressed in a tissue-specific pattern. They function as posttranscriptional regulators of gene expression. The seed sequence of 6-8 nucleotides in the miRNA can bind the 3’ UTR region of target mRNAs, resulting in repression or breakdown of the transcript. Interestingly, since partial complementation is sufficient, many different interactions can occur: a single miRNA can target multiple mRNAs, and a single mRNA can also be targeted by multiple miRNAs. Therefore, miRNAs play a complex and important role in cellular processes.

In genome-wide studies, many dysregulated miRNAs have been associated with cervical cancer development. Both upregulated and downregulated miRNAs have been described. Few upregulated miRNAs are associated with copy number gains and several downregulated miRNAs are associated with DNA methylation-mediated silencing. A recent meta-analysis by He et al. identified 42 upregulated and 21 downregulated miRNAs at different stages of cervical carcinogenesis. A subset, 5 upregulated (miR-10a-5p, miR-16-5p, miR-25-5p, miR-92a-3p and miR-196a-5p) and 7 downregulated (miR-29a, miR-34a, miR-99a-5p, miR-100-5p, miR-199a-3p, miR-203 and miR-218-5p) miRNAs, showed a continuous up- or downregulated miRNA profile in CIN1, CIN2, CIN3 and cervical cancer.

DNA methylation

DNA methylation is a well-studied epigenetic event during cervical carcinogenesis. In general, loss of DNA methylation occurs in the genome during carcinogenesis, resulting in chromosomal instability. However, at specific genomic regions such as promoter regions of tumour suppressor genes (TSGs), increased DNA methylation levels are observed in cancer cells. DNA methyl transferases (DNMTs) add a methyl-group to cytosines at CpG dinucleotides. Genomic regions with a high density of these CpG sites are called CpG islands and are frequently found in promoter regions. Methylated promoter regions generally result in repression of gene expression, which can play an important role in cervical carcinogenesis when TSGs are affected (Fig. 3B). Besides protein-coding genes, other genomic features such as non-coding miRNAs can also be a target of DNA methylation-mediated silencing. Both hrHPV E6 and E7 oncoproteins can influence the expression and activity of DNMT1. E7 can directly bind to and activate DNMT1, as well as indirectly through binding to PRB resulting in the release of EZF. EZF regulates the promoter activity of DNMT1, which results in overexpression of DNMT1. E6 can upregulate DNMT1 via degradation of p53. Conversely, silencing of E6 and E7 resulted in reduction of DNA methylation of TSGs and the transformed phenotype was restored in cervical cancer cells. Furthermore, altered DNA methylation levels were independent of hrHPV type.

The number of studies on altered DNA methylation patterns in CIN lesions and cervical cancer is rising rapidly. Methylated genes identified by both genome-wide or targeted gene approaches include CADM1, EPB41L3, FAM19A4, MAL, mir-124-2 and PAX1.
DNA methylation analysis

Genome-wide methods

DNA methylation can be studied in a comprehensive manner using genome-wide DNA methylation profiling techniques. Both array-based and next generation sequencing (NGS)-based techniques are frequently used44–47. The most common used method is the Infinium HumanMethylation450 BeadChip array (Infinium 450K array)44. Other techniques, such as methyl binding domain (MBD) protein-enriched DNA sequencing (MBD-seq), reduced representation bisulphite sequencing (RRBS), MeDIP-seq or MeDIP-chip and whole genome bisulphite sequencing (WGBS) are also widely used44–47.

In this thesis, we applied two different genome-wide DNA methylation profiling techniques, MBD-seq and the Infinium 450K array. Using MBD-seq, DNA is fragmented and enriched for methylated DNA by MBD proteins. Subsequently, these MBD-enriched DNA fragments are analysed using NGS (Fig. 4)46,47. The Infinium 450K array does not depend on DNA enrichment, but involves bisulphite treatment and amplification of the DNA followed by hybridisation to the array. During bisulphite treatment, unmethylated cytosines are deaminated into uracils, which become thymines after amplification. The methylated cytosines are unaffected and remain as cytosines. The array contains more than 450,000 probes which measure DNA methylation at single CpG sites, covering 99% of RefSeq genes (including miRNAs) and 96% of CpG islands with multiple probes48. Two different probe types are used on the array to measure DNA methylation by single base extension with fluorescently-labelled nucleotides. Single base extension occurs either next to the CpG site depending on the methylation status or within the CpG site using two different colours (Fig. 5)48. These genome-wide techniques are useful for discovery studies, to determine DNA methylation profiles and to identify promising DNA methylation targets.

