

# Genome-wide methylome analysis using MethylCap-seq uncovers 4 hypermethylated markers with high sensitivity for both adeno- and squamous-cell cervical carcinoma

Rong Wang<sup>1,3,\*</sup>, Robert W. van Leeuwen<sup>1,\*</sup>, Aniek Boers<sup>1</sup>, Harry G. Klip<sup>1</sup>, Tim de Meyer<sup>4</sup>, Renske D. M. Steenbergen<sup>5</sup>, Wim van Criekinge<sup>4</sup>, Ate G. J. van der Zee<sup>1</sup>, Ed Schuurin<sup>2</sup>, G. Bea A. Wisman<sup>1</sup>

<sup>1</sup>Department of Gynecologic Oncology, University of Groningen, University Medical Centre Groningen, Cancer Research Centre Groningen, Groningen, The Netherlands

<sup>2</sup>Department of Pathology, University of Groningen, University Medical Centre Groningen, Cancer Research Centre Groningen, Groningen, The Netherlands

<sup>3</sup>Department of Laboratory Medicine, Tianjin Medical University, Tianjin, China

<sup>4</sup>Department of Mathematical Modeling, Statistics and Bio-informatics, University of Ghent, Ghent, Belgium

<sup>5</sup>Department of Pathology, VU University Medical Centre, Amsterdam, The Netherlands

\*These authors have contributed equally to this work

**Correspondence to:** Ed Schuurin, **email:** e.schuuring@umcg.nl  
G. Bea A. Wisman, **email:** g.b.a.wisman@umcg.nl

**Keywords:** uterine cervical neoplasms, DNA methylation biomarkers, MethylCap-seq, adenocarcinoma (in situ), (quantitative) methylation-specific PCR ((Q)MSP)

**Received:** July 14, 2016

**Accepted:** September 29, 2016

**Published:** October 12, 2016

## ABSTRACT

**Background:** Cytology-based screening methods for cervical adenocarcinoma (ADC) and to a lesser extent squamous-cell carcinoma (SCC) suffer from low sensitivity. DNA hypermethylation analysis in cervical scrapings may improve detection of SCC, but few methylation markers have been described for ADC. We aimed to identify novel methylation markers for the early detection of both ADC and SCC.

**Results:** Genome-wide methylation profiling for 20 normal cervixes, 6 ADC and 6 SCC using MethylCap-seq yielded 53 candidate regions hypermethylated in both ADC and SCC. Verification and independent validation of the 15 most significant regions revealed 5 markers with differential methylation between 17 normals and 13 cancers. Quantitative methylation-specific PCR on cervical cancer scrapings resulted in detection rates ranging between 80% and 92% while between 94% and 99% of control scrapings tested negative. Four markers (*SLC6A5*, *SOX1*, *SOX14* and *TBX20*) detected ADC and SCC with similar sensitivity. In scrapings from women referred with an abnormal smear (n=229), CIN3+ sensitivity was between 36% and 71%, while between 71% and 93% of adenocarcinoma in situ (AdCIS) were detected; and CIN0/1 specificity was between 88% and 98%. Compared to hrHPV, the combination *SOX1/SOX14* showed a similar CIN3+ sensitivity (80% vs. 75%, respectively, P>0.2), while specificity improved (42% vs. 84%, respectively, P < 10<sup>-5</sup>).

**Conclusion:** *SOX1* and *SOX14* are methylation biomarkers applicable for screening of all cervical cancer types.

## INTRODUCTION

Cervical cancer is one of the most common female cancers in the world, with more than 525,000 new cases and over 265,000 deaths occurring globally each year [1, 2]. Cervical squamous-cell carcinoma (SCC) and

cervical adenocarcinoma (ADC) are two main histological subtypes of invasive cervical cancer, which account for 75 - 90% and 10 - 25% of cases, respectively [3-5]. Currently, the incidence of SCC is declining in most developed countries. In contrast, there is a rise in the absolute and relative incidence of ADC [6]. In Europe,

ADC is increasing rapidly, especially in younger women [7, 8]. In the Netherlands, the absolute incidence rate of ADC increased by 15.8% in women aged 15 - 29 years and 2.5% in women aged 30 - 44 years [7]. Moreover, compared to SCC, ADC is mainly diagnosed in more advanced stages, appears to be less sensitive to (chemo) radiation therapy and is associated with a worse prognosis [9–12].

Both the upward trend in relative incidence as well as delayed detection of ADC are probably due to the present population-based screening programs, which are more effective in the detection of the precursors of SCC than those of ADC. This may be because ADC arises in a more cranial localization in the cervix where it is more difficult to either obtain representative cytology samples or to detect ADC or its precursors by colposcopy [13]. High-risk human papillomavirus (hrHPV) is widely accepted as the predominant etiological agent of cervical cancer [14] and its detection is clinically more sensitive for the detection of cervical adenocarcinoma in situ (AdCIS) and ADC compared to cytology [15]. As hrHPV testing has a relatively low positive predictive value in population-based screening programs [16, 17], women who are tested positive will require an additional test that is equally sensitive to detect SCC and ADC and can also sufficiently detect their precursors to ensure correct referral to the gynecologist for colposcopy [18]. Therefore, novel biomarkers for cervical cancer are required that ideally will identify both ADC and SCC as well as their precursors with high sensitivity.

Aberrant gene expression caused by epigenetic mechanisms is a prominent feature of many types of cancer [19], and DNA promoter methylation of tumor suppressor genes (TSG) has been reported to be an early event in carcinogenesis [20]. DNA methylation markers might be exploited in cancer diagnosis as variations in DNA methylation are observed more frequently than other genetic variations [21]. Although we [22, 23] and others [24] have reported many methylation markers associated with cervical cancer, many of these markers are more frequently methylated in SCC compared to ADC [23]. Moreover, so far only a limited number of methylated genes have been identified that are specifically associated with ADC. Most of these markers have lower sensitivity for both ADC and SCC or either one [25–27].

In the past ten years, advances in whole-genome methylation profiling technologies have revolutionized the field of cancer research. In order to identify cervical cancer-specific methylation markers, approaches such as a pharmacological unmasking expression microarray [28] or immunoprecipitation combined with oligonucleotide microarrays have been performed [29–31]. Nevertheless, microarray-based screening has drawbacks regarding their design and production, and also the inaccurate hybridization signals and variable immunoprecipitation step leave room for further improvement. Reductions

in costs have spurred the adoption of next-generation sequencing (NGS) platforms with higher sensitivity and accuracy compared to traditional microarray profiling [32]. Recently, an affinity-based methylation capture assay using methyl-binding domain (MBD) complexes coupled with NGS (MethylCap-seq) has been reported to be an effective technique to comprehensively analyze the methylome in lung cancer, ovarian cancer, head and neck cancer, and high-grade cervical intraepithelial neoplasia (CIN) [33–37]. These technologies have facilitated the discovery of potential DNA methylation biomarkers for disease development and progression as well as our understanding of the complex, underlying molecular mechanisms that lead to cancer.

Until now, no cervical cancer DNA methylome analysis has been performed using ADC. In this study, MethylCap-seq was applied to perform a genome-wide DNA methylation screening of cervical cancer, including both ADC and SCC, and normal cervix tissues. With this approach, we sought to identify genome-wide aberrant methylation patterns of cervical cancer-specific markers with high sensitivity to detect both ADC and SCC and their precursors in cervical scrapings.

## RESULTS

### Identification of methylated candidates by MethylCap-seq

To discover markers that are methylated in both SCC and ADC, but not in normal cervix, we generated a methylome of 6 SCC and 6 ADC and used a stepwise selection approach as outlined in Figure 1. In total, 6,231 candidate differentially methylated regions (DMRs) for ADC compared to normal cervixes and 10,724 candidate DMRs for SCC compared to normal cervixes were identified after applying our selection criteria (see methods section for detailed description of the identification step, Figure 2). In ADC as well as in SCC hypomethylation was more frequently observed compared to hypermethylation (Figure 3). We focused on the hypermethylated DMRs, as these are more easily translated into methylation-specific PCR (MSP) assays, which can be implemented as clinical diagnostic tests. Overall 446 candidate hypermethylated DMRs, comprising 357 genes, were identified in ADC and 93 DMRs, comprising 89 genes, in SCC. Gene ontology (GO) functional analysis for these DMRs was performed to determine if similar pathways were affected in both cancer histological types. There were in total 328 and 49 GO terms enriched in ADC and SCC, respectively (Supplementary Table S1 and S2). Most GO terms enriched in SCC were also enriched in ADC, as 37/38 of the biological processes, 4/5 of the cellular components and 5/6 of the molecular functions, were also shown in ADC. This observation suggests that similar pathways are disrupted in the carcinogenesis of both histological types.

Figure 4 shows the most significant GO terms enriched in ADC together with the associated significance in SCC. We identified 53 regions, comprising 50 genes that were hypermethylated in both ADC and SCC (Figure 2).

### Verification and validation of the top 15 candidates

After identification of 53 candidate regions, the top 15 regions (Supplementary Table S3) were selected for verification of the MethylCap-seq data by either MSP or pyrosequencing analysis. To obtain the top 15 genes, candidate regions were ranked on the number of (i) methylated ADC, (ii) methylated SCC, and on (iii) unmethylated normal cervixes. Using the same DNA as used for MethylCap-seq, 10 genes (*SOX1*, *GFRA1*, *SLC6A5*, *TBX5*, *OLIG2*, *AC004963.1*, *TBX20*, *RP11-100E13.1*, *RP1-241P17.1*, and *SOX14*) showed a significant correlation between MSP or pyrosequencing and the number of reads from the MethylCap-seq (Supplementary Table S4).