Multiplex quantitative methylation-specific PCR

A targeted and more sensitive method is useful for validation studies and as an assay for routine screening or diagnostic purposes. Targeted detection of multiple methylated genes together with a reference gene in one assay can be performed by multiplex quantitative methylation-specific PCR (qMSP). Multiplex qMSP is highly reproducible and requires minimal amount of bisulphite-modified DNA as input, which makes it an efficient method and suitable for high-throughput settings49.
General Introduction

Chapter 1

Figure 5. The two different probe types on the Infinium HumanMethylation450 BeadChip array to measure DNA methylation at a single CpG site. (A) Two bead types are present on the array, one specific for the methylated state and one specific for the unmethylated state of the corresponding CpG site. Depending on the methylation status of the CpG site, single base extension occurs at one bead type and gives a signal. (B) One bead type corresponding to a CpG site. Methylation status is measured by single base extension within the CpG locus using two colours, Cy5-red (labelled “A” complementary to unmethylated “T”) and Cy3-green (labelled “G” complementary to methylated “C”). Adapted from Bibikova et al.

CERVICAL SCREENING

HPV-based cervical screening in The Netherlands

The large time window from a hrHPV infection to the development of cervical cancer allows for successful screening. When screening detects CIN2/3 or cancer at an early stage, treatment has a high potential for cancer prevention and cure, respectively.

The Netherlands has a well-organised screening programme which has recently, in 2017, been changed from cytology-based screening to primary hrHPV testing. Multiple studies have shown that primary hrHPV testing is more sensitive in detection of CIN2/3 and cervical cancer compared with primary cytology testing. Secondly, the new screening programme offers self-collection of cervico-vaginal specimens (self-sampling) for hrHPV testing to increase the participation rate and thereby effectiveness of the programme. Previous studies have shown that up to 30% of original non-attendees return a self-sample to the laboratory for hrHPV testing. Notably, hrHPV testing on self-samples for detection of CIN3 and cervical cancer has a similar clinical performance and accuracy as hrHPV testing on physician-collected cervical scrapes, when the right combination of self-sampling device and hrHPV test is used.

Primary hrHPV testing using a clinically validated assay has a high sensitivity for CIN2/3 and cervical cancer, however, its specificity is 3-5% lower compared with cytology, due to detection of transient hrHPV infections. Therefore, a triage test is necessary to identify only women with clinically relevant infections associated with cervical lesions that require further referral to the gynaecologist. This prevents over-referral and overtreatment. Cytology either alone or combined with HPV16/18 genotyping are the most frequently used triage methods. Currently, in The Netherlands, repeat cytology is the used triage tool comprising cytology at baseline with reflex cytology at 6 months in women negative for intraepithelial lesion or malignancy (NILM) at baseline. Cytology can be performed on the same physician-taken cervical scrape material as used for primary hrHPV testing. However, cytology cannot be reliably performed on self-sampled material. Consequently, women with a hrHPV-positive self-sample need to visit a physician for an additional cervical scrape for cytology triage testing. This may lead to loss to follow-up of 10-40% of women and delays in the diagnostic track.

Alternative triage methods

To improve the management of hrHPV-positive women, alternative triage methods applicable on both cervical scrapes and self-sampled material are necessary. From the methods which have been described, P16/Ki-67 co-expression (CINtec® PLUS; Roche) and
FAM19A4/miR124-2 DNA methylation (QIASure®; Qiagen) are the most promising assays and have been clinically validated.

In contrast to microscopy-based triage methods such as CINtec, the applicability of molecular tests on self-samples in addition to physician-taken scrapes makes this a very attractive triage method. Molecular markers based on DNA methylation of host cell genes have been extensively analysed with feasible and effective results for detection of CIN2/3 lesions and cervical cancer in both cervical scrapes and self-samples\textsuperscript{21,64,65}. Other genetic or epigenetic host cell alterations might also provide interesting molecular triage markers.

**THESIS OUTLINE**

Besides a persistent hrHPV infection, additional genetic and epigenetic alterations in host cell genes are necessary for the progression of a CIN2/3 to cervical cancer. Improved knowledge on these host cell changes may enable a better risk stratification of hrHPV-positive women for the development of cervical cancer, which is the topic of this thesis. As detailed below, we evaluated DNA mutations, DNA methylation, miRNA expression and DNA methylation-mediated silencing of miRNAs with respect to their association with hrHPV-induced carcinogenesis. Furthermore, we translated this knowledge into the development of molecular triage tools for hrHPV-positive physician-taken and self-collected samples. To achieve these goals, the following studies were performed.

In **Chapter 2**, we analysed six previously identified downregulated miRNAs, which are located in a CpG island. We determined whether these miRNAs were targeted by DNA methylation-mediated silencing during cervical carcinogenesis. Three out of six miRNA’s (miR149, miR203 and miR375) showed repression due to DNA methylation. DNA methylation analysis of miR203 showed an increase in DNA methylation levels associated with severity of CIN lesions to cervical cancer in tissue specimens as well as in a small series of cervical scrapes. Furthermore, we showed that miR203 has a tumour suppressive role.

DNA mutations are comprehensively analysed in cervical cancer, however, the mutation profile in CIN2/3 lesions is largely unknown. Therefore, in **Chapter 3**, we tested CIN2/3 and cervical cancer for the presence of somatic hotspot mutations in cancer-related genes by targeted NGS. We found PIK3CA exon 9 as the most frequently mutated gene locus in cervical cancer and as the only detected mutated locus in CIN2/3. Evaluation of PIK3CA exon 9 in independent large series of cervical tissue specimens showed that this gene locus was mutated in a substantial subset of cervical cancers, while only rarely in CIN3 and not in lower-grade CIN lesions, suggesting that somatic mutations in PIK3CA represent a late event in cervical carcinogenesis.

Previously, it has been shown that altered methylated genes can function as promising triage markers to identify hrHPV-positive women at risk for progression to cervical cancer\textsuperscript{21}. However, the clinical performance can still be increased and identification of novel DNA methylation markers remains necessary to improve the management of hrHPV-positive women. In **Chapter 4**, we performed an unbiased genome-wide DNA methylation screen followed by comprehensive verification and validation steps using in vitro and patient-
derived samples. We identified 3 novel biologically relevant DNA methylation markers, GHSR, SST and ZIC1, which showed a good clinical performance for detection of CIN3 and cervical cancer in hrHPV-positive cervical scrapes. Interestingly, within the heterogeneous CIN2/3 lesions, all 3 genes mainly detect the clinically relevant advanced CIN2/3, defined as CIN2/3 associated with a persistent hrHPV infection of ≥ 5 years, and are associated with a 3q gain.

Since self-sampling is offered in the Dutch HPV-based screening programme, it is important to identify DNA methylation markers which are also applicable on self-sampled material with a good clinical performance for detection of CIN2/3 and cervical cancer. From previous research we know that DNA methylation markers originally discovered in tissue specimens and evaluated in cervical scrapes are not necessarily of good clinical value in self-samples, due to the dilution of disease-related cells by vaginal cells. In Chapter 5, we therefore performed a second genome-wide DNA methylation marker discovery screen directly on self-sampled material to identify the most informative DNA methylation markers for hrHPV-positive self-samples. This screen revealed 12 DNA methylation targets for detection of CIN3 and were subsequently analysed by multiplex qMSP in independent large series of hrHPV-positive lavage and brush self-samples to build a DNA methylation classifier applicable to self-samples of both devices. Using logistic regression analysis, we identified a highly effective 3-gene methylation classifier, consisting of ASCL1, LHX8 and ST6GALNAC5. This 3-gene methylation classifier showed in the independent large hrHPV-positive validation set a very good and reproducible clinical performance in both lavage and brush self-samples for CIN3 detection. Importantly, all self-samples from women with cervical cancer scored DNA methylation-positive by the 3-gene methylation classifier.

In Chapter 6, we systematically analysed 12 genes identified in multiple earlier studies to compare the onset and DNA methylation pattern during cervical carcinogenesis. The 12 genes showed an increased DNA methylation profile and varying onset of DNA methylation in the premalignant stage in an in vitro model. In cervical scrapes, the majority of methylated genes showed a highly comparable and increased DNA methylation profile associated with the increased severity of CIN lesions to cervical cancer. Furthermore, a heterogeneous DNA methylation profile was observed in scrapes from women with CIN2 and CIN3, with a subset showing a cancer-like methylation-high profile.

Besides DNA methylation, deregulated miRNAs could also serve as a potential molecular triage tool for hrHPV-positive women. Therefore, in Chapter 7, we determined miRNA profiles by genome-wide small RNA sequencing on hrHPV-positive self-sampled material to identify the most promising miRNAs that can predict the presence of CIN3 and cervical cancer in self-samples. This screen yielded a miRNA panel of 9 miRNAs for CIN3 detection. Further validation by qPCR resulted in a miRNA classifier of 5 miRNAs with a good clinical performance for detection of CIN3 and cervical cancer.

In Chapter 8, we present a summary and discuss our findings with future applications.
REFERENCES