MSP primers were designed for these 10 markers. Four MSP assays showed high methylation levels in DNA from leukocytes and whole-genome amplified (WGA) DNA and were therefore excluded from further validation (Supplementary Table S4). MSP of the remaining 6 genes was performed on DNA from an independent series of 17 normal cervix and 13 cervical cancer formalin-fixed paraffin-embedded (FFPE) tissue samples (6 ADC and 7 SCC). Except for *TBX5*, all 5 genes (*GFRA1*, *SLC6A5*, *SOX1*, *SOX14* and *TBX20*) were significantly differentially methylated between normal and cancer tissues with a high methylation frequency in both ADC and SCC (Table 1).

### Diagnostic evaluation on scrapings from healthy cervixes and cervical cancer scrapings

QMSP was designed for 5 genes (*GFRA1*, *SLC6A5*, *SOX1*, *SOX14*, *TBX20*) and their diagnostic potential was evaluated on scrapings from a large series of cervical cancer patients (n=125: 57 ADC and 68 SCC) and controls with similar age. The level of DNA methylation for all five genes

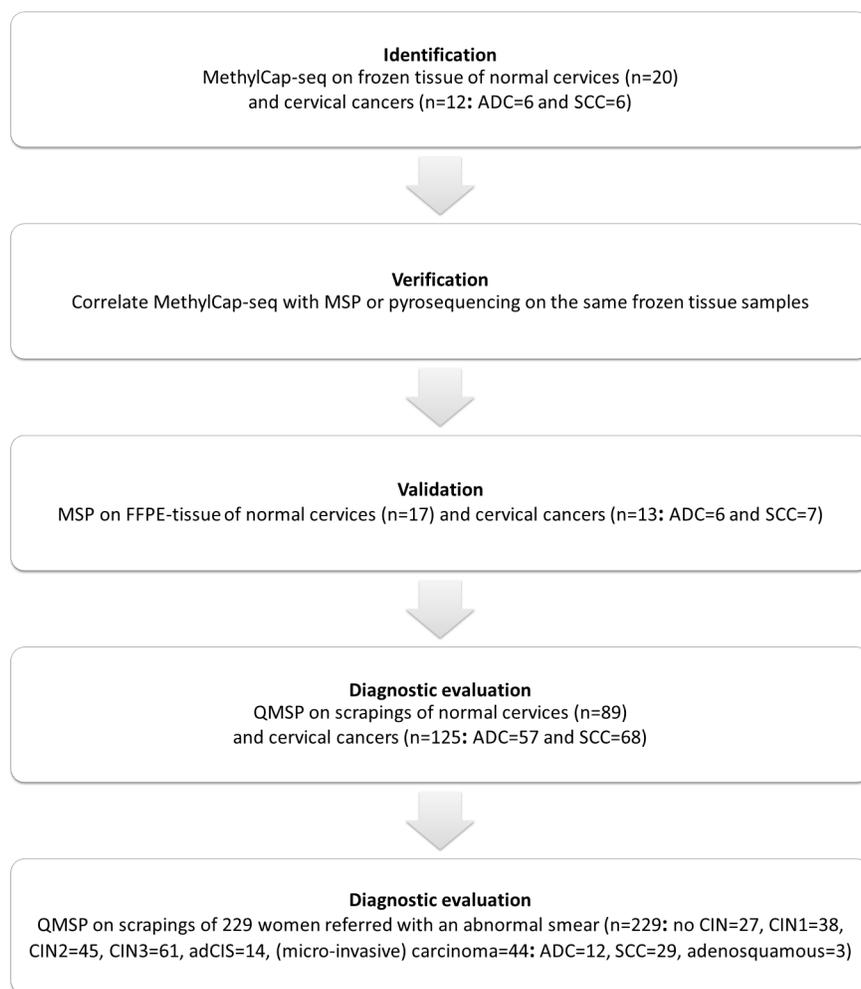
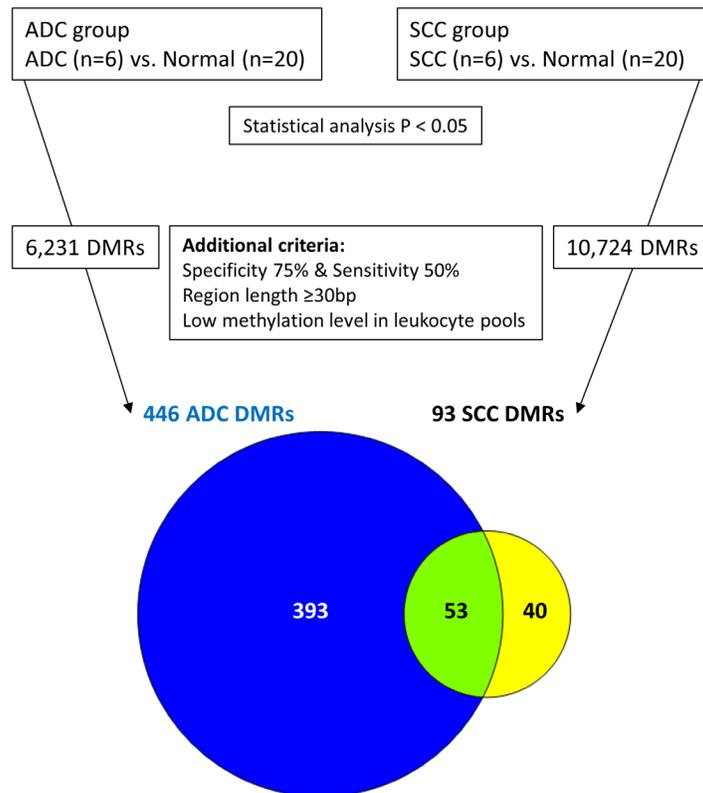
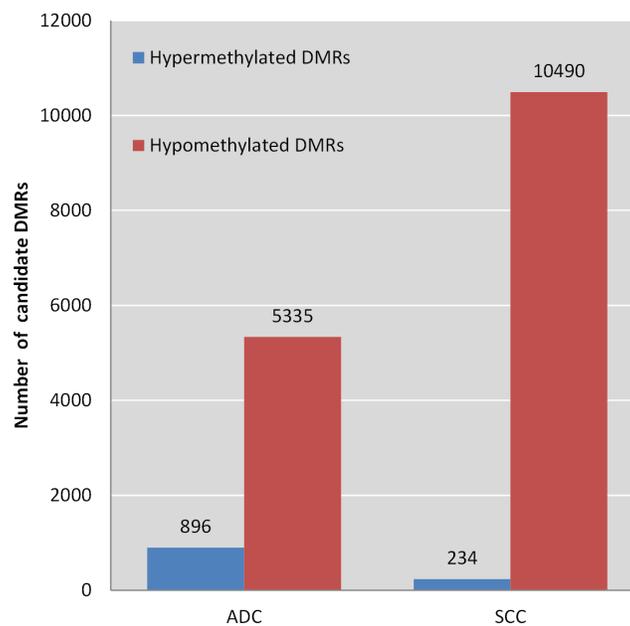


Figure 1: Flow scheme for the identification of new cervical cancer markers.



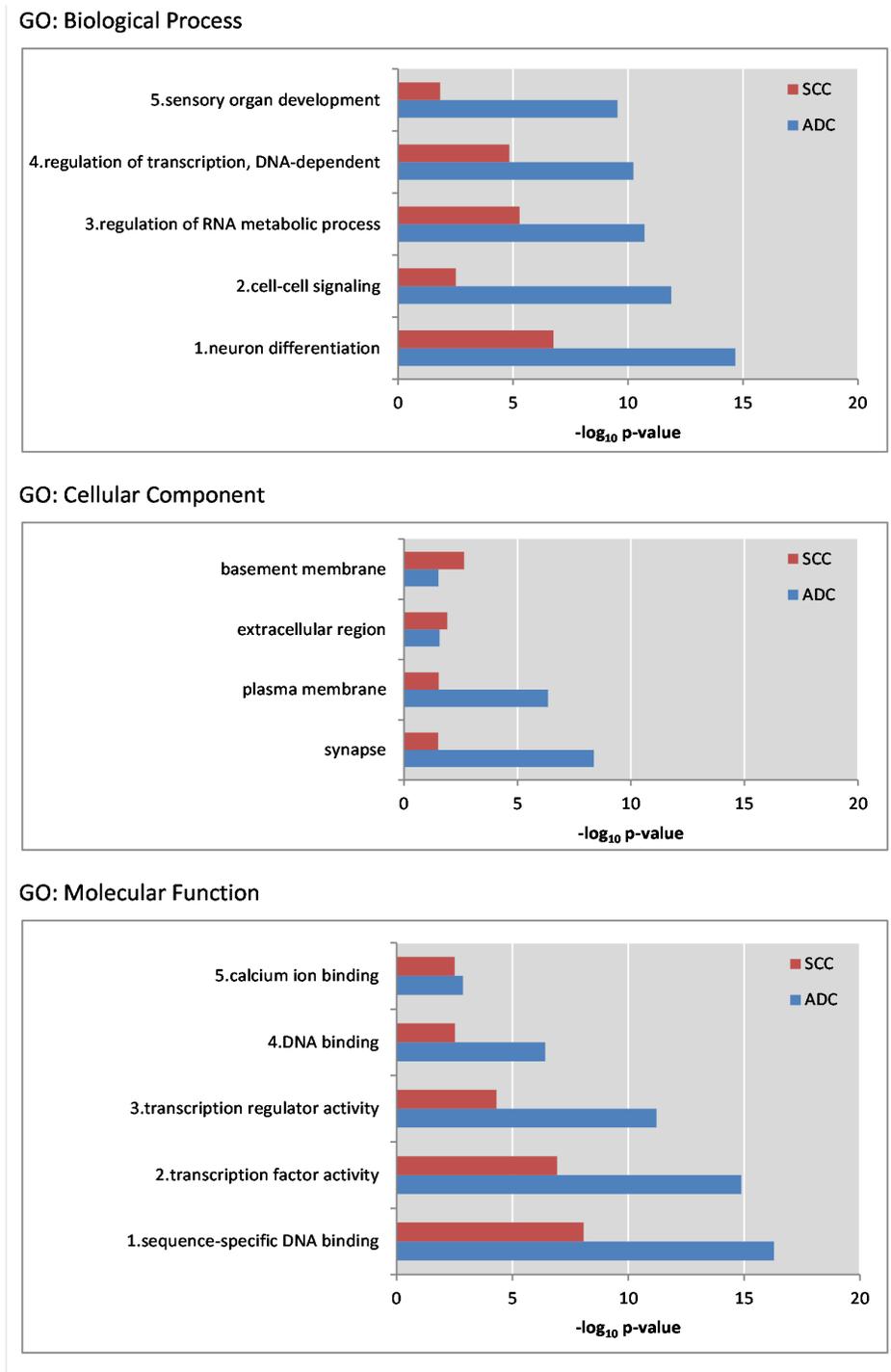
**Figure 2: Identification of methylated candidates by MethylCap-seq.**



**Figure 3: Frequencies of hyper- and hypomethylated regions in ADC and SCC.**

was significantly different in cancer scrapings compared to normal scrapings (each  $P < 10^{-24}$ ), while methylation levels in ADC and SCC (Figure 5) were similar (all  $P > 0.10$ ). Because many markers were also methylated in normal scrapings, albeit at lower levels as observed in cancer scrapings, a threshold was set by maximizing Youden's index J using receiver-operator characteristic (ROC)

analysis of the individual genes. Hereafter, the specificity and the sensitivity for ADC and SCC were determined for all individual genes (Table 2). The sensitivity of the 5 QMSP assays ranged from 79% to 88%, while the specificity ranged from 94% to 99%. Except for *GFR1*, all markers (*SLC6A5*, *SOX1*, *SOX14* and *TBX20*) detected ADC and SCC with a similar sensitivity.

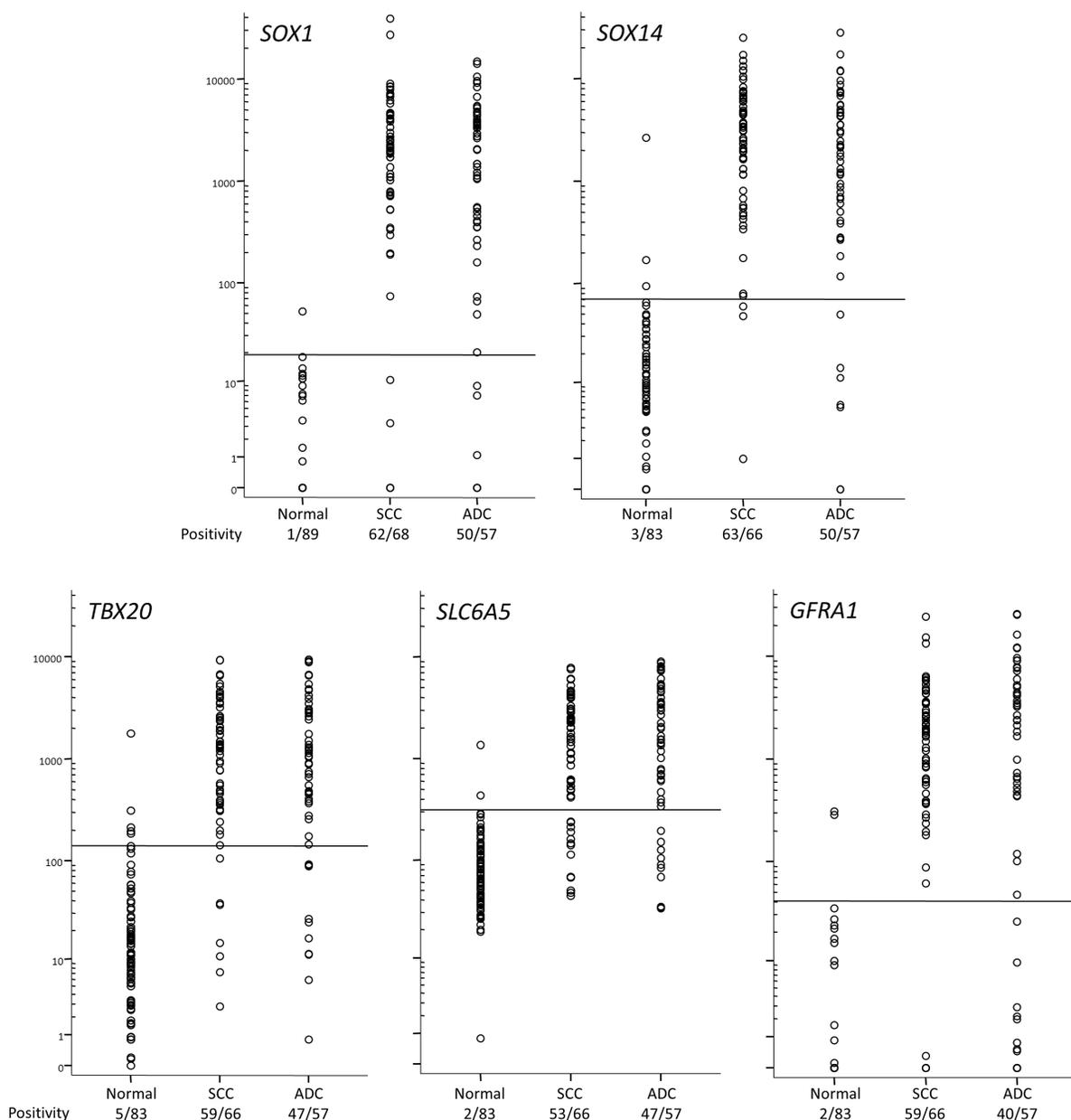


**Figure 4:** For three GO themes the five most significant GO terms enriched in ADC are shown along with the P-value of that term in SCC. Within the cellular component GO terms, only 4 terms were significant in both ADC and SCC. All depicted P-values are below 0.05.

**Table 1: MSP positivity in the external validation cohort of FFPE tissue samples**

Gene	Normal		Cancer		ADC		SCC	
<i>TBX5</i>	56%	(9/16)	67%	(8/12)	60%	(3/5)	71%	(5/7)
<i>SOX14</i> *	25%	(4/16)	85%	(11/13)	83%	(5/6)	86%	(6/7)
<i>SOX1</i> *	0%	(0/15)	92%	(11/12)	100%	(5/5)	86%	(6/7)
<i>TBX20</i> *	6%	(1/17)	83%	(10/12)	100%	(5/5)	71%	(5/7)
<i>SLC6A5</i> *	7%	(1/15)	83%	(10/12)	100%	(5/5)	71%	(5/7)
<i>GFRA1</i> *	0%	(0/11)	83%	(10/12)	83%	(5/6)	83%	(5/6)

\* the positive rates in normal and cancer samples differ ( $P < 0.05$ ).



**Figure 5: DNA methylation levels in normal and cancer scrapings determined by QMSP.** The horizontal lines represent optimal thresholds (see Table 2). The positive rate is depicted below the class labels.

**Table 2: Diagnostic performance of five QMSPs on normal and cancer scrapings**

Gene	AUC (95% CI)	J	Cutoff	Positive normal	Positive cancer	P†	Positive ADC	Positive SCC	P‡
<i>SOX1</i>	0.96 (0.93–0.99)	0.885	19	1%	90%	10 <sup>-37</sup>	88%	91%	0.528
<i>SOX14</i>	0.96 (0.93–0.99)	0.883	70	4%	92%	10 <sup>-36</sup>	88%	95%	0.118
<i>TBX20</i>	0.94 (0.90–0.97)	0.802	140	6%	86%	10 <sup>-29</sup>	82%	89%	0.266
<i>SLC6A5</i>	0.93 (0.89–0.96)	0.789	315	2%	81%	10 <sup>-28</sup>	82%	80%	0.760
<i>GFR1</i>	0.92 (0.88–0.96)	0.781	41	2%	80%	10 <sup>-28</sup>	70%	89%	0.007

CI is the confidence interval, † frequency normal vs. cancer, ‡ frequency ADC vs. SCC.

### Diagnostic evaluation on scrapings of women referred with an abnormal smear

In the second diagnostic evaluation 229 scrapings were analyzed for *SLC6A5*, *SOX1*, *SOX14* and *TBX20* hypermethylation using the QMSP threshold values that were set in the normal and cancer samples (Table 2). For each gene both the QMSP ratios and the positive rates increased with the severity of the underlying lesion ( $P < 10^{-9}$  for each gene, Figure 6). Also in this cohort no differences in QMSP levels and positivity between ADC and SCC were found (data not shown). Of note, 10 to 13 out of 14 AdCIS scrapings were methylation positive (Figure 6).

Analysis of the sensitivity and specificity of individual genes indicated that *SOX1* was the most sensitive, whereas *SLC6A5* and *TBX20* were the most specific (Table 3, Supplementary Table S5). Next, all possible gene combinations were generated in order to improve the diagnostic performance. The combination with the highest sensitivity and specificity was *SOX1/SOX14/TBX20* methylation, in which *SLC6A5* was not additive. The best 2 gene combination was *SOX1/SOX14*.

Analysis of hrHPV in the same scrapings revealed that hrHPV detection was more frequent when the underlying lesion was more severe ( $P = 0.005$ , see Table 3, Supplementary Table S5). Also, hrHPV was detected at similar rates in both ADC and SCC (75% vs. 68% respectively,  $P > 0.7$ , data not shown).

The hrHPV test classified samples differently than *SOX1/SOX14* hypermethylation ( $P < 10^{-4}$ ). In this population of women referred with abnormal cytology, methylation analysis was less sensitive to detect CIN2+ ( $P < 0.001$ ), however, equally sensitive for CIN3+ ( $P > 0.2$ ), and produced less false-positives compared to hrHPV analysis ( $P < 10^{-5}$ ).

## DISCUSSION

In our study, we used a genome-wide DNA methylation screening strategy detecting ADC as well as SCC with its precursor lesion in cervical scrapings. In addition, many differentially methylated regions were

observed when normal cervixes were compared with both ADC and SCC. The observation that some regions were differentially methylated exclusively in ADC or SCC could reflect exposure to different environmental factors [38]. Specifically, this observation may be explained by the difference in risk factors – smoking and high parity are risk factors for SCC [39] and obesity is a risk factor for ADC [40].

GO analysis pointed out that most of the pathways affected by hypermethylation in SCC were also affected in ADC, indicating similar pathways are deregulated by hypermethylation during carcinogenesis independent of histological cancer subtype. Pathways identified were all known to be involved in carcinogenesis [41–43]. Of the 53 differentially methylated candidates that were found in both ADC and SCC, 20 genes (Supplementary Table S3) were described previously in literature as being more frequently methylated in cancer, and 6 genes in (squamous-cell) cervical cancer (*SOX1*, *SOX14*, *ONECUT1* and *WT1*) [44] or high-grade CIN (*GFR1*, *SOX1* and *TBX20*) [34, 45].

Compared to the gene panels recently reported by our group [34] the combination of *SOX1/SOX14* methylation showed less positive test results, both in scrapings from women without cervical disease and in women with high-grade CIN, but not in women with cervical cancer. Further research (e.g. by decision tree algorithms or latent structure analysis) is necessary to evaluate whether these markers are additive to each other. Furthermore, we cannot exclude the possibility that other interesting candidate genes are present in the highest ranking genes beyond the top 15 regions (Supplementary Table S3). Alternatively unmethylated CIN2+ samples may be analyzed by genome-wide methods to develop complementary assays that increase the clinical sensitivity. In addition, it can be hypothesized that the positive CIN2/3 samples are more similar to the cervical cancer lesions compared to the negative CIN samples; possibly reflecting the percentage of women which might develop cancer when left untreated.

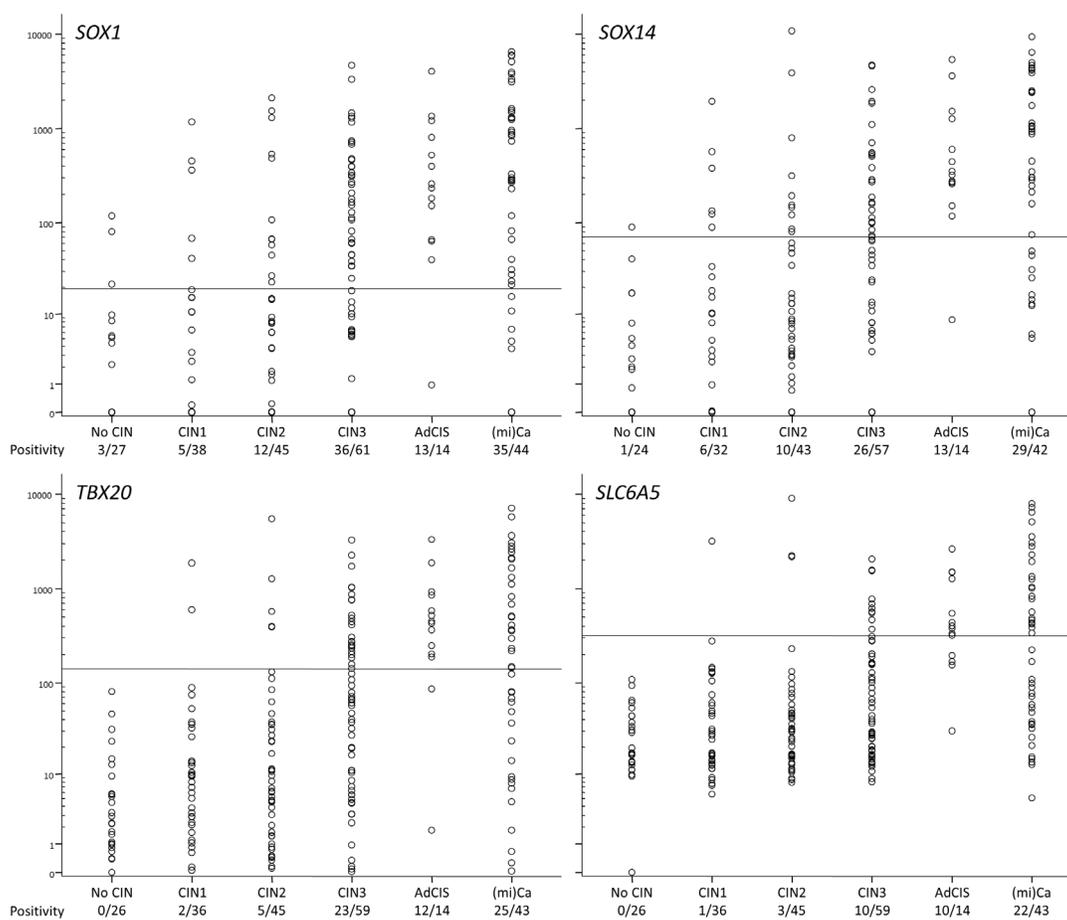
In this study, we used MethylCap-seq to draw detailed methylome maps. Enrichment for methylated DNA by either MBD proteins (MethylCap) or antibodies

(MeDIP) allows comparable distinction between methylated and unmethylated regions as bisulfite-based methods, but is less accurate to quantify the DNA methylation levels in partially methylated genomic regions. However, MethylCap is able to detect roughly twice as many DMRs compared to MeDIP at comparable sequencing depths [46]. Recent data showed that although MethylCap-seq was less sensitive compared to the array-based method of Infinium, more regions could be identified genome-wide [47]. Pyrosequencing did offer single-base resolution and could verify the MethylCap-seq results. Subsequently, primers for (Q)MSP were designed as these assays are more appropriate for high-throughput diagnostics. All five identified candidate markers discriminated between normal epithelium and cancer. A relatively good specificity was observed when using a threshold.

When we further validated the best four QMSPs on a series of scrapings of women referred with an abnormal smear, we observed that *SOX1* and *SOX14* provided a relatively good sensitivity, whereas *TBX20* and *SLC6A5* provided a relatively good specificity. Combining biomarker test results is a common choice to enhance the accuracy of clinical diagnosis [48]. Combining only

*SOX1* and *SOX14* seems to be sufficient to provide the highest sensitivity and specificity. As to cervical cancer diagnostics, an important advantage of DNA methylation markers is that they can be tested on the same material as used for HPV analyses [49, 50]. When comparing the detection of disease using hrHPV testing with the gene combination *SOX1* and *SOX14*, we observed no difference in CIN3+ sensitivity, but a higher specificity in the methylation test. However, these findings need to be validated in population-based screening cohorts; particularly because this population is not representative of a referral population. Besides, the remaining 20% of CIN2+ samples are considered by us as hrHPV test-negative and not definitely hrHPV-negative [51]. Possibly these samples were infected with a HPV type that is not detected by the assays or infected with multiple types (e.g. high-risk and low-risk) leading to a reduced sensitivity [52].

So far only a limited number of methylated genes have been examined in ADC, especially using cervical scrapings in a large series. These studies revealed markers with a different clinical utility, i.e. with a lower sensitivity for both ADC and SCC or either one [24, 26, 27, 53–58]. Two genes of the Wnt pathway, *DKK3* and *SFRP2*, showed



**Figure 6: Methylation levels in cervical scrapings of women referred with an abnormal smear.** The positive rate is depicted below the histological class labels.

**Table 3: Diagnostic performance of individual genes, gene combinations and hrHPV detection in cervical scrapings of women referred with an abnormal smear (ranked on sensitivity CIN3+)**

	specificity CIN0/1	sensitivity			P <sup>~</sup>
		CIN2+	CIN3+	(mi)Ca	
<b>Individual genes</b>					
<i>SOX1</i>	88%	59%	71%	80%	5×10 <sup>-15</sup>
<i>SOX14</i>	88%	50%	60%	69%	8×10 <sup>-11</sup>
<i>TBX20</i>	97%	40%	52%	58%	2×10 <sup>-12</sup>
<i>SLC6A5</i>	98%	28%	36%	51%	4×10 <sup>-10</sup>
<b>Gene combinations</b>					
<i>SOX1/SOX14/TBX20</i>	84%	63%	76%	83%	8×10 <sup>-15</sup>
<i>SLC6A5/SOX1/SOX14/TBX20</i>	84%	63%	76%	83%	8×10 <sup>-15</sup>
<i>SOX1/SOX14</i>	84%	63%	75%	83%	1×10 <sup>-14</sup>
<i>SLC6A5/SOX1/SOX14</i>	84%	63%	75%	83%	1×10 <sup>-14</sup>
<i>SOX1/TBX20</i>	87%	60%	72%	81%	4×10 <sup>-15</sup>
<i>SLC6A5/SOX1/TBX20</i>	87%	60%	72%	81%	4×10 <sup>-15</sup>
<i>SLC6A5/SOX1</i>	87%	58%	71%	81%	1×10 <sup>-14</sup>
<i>SOX14/TBX20</i>	88%	53%	64%	69%	2×10 <sup>-11</sup>
<i>SLC6A5/SOX14/TBX20</i>	88%	53%	64%	69%	2×10 <sup>-11</sup>
<i>SLC6A5/SOX14</i>	88%	50%	60%	69%	8×10 <sup>-11</sup>
<i>SLC6A5/TBX20</i>	97%	40%	52%	58%	2×10 <sup>-12</sup>
<b>hrHPV detection</b>					
GP5+/6+ and Cobas	42%	80%	80%	72%	5×10 <sup>-3</sup>

<sup>~</sup> linear-by-linear association test.

more methylation in ADC tissue compared to SCC tissue (82% vs. 18% and 84% vs. 39%) and combined analysis in cervical scrapes (n=8) detected all AdCIS and ADC [27]. Recently, *PAX1*, *PTPRR*, *SOX1* and *ZNF582*, previously reported to be frequently methylated in scrapings of SCC patients, were also analyzed in scrapings of ADC patients and showed a sensitivity of the single genes of 82% - 93% with a specificity of 81% - 95% in a Taiwanese population [26]. However, data on screening large cohorts with these markers is currently unavailable.

All of the 4 genes that we identified have previously been reported to be methylated in cancer. *SOX1* and *SOX14* belong to the SOX family of transcription factors having similar DNA binding specificities yet with divergent functions [59]. *SOX1* encodes a transcription factor implicated in the regulation of embryonic development and in the determination of cell fate. DNA methylation of *SOX1* in cervical cancer has been reported by Lai *et al.* [26, 44] albeit in a different region. Furthermore,

*SOX1* was identified as a tumor suppressor gene, because it interfered with Wnt/ $\beta$ -catenin signaling in cervical cancer cells [60] and hepatocellular carcinoma [61]. Hypermethylated *SOX1* was also found in ovarian cancer cells that are chronically exposed to cisplatin [62]. *SOX1* methylation, at least in part, is responsible for cisplatin resistance in human non-small cell lung cancer (NSCLC) [25, 63]. *SOX14*, in contrast to our data, has been reported to be a potential marker to differentiate between ADC and SCC, with more methylation or mutation in SCC as determined by NotI-microarrays [55].

T-box (TBX) transcription factors belong to an ancient gene family with critical roles in embryogenesis, in early cell fate decisions and in control of differentiation and organogenesis [64]. *TBX20* methylation has previously been related to specific bladder cancers [65], late stage hepatocellular carcinoma development [66], recurrence of lung adenocarcinoma [67] and cervical cancer [45]. *TBX20* expression has been related to colorectal cancer [68].

So far, methylation of *SLC6A5* (also known as the glycine transporter gene *GLYT2*) was associated with glioma [69], prostate cancer [70, 71], oral and pharyngeal cancer [72]; and the expression was down-regulated during rat liver regeneration [73]. Additionally, the *SLC6A5* gene can form multigene complexes under the influence of TNF signaling, and is subsequently activated by the NF- $\kappa$ B transcription factor [74].

Overall, our approach resulted in four new cervical cancer methylation markers with high specificity and high sensitivity for both cervical ADC and SCC. These results indicate that especially *SOX1* and *SOX14* are meaningful for cervical cancer screening.

## MATERIALS AND METHODS

### General strategy

In order to identify and validate cervical cancer-specific methylation markers both for ADC and SCC, the following strategy was applied (schematically represented in Figure 1). Step 1, DNA from snap-frozen tissue of 12 cervical cancers (ADC=6, SCC=6) and 20 normal cervixes was analyzed using MethylCap-seq. Subsequently, the differentially methylated regions (DMRs) were identified between normal cervixes and both cancer subtypes. Step 2, among the methylation candidates, the top 15 was selected for verification by methylation-specific PCR (MSP) or pyrosequencing on the same frozen tissues from step 1. Step 3, using the selected candidates from step 2, MSP was performed on DNA of FFPE tissues from an independent series of 17 normal and 13 cancer samples (ADC=6, SCC=7). Step 4, the candidate regions that showed more methylation in cancer tissues were selected for further clinical validation by quantitative MSP (QMSP) on cervical scrapings from a large series of cervical cancer patients (n=125: comprising 57 ADC and 68 SCC) and 89 controls of comparable age. Step 5, the markers that best distinguished scrapings from normal cervixes and cervical cancers were selected for a second diagnostic evaluation, provided the markers could sufficiently detect *both* ADC and SCC. QMSP was performed on scrapings of women referred with an abnormal smear and with known histological diagnosis (n=229: no CIN=27, CIN1=38, CIN2=45, CIN3=61, AdCIS=14 and (mi)Ca=44: ADC=12, SCC=29, adenosquamous=3).

### Patients

Patients with cervical cancer referred to the outpatient clinic of the University Medical Centre Groningen (UMCG) are asked to participate in our ongoing 'Methylation study' that has been approved by the Institutional Review Board of University Medical Centre Groningen, the Netherlands. All patients from whom

material was obtained gave written informed consent. Snap-frozen tissue, FFPE tissue and scrapings for this study are prospectively collected and stored in our tissue bank.

Within our 'Methylation study', normal tissue samples and normal scrapings are also collected from patients planned to undergo a hysterectomy for non-malignant reasons. All cervical tissue that was used for the normal control group was judged as histopathological normal. Additionally, all women without cervical disease never had an abnormal cervical smear prior to inclusion. Patients referred with cervical cancer are staged according to the FIGO criteria with pelvic examination and biopsies under general anesthesia. All cervical scrapings are collected prior to treatment. All cervical tissue samples were scored by an experienced gynecologic pathologist and the histological classification was used as the reference standard. All clinicopathological data were retrieved from patient files and stored in our large anonymous password-protected institutional Gynecologic Oncology database.

For MethylCap-seq and pyrosequencing, frozen tissue specimens were collected from 20 patients with a normal cervix (median age: 43 years, IQR 33-45) and 12 cancer patients (median age: 44 years, IQR 27-69) composed of 6 SCC and 6 ADC. For MSP analysis, FFPE tissue was collected from 17 patients with a normal cervix (median age: 43 years, IQR 40-44) and 13 cervical cancer patients (median age: 49 years, IQR 42-54) including 6 ADC and 7 SCC. For QMSP, scrapings were collected from 89 patients with normal cervixes (median age: 47 years, IQR 43-53), and from 125 cervical cancer patients (median age: 49 years, IQR 39-63, P=0.32) comprising 68 SCC and 57 ADC. See Supplementary Table S6 for the histological classification, age and FIGO stage of all cervical cancer patients in this study.

Likewise, scrapings were selected from women in whom an abnormal cervical smear was found (n=229). These scrapings are prospectively collected from women referred to our outpatient clinic for colposcopy after being tested with an abnormal smear in the population-based screening program. Here we randomly selected scrapings – i.e. within each diagnosis group – from 27 women diagnosed without CIN (median age: 36, IQR 30-48), 38 with CIN1 (median age: 40, IQR 31 to 45), 45 CIN2 (median age: 35, IQR 31-40), 61 CIN3 (median age: 35, IQR 32-40), 14 adenocarcinoma in situ (AdCIS) (median age: 36, IQR 32-44) and 44 with (micro-invasive) carcinoma (median age: 40, IQR 36-50). Of the AdCIS scrapings 3 were collected at the Meander Medical Centre (Amersfoort, the Netherlands) and 6 at the VU University Medical Centre (Amsterdam, the Netherlands) [75, 76]. The (mi)Ca cases comprised 12 ADC (median age: 40, IQR 36 to 49), 29 SCC (median age: 40, IQR 36 to 50) and 3 adenosquamous carcinomas (see Supplementary Table S5).

## Sample collection and DNA isolation

Preparation of tissue slides, enrichment for epithelial (tumor) cells by macrodissection, collection and processing of cervical scrapings, the DNA isolation and the assessment of the DNA's structural integrity were as described previously [22, 23, 77, 78]. DNA concentrations and 260/280 ratios were measured using a Nanodrop ND-1000 spectrophotometer (Thermo Scientific, Waltham, MA, USA). A 260/280 ratio around 1.8 and the capability to produce amplicons of at least 300 base pairs (bp) was required for all DNA samples. For the MethylCap-seq samples, DNA concentrations were measured using the Quant-iT™ PicoGreen® dsDNA Assay kit (Invitrogen, Carlsbad, CA, USA).

## MethylCap-seq

To assess genome-wide methylation patterns, MethylCap-seq was performed at the University of Ghent using the MethylCap kit according to manufacturer's instructions (Diagenode, Liège, Belgium) as previously described [33, 34]. Briefly, DNA samples (500 ng) were sheared to a size range of 300 - 1000 bp using a Bioruptor™ UCD-200 (Diagenode, Liège, Belgium) and fragments of approximately 300 bp were isolated and subsequently captured. Captured DNA was paired-end-sequenced on the Illumina Genome Analyzer II platform according to protocol (Illumina, San Diego, CA, USA). Leukocyte DNA of 4 healthy women was also included in 2 sets of 2 samples. Results were mapped using Bowtie software [79], visualized using BioBix' H2G2 browser (<http://h2g2.ugent.be/>) and processed using the human reference genome (NCBI build 37). The paired-end fragments were unique and located within 400 bp of each other [80].

## MethylCap-seq analysis

Read data of the promoters and exons were retrieved and dichotomized into methylation positive (if  $\geq 3$  reads) or methylation negative (if 0 or 1 read). Subsequently, Fisher's exact test was performed to identify DMRs between ADC and normal and between SCC and normal. To downsize the number of DMRs and to pinpoint candidate methylation markers in cervical cancer the following additional criteria were applied (depicted in Figure 2): 1) at least 75% (15/20) of the normal cervix group was methylation negative; 2) at least 50% (3/6) of ADC as well as at least 50% (3/6) of SCC was methylation positive; 3) no methylation in leukocytes, i.e. no more than 1 read for both sample pools or more than 2 reads for any individual leukocyte pool; 4) the region is at least 30 bp long.

## Gene ontology analysis

The Ensembl gene identifiers that were coupled to the DMRs were used for functional classification and GO

analyses using DAVID v6.7 [81] including the annotated hypermethylated DMRs in ADC or SCC.

## Bisulfite treatment

Sodium bisulfite modification of denatured genomic DNA was performed as previously reported [82]. One microgram of genomic DNA per sample was converted using the EZ DNA methylation kit (Zymo Research Corp, Irvine, US-CA). Leukocyte DNA from healthy women and WGA (illustra ready-to-go GenomiPhi HY kit, GE healthcare, Little Chalfont, UK), were used as negative controls for methylation, whereas *in vitro* methylated leukocyte DNA, produced using M. SssI methyltransferase (New England Biolabs, Ipswich, US-MA), served as a positive control.

## Pyrosequencing

Bisulfite-modified DNA (BS-DNA) was amplified using PyroMark PCR kit reagents and conditions (Qiagen, Hilden, Germany), yet we used a universal biotinylated primer as previously described [33, 83]. Sample preparation and pyrosequencing was performed on a PyroMark Q24 platform using PyroGold Q24 reagents (Qiagen, Hilden, Germany). Experiments were designed with PyroMark Assay Design 2.0; primers were checked for specificity with BiSearch [84] and sequences are available upon request. Non-template control (water), positive and negative controls were used in each reaction.

## Methylation-specific PCR

Each reaction was performed in 30  $\mu$ l total reaction volume, containing 600 nM of each primer, 1.5  $\mu$ l BS-DNA (approximately 15 ng), and 0.5 U AmpliTaq Gold DNA polymerase (Applied Biosystems, Carlsbad, CA, USA). The thermal profile of the MSP was 10' hot-start at 95°C, 40 cycles of 95°C for 60", 60°C for 60", 72°C 60", and finally an elongation step of 7' at 72°C. PCR products were separated on a 2.5% agarose gel, pre-stained with 0.5  $\mu$ g/ml ethidium bromide and visualized by UV transillumination. Non-template control (water) and positive/negative controls were used in each reaction. MSP primers were designed with Methyl Primer Express version 1.0 (Applied Biosystems, Carlsbad, CA, USA), were checked for specificity with BiSearch [84] and are available upon request. Separate reactions were performed to detect either unmethylated or methylated template.

## Quantitative methylation-specific PCR

QMSP was performed as we described previously [22] with a double-quenched hybridization probe (Integrated DNA Technologies, Leuven, Belgium). Probe sequences are available upon request. The *ACTB* gene was used as a methylation independent reference reaction. QMSP was performed in 10  $\mu$ l containing 300

nM of each primer, 200 nM probe, QuantiTect Probe PCR Master Mix (Qiagen, Hilden, Germany) and 2.5 µl BS-DNA (approximately 25 ng). Each sample was analyzed in triplicate by ABI PRISM® 7900HT Sequence Detection System (Applied Biosystems, Carlsbad, CA, USA). Serial dilutions of *in vitro* methylated leukocyte DNA enabled absolute quantification of (methylated) template. A DNA sample was considered methylated if at least 2 out of the 3 wells were methylation positive with a  $C_q$  below 50 with at least 225 pg *ACTB* input. The relative level of methylation of the region of interest was expressed as: (average quantity of methylated DNA / average quantity of *ACTB*) x 10000 [85].

### High-risk HPV testing

The presence of clinically relevant levels of hrHPV DNA was assessed using GP5+/6+ PCR and subsequently by Cobas HPV PCR as described previously [34]. The 6 AdCIS samples that were collected at the VU University Medical Centre (Amsterdam, the Netherlands) were previously tested hrHPV-positive with a GP5+/6+ enzyme immunoassay [86].

### Statistical analysis

Statistical analysis was performed using IBM SPSS Statistics 22 (IBM Corporation, New York, US-NY). Chi-square test and Fisher's exact test for small numbers were used to analyze the different methylation frequency between normal and cancer. The average methylation level of each frozen tissue sample and MethylCap-seq reads were correlated using Spearman's rank test. The Mann-Whitney U test was used to determine differences in median methylation levels between 2 groups. The sensitivity, specificity, ROC curves and area under the ROC curve (AUC) were calculated for the first diagnostic evaluation (normal vs. cervical cancer) [87]. The optimal threshold was calculated based on the largest Youden's index J [88, 89]. The Jonckheere-Terpstra test was used to assess whether the methylation levels changed with the severity of the underlying lesion. The chi-square linear-by-linear test was applied to analyze the dichotomous results across histological types. The McNemar-Bowker test was employed to assess differences in test classification, and the McNemar  $X^2$  test was subsequently used to attribute differences to either sensitivity or specificity or both. A gene combination labeled a sample positive if at least one of those QMSPs produced a positive test. A P-value below 0.05 was considered to be significant.

### ACKNOWLEDGMENTS

RWvL, RW and AB performed the (Q)MSP experiments. HGK and RDMS collected all cervical scrapings and clinical data. RW, WvC, TdM and GBAW

performed and analyzed the MethylCap-seq data. RWvL, RW, AB and GBAW performed the statistical analyses. AGJvdZ, ES and GBAW conceived and designed the experiments. RW, RWvL, AGJvdZ, ES and GBAW wrote the paper. All authors read and approved the final manuscript.

### GRANT SUPPORT

This project was financially supported by the Graduate School of Medical Sciences, University Medical Centre Groningen, Groningen, the Netherlands and by the Dutch Cancer Society (RUG-NKB2009-4577).

### CONFLICTS OF INTEREST

RDMS, WvC, AGJvdZ, ES and GBAW hold patents related to the content of the manuscript. ES is a member of the scientific advisory board of Roche, Hologic and QCMD, and received travel reimbursements from Roche, Abbott, Hologic Inc. and QCMD. WvC is an employee of MDxHealth Inc. (Irvine, CA, USA). RDMS has a minority stake in Self-Screen B.V., a spin-off company of the VU University Medical Center.

### REFERENCES

1. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *International Journal of Cancer*. 2010; 127: 2893-2917.
2. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA: A Cancer Journal for Clinicians*. 2015; 2015; 65: 87-108.
3. Vizcaino AP, Moreno V, Bosch FX, Munoz N, Barros-Dios XM, Borras J, Parkin DM. International trends in incidence of cervical cancer: II. Squamous-cell carcinoma. *International Journal of Cancer*. 2000; 86: 429-435.
4. Vizcaino AP, Moreno V, Bosch FX, Munoz N, Barros-Dios XM, Parkin DM. International trends in the incidence of cervical cancer: I. Adenocarcinoma and adenosquamous cell carcinomas. *International Journal of Cancer*. 1998; 75: 536-545.
5. Seoud M, Tjalma WA, Ronsse V. Cervical adenocarcinoma: moving towards better prevention. *Vaccine*. 2011; 29: 9148-9158.
6. Vinh-Hung V, Bourgain C, Vlastos G, Cserni G, De Ridder M, Storme G, Vlastos AT. Prognostic value of histopathology and trends in cervical cancer: a SEER population study. *BMC Cancer*. 2007; 7: 164.
7. Bulk S, Visser O, Rozendaal L, Verheijen RH, Meijer CJ. Cervical cancer in the Netherlands 1989-1998: Decrease of squamous cell carcinoma in older women, increase of

- adenocarcinoma in younger women. *International Journal of Cancer*. 2005; 113: 1005-1009.
8. Pettersson BF, Hellman K, Vaziri R, Andersson S, Hellström A. Cervical cancer in the screening era: who fell victim in spite of successful screening programs? *Journal of Gynecologic Oncology*. 2011; 22: 76.
  9. Lai CH, Hsueh S, Hong JH, Chang TC, Tseng CJ, Chou HH, Huang KG, Lin JD. Are adenocarcinomas and adenosquamous carcinomas different from squamous carcinomas in stage IB and II cervical cancer patients undergoing primary radical surgery? *International Journal of Gynecological Cancer*. 1999; 9: 28-36.
  10. Eifel PJ, Burke TW, Morris M, Smith TL. Adenocarcinoma as an independent risk factor for disease recurrence in patients with stage IB cervical carcinoma. *Gynecologic Oncology*. 1995; 59: 38-44.
  11. Hopkins MP, Morley GW. A comparison of adenocarcinoma and squamous cell carcinoma of the cervix. *Obstetrics and Gynecology*. 1991; 77: 912-917.
  12. Chen RJ, Lin YH, Chen CA, Huang SC, Chow SN, Hsieh CY. Influence of histologic type and age on survival rates for invasive cervical carcinoma in Taiwan. *Gynecologic Oncology*. 1999; 73: 184-190.
  13. Gien LT, Beauchemin MC, Thomas G. Adenocarcinoma: a unique cervical cancer. *Gynecologic Oncology*. 2010; 116: 140-146.
  14. zur Hausen H. Papillomaviruses causing cancer: evasion from host-cell control in early events in carcinogenesis. *Journal of the National Cancer Institute*. 2000; 92: 690-698.
  15. Costa S, Venturoli S, Origoni M, Preti M, Mariani L, Cristoforoni P, Sandri MT. Performance of HPV DNA testing in the follow-up after treatment of high-grade cervical lesions, adenocarcinoma in situ (AIS) and microinvasive carcinoma. *Ecancermedalscience*. 2015; 9: 528.
  16. Arbyn M, Verdoodt F, Snijders PJ, Verhoef VM, Suonio E, Dillner L, Minozzi S, Bellisario C, Banzi R, Zhao FH, Hillemanns P, Anttila A. Accuracy of human papillomavirus testing on self-collected versus clinician-collected samples: a meta-analysis. *The Lancet Oncology*. 2014; 15: 172-183.
  17. Pileggi C, Flotta D, Bianco A, Nobile CGA, Pavia M. Is HPV DNA testing specificity comparable to that of cytological testing in primary cervical cancer screening? Results of a meta-analysis of randomized controlled trials. *International Journal of Cancer*. 2014; 2014; 135: 166-177.
  18. Kinney W, Fetterman B, Cox JT, Lorey T, Flanagan T, Castle PE. Characteristics of 44 cervical cancers diagnosed following Pap-negative, high risk HPV-positive screening in routine clinical practice. *Gynecologic Oncology*. 2011; 121: 309-313.
  19. Hoque MO, Kim MS, Ostrow KL, Liu J, Wisman GB, Park HL, Poeta ML, Jeronimo C, Henrique R, Lendvai A, Schuurin E, Begum S, Rosenbaum E, et al. Genome-wide promoter analysis uncovers portions of the cancer methylome. *Cancer Research*. 2008; 68: 2661-2670.
  20. Baylin SB, Ohm JE. Epigenetic gene silencing in cancer - a mechanism for early oncogenic pathway addiction? *Nature Reviews Cancer*. 2006; 6: 107-116.
  21. Esteller M, Corn PG, Baylin SB, Herman JG. A gene hypermethylation profile of human cancer. *Cancer Research*. 2001; 61: 3225-3229.
  22. Eijnsink JJ, Lendvai A, Deregowski V, Klip HG, Verpooten G, Dehaspe L, de Bock GH, Hollema H, van Criekinge W, Schuurin E, van der Zee AG, Wisman GB. A four-gene methylation marker panel as triage test in high-risk human papillomavirus positive patients. *International Journal of Cancer*. 2012; 130: 1861-1869.
  23. Yang N, Nijhuis ER, Volders HH, Eijnsink JJ, Lendvai A, Zhang B, Hollema H, Schuurin E, Wisman GB, van der Zee AG. Gene promoter methylation patterns throughout the process of cervical carcinogenesis. *Cellular Oncology*. 2010; 32: 131-143.
  24. Wentzensen N, Sherman ME, Schiffman M, Wang SS. Utility of methylation markers in cervical cancer early detection: appraisal of the state-of-the-science. *Gynecologic Oncology*. 2009; 112: 293-299.
  25. Kang S, Kim JW, Kang GH, Lee S, Park NH, Song YS, Park SY, Kang SB, Lee HP. Comparison of DNA hypermethylation patterns in different types of uterine cancer: cervical squamous cell carcinoma, cervical adenocarcinoma and endometrial adenocarcinoma. *International Journal of Cancer*. 2006; 118: 2168-2171.
  26. Chang CC, Huang RL, Wang HC, Liao YP, Yu MH, Lai HC. High methylation rate of LMX1A, NKX6-1, PAX1, PTPRR, SOX1, and ZNF582 genes in cervical adenocarcinoma. *International Journal of Gynecological Cancer*. 2014; 24: 201-209.
  27. van der Meide WF, Snellenberg S, Meijer CJ, Baalbergen A, Helmerhorst TJ, van der Sluis WB, Snijders PJ, Steenbergen RD. Promoter methylation analysis of WNT/beta-catenin signaling pathway regulators to detect adenocarcinoma or its precursor lesion of the cervix. *Gynecologic Oncology*. 2011; 123: 116-122.
  28. Ongenaert M, Wisman GB, Volders HH, Koning AJ, Zee AG, van Criekinge W, Schuurin E. Discovery of DNA methylation markers in cervical cancer using relaxation ranking. *BMC Medical Genomics*. 2008; 1: 57-8794-1-57.
  29. Wilson IM, Davies JJ, Weber M, Brown CJ, Alvarez CE, MacAulay C, Schubeler D, Lam WL. Epigenomics: mapping the methylome. *Cell Cycle*. 2006; 5: 155-158.
  30. Weber M, Davies JJ, Wittig D, Oakeley EJ, Haase M, Lam WL, Schubeler D. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nature Genetics*. 2005; 37: 853-862.
  31. Lendvai A, Johannes F, Grimm C, Eijnsink JJ, Wardenaar R, Volders HH, Klip HG, Hollema H, Jansen RC, Schuurin

- E, Wisman GB, van der Zee AG. Genome-wide methylation profiling identifies hypermethylated biomarkers in high-grade cervical intraepithelial neoplasia. *Epigenetics*. 2012; 7: 1268-1278.
32. Hurd PJ, Nelson CJ. Advantages of next-generation sequencing versus the microarray in epigenetic research. *Briefings in Functional Genomics & Proteomics*. 2009; 8: 174-183.
  33. Clausen MJ, Melchers LJ, Mastik MF, Slagter-Menkema L, Groen HJ, van der Laan, Bernard FAM, van Criekinge W, de Meyer T, Denil S, Wisman GBA. Identification and validation of WISP1 as an epigenetic regulator of metastasis in oral squamous cell carcinoma. *Genes, Chromosomes and Cancer*. 2016; 55: 45-59.
  34. Boers A, Wang R, van Leeuwen R, Klip H, de Bock G, Hollema H, van Criekinge W, de Meyer T, Denil S, van der Zee, AG J. Discovery of new methylation markers to improve screening for cervical intraepithelial neoplasia grade 2/3. *Clinical Epigenetics*. 2016; 8: 1.
  35. Huang RL, Gu F, Kirma NB, Ruan J, Chen CL, Wang HC, Liao YP, Chang CC, Yu MH, Pilrose JM, Thompson IM, Huang HC, Huang TH, et al. Comprehensive methylome analysis of ovarian tumors reveals hedgehog signaling pathway regulators as prognostic DNA methylation biomarkers. *Epigenetics*. 2013; 8: 624-634.
  36. Huang TT, Gonzales CB, Gu F, Hsu YT, Jadhav RR, Wang CM, Redding SW, Tseng CE, Lee CC, Thompson IM, Chen HR, Huang TH, Kirma NB. Epigenetic deregulation of the anaplastic lymphoma kinase gene modulates mesenchymal characteristics of oral squamous cell carcinomas. *Carcinogenesis*. 2013; 34: 1717-1727.
  37. Carvalho RH, Haberle V, Hou J, van Gent T, Thongjuea S, van Ijcken W, Kockx C, Brouwer R, Rijkers E, Sieuwerts A, Foekens J, van Vroonhoven M, Aerts J, et al. Genome-wide DNA methylation profiling of non-small cell lung carcinomas. *Epigenetics & Chromatin*. 2012; 5: 9-8935-5-9.
  38. Wild CP, Scalbert A, Herceg Z. Measuring the exposome: a powerful basis for evaluating environmental exposures and cancer risk. *Environmental and Molecular Mutagenesis*. 2013; 54: 480-499.
  39. Lacey JV, Jr, Frisch M, Brinton LA, Abbas FM, Barnes WA, Gravitt PE, Greenberg MD, Greene SM, Hadjimichael OC, McGowan L, Mortel R, Schwartz PE, Zaino RJ, et al. Associations between smoking and adenocarcinomas and squamous cell carcinomas of the uterine cervix (United States). *Cancer Causes & Control*. 2001; 12: 153-161.
  40. Lacey JV, Jr, Swanson CA, Brinton LA, Altekruze SF, Barnes WA, Gravitt PE, Greenberg MD, Hadjimichael OC, McGowan L, Mortel R, Schwartz PE, Kurman RJ, Hildesheim A. Obesity as a potential risk factor for adenocarcinomas and squamous cell carcinomas of the uterine cervix. *Cancer*. 2003; 98: 814-821.
  41. Robertson KD, Jones PA. DNA methylation: past, present and future directions. *Carcinogenesis*. 2000; 21: 461-467.
  42. Herman JG, Baylin SB. Gene silencing in cancer in association with promoter hypermethylation. *The New England Journal of Medicine*. 2003; 349: 2042-2054.
  43. Baylin SB, Herman JG. DNA hypermethylation in tumorigenesis: epigenetics joins genetics. *Trends In Genetics*. 2000; 16: 168-174.
  44. Lai HC, Lin YW, Huang TH, Yan P, Huang RL, Wang HC, Liu J, Chan MW, Chu TY, Sun CA, Chang CC, Yu MH. Identification of novel DNA methylation markers in cervical cancer. *International Journal of Cancer*. 2008; 123: 161-167.
  45. Farkas SA, Milutin-Gašperov N, Grce M, Nilsson TK. Genome-wide DNA methylation assay reveals novel candidate biomarker genes in cervical cancer. *Epigenetics*. 2013; 8: 1213-1225.
  46. Bock C, Tomazou EM, Brinkman AB, Muller F, Simmer F, Gu H, Jager N, Gnirke A, Stunnenberg HG, Meissner A. Quantitative comparison of genome-wide DNA methylation mapping technologies. *Nature Biotechnology*. 2010; 28: 1106-1114.
  47. De Meyer T, Bady P, Trooskens G, Kurscheid S, Bloch J, Kros JM, Hainfellner JA, Stupp R, Delorenzi M, Hegi ME, Van Criekinge W. Genome-wide DNA methylation detection by MethylCap-seq and Infinium HumanMethylation450 BeadChips: an independent large-scale comparison. *Scientific Reports*. 2015; 5: 15375.
  48. Steenbergen RD, Snijders PJ, Heideman DA, Meijer CJ. Clinical implications of (epi)genetic changes in HPV-induced cervical precancerous lesions. *Nature Reviews Cancer*. 2014; 14: 395-405.
  49. Boers A, Bosgraaf RP, van Leeuwen RW, Schuurings E, Heideman DA, Massuger LF, Verhoef VM, Bulten J, Melchers WJ, van der Zee AG, Bekkers RL, Wisman GB. DNA methylation analysis in self-sampled brush material as a triage test in hrHPV-positive women. *British Journal of Cancer*. 2014; 111: 1095-1101.
  50. Eijssink JJ, Yang N, Lendvai A, Klip HG, Volders HH, Buikema HJ, van Hemel BM, Voll M, Coelingh Bennink HJ, Schuurings E, Wisman GB, van der Zee AG. Detection of cervical neoplasia by DNA methylation analysis in cervico-vaginal lavages, a feasibility study. *Gynecologic Oncology*. 2011; 120: 280-283.
  51. Evans MF, Adamson CS, Schned LM, St John TL, Leiman G, Ashikaga T, Cooper K. HPV is detectable in virtually all abnormal cervical cytology samples after reinvestigation of HPV negatives with multiple alternative PCR tests. *Diagnostic Molecular Pathology, part B*. 2010; 19: 144-150.
  52. Mori S, Nakao S, Kukimoto I, Kusumoto-Matsuo R, Kondo K, Kanda T. Biased amplification of human papillomavirus DNA in specimens containing multiple human papillomavirus types by PCR with consensus primers. *Cancer Science*. 2011; 102: 1223-1227.
  53. Blanco-Luquin I, Guarch R, Ojer A, Pérez-Janices N, Martín-Sánchez E, Maria-Ruiz S, Monreal-Santesteban I, Blanco-Fernandez L, Pernaut-Leza E, Escors D. Differential

- role of gene hypermethylation in adenocarcinomas, squamous cell carcinomas and cervical intraepithelial lesions of the uterine cervix. *Pathology International*. 2015; 65: 476-485.
54. Banzai C, Nishino K, Quan J, Yoshihara K, Sekine M, Yahata T, Tanaka K. Promoter methylation of DAPK1, FHIT, MGMT, and CDKN2A genes in cervical carcinoma. *International Journal of Clinical Oncology*. 2014; 19: 127-132.
  55. Senchenko VN, Kisseljova NP, Ivanova TA, Dmitriev AA, Krasnov GS, Kudryavtseva AV, Panasenkov GV, Tsitrin EB, Lerman MI, Kisseljov FL, Kashuba VI, Zabarovsky ER. Novel tumor suppressor candidates on chromosome 3 revealed by NotI-microarrays in cervical cancer. *Epigenetics*. 2013; 8: 409-420.
  56. Kim J, Choi YD, Lee JS, Lee JH, Nam JH, Choi C, Kweon S, Fackler MJ, Sukumar S. Quantitative assessment of DNA methylation for the detection of cervical neoplasia in liquid-based cytology specimens. *Virchows Archiv*. 2010; 457: 35-42.
  57. Lee E, McClelland M, Wang Y, Long F, Choi S, Lee J. Distinct DNA methylation profiles between adenocarcinoma and squamous cell carcinoma of human uterine cervix. *Oncology Research Featuring Preclinical and Clinical Cancer Therapeutics*. 2009; 18: 401-408.
  58. Lin Y, Chung M, Lai H, De Yan M, Shih Y, Chang C, Yu M. Methylation analysis of SFRP genes family in cervical adenocarcinoma. *Journal of cancer research and clinical oncology*. 2009; 135: 1665-1674.
  59. Malas S, Duthie S, Deloukas P, Episkopou V. The isolation and high-resolution chromosomal mapping of human SOX14 and SOX21; two members of the SOX gene family related to SOX1, SOX2, and SOX3. *Mammalian Genome*. 1999; 10: 934-937.
  60. Lin YW, Tsao CM, Yu PN, Shih YL, Lin CH, Yan MD. SOX1 suppresses cell growth and invasion in cervical cancer. *Gynecologic Oncology*. 2013; 131: 174-181.
  61. Tsao CM, Yan MD, Shih YL, Yu PN, Kuo CC, Lin WC, Li HJ, Lin YW. SOX1 functions as a tumor suppressor by antagonizing the WNT/beta-catenin signaling pathway in hepatocellular carcinoma. *Hepatology*. 2012; 56: 2277-2287.
  62. Yu W, Jin C, Lou X, Han X, Li L, He Y, Zhang H, Ma K, Zhu J, Cheng L, Lin B. Global analysis of DNA methylation by Methyl-Capture sequencing reveals epigenetic control of cisplatin resistance in ovarian cancer cell. *PLoS One*. 2011; 6: e29450.
  63. Li N, Li X, Li S, Zhou S, Zhou Q. Cisplatin-induced downregulation of SOX1 increases drug resistance by activating autophagy in non-small cell lung cancer cell. *Biochemical and Biophysical Research Communications*. 2013; 439: 187-190.
  64. Naiche L, Harrelson Z, Kelly RG, Papaioannou VE. T-box genes in vertebrate development. *Annual Reviews Genetics*. 2005; 39: 219-239.
  65. Zwarthoff EC, Van Tilborg AG. Method of diagnosing bladder cancer. 2010; US20120101023 A1.
  66. Zhao Y, Xue F, Sun J, Guo S, Zhang H, Qiu B, Geng J, Gu J, Zhou X, Wang W. Genome-wide methylation profiling of the different stages of hepatitis B virus-related hepatocellular carcinoma development in plasma cell-free DNA reveals potential biomarkers for early detection and high-risk monitoring of hepatocellular carcinoma. *Clinical Epigenetics*. 2014; 6: 1.
  67. Sato T, Arai E, Kohno T, Tsuta K, Watanabe S, Soejima K, Betsuyaku T, Kanai Y. DNA methylation profiles at precancerous stages associated with recurrence of lung adenocarcinoma. *PLoS One*. 2013; 8: e59444.
  68. Liu J, Wu S, Li M, Wang X, Tang Y. LncRNA expression profiles reveal the co-expression network in human colorectal carcinoma. *International Journal of Clinical and Experimental Pathology*. 2016; 9: 1885-1892.
  69. Hopp L, Willscher E, Löffler-Wirth H, Binder H. Function shapes content: DNA-methylation marker genes and their impact for molecular mechanisms of glioma. *Journal of Cancer Research Updates*. 2015; 4: 127-148.
  70. Bert SA, Robinson MD, Strbenac D, Statham AL, Song JZ, Hulf T, Sutherland RL, Coolen MW, Stirzaker C, Clark SJ. Regional activation of the cancer genome by long-range epigenetic remodeling. *Cancer Cell*. 2013; 23: 9-22.
  71. Schweiger M, Lehrach H, Sörnö S, Schlomm T, Sültmann H, Suter G. Prostate Cancer Markers. 2014; US 20140051082 A1.
  72. Langevin SM, Eliot M, Butler RA, Cheong A, Zhang X, McClean MD, Koestler DC, Kelsey KT. CpG island methylation profile in non-invasive oral rinse samples is predictive of oral and pharyngeal carcinoma. *Clinical Epigenetics*. 2015; 7: 1.
  73. Xu C, Chang C. Expression profiles of the genes associated with metabolism and transport of amino acids and their derivatives in rat liver regeneration. *Amino Acids*. 2008; 34: 91-102.
  74. Nguyen HQ, Bosco G. Gene Positioning Effects on Expression in Eukaryotes. *Annual Review of Genetics*. 2015; 49: 627-646.
  75. Rijkaart DC, Berkhof J, van Kemenade FJ, Rozendaal L, Verheijen RH, Bulk S, Herreijlers ME, Verweij WM, Snijders PJ, Meijer CJ. Comparison of HPV and cytology triage algorithms for women with borderline or mild dyskaryosis in population-based cervical screening (VUSA-screen study). *International Journal of Cancer*. 2010; 126: 2175-2181.
  76. Bulkman NW, Berkhof J, Rozendaal L, van Kemenade FJ, Boeke AJ, Bulk S, Voorhorst FJ, Verheijen RH, van Groningen K, Boon ME, Ruitinga W, van Ballegooijen M, Snijders PJ, et al. Human papillomavirus DNA testing for the detection of cervical intraepithelial neoplasia grade 3 and cancer: 5-year follow-up of a randomised controlled implementation trial. *Lancet*. 2007; 370: 1764-1772.

77. Wisman GB, Hollema H, de Jong S, ter Schegget J, Tjong-A-Hung SP, Ruiters MH, Krans M, de Vries EG, van der Zee AG. Telomerase activity as a biomarker for (pre) neoplastic cervical disease in scrapings and frozen sections from patients with abnormal cervical smear. *Journal of Clinical Oncology*. 1998; 16: 2238-2245.
78. van Dongen JJ, Langerak AW, Bruggemann M, Evans PA, Hummel M, Lavender FL, Delabesse E, Davi F, Schuurin E, Garcia-Sanz R, van Krieken JH, Droese J, Gonzalez D, et al. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. *Leukemia*. 2003; 17: 2257-2317.
79. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*. 2009; 10: R25-2009-10-3-r25. Epub 2009 Mar 4.
80. De Meyer T, Mampaey E, Vlemmix M, Denil S, Trooskens G, Renard J, De Keulenaer S, Dehan P, Menschaert G, Van Criekinge W. Quality evaluation of methyl binding domain based kits for enrichment DNA-methylation sequencing. *PLoS One*. 2013; 8: e59068.
81. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*. 2009; 4: 44-57.
82. Wisman GB, Nijhuis ER, Hoque MO, Reesink-Peters N, Koning AJ, Volders HH, Buikema HJ, Boezen HM, Hollema H, Schuurin E, Sidransky D, van der Zee AG. Assessment of gene promoter hypermethylation for detection of cervical neoplasia. *International Journal of Cancer*. 2006; 119: 1908-1914.
83. Colella S, Shen L, Baggerly KA, Issa JP, Krahe R. Sensitive and quantitative universal Pyrosequencing methylation analysis of CpG sites. *BioTechniques*. 2003; 35: 146-150.
84. Aranyi T, Varadi A, Simon I, Tusnady GE. The BiSearch web server. *BMC Bioinformatics*. 2006; 7: 431.
85. Reesink-Peters N, Wisman GB, Jeronimo C, Tokumaru CY, Cohen Y, Dong SM, Klip HG, Buikema HJ, Suurmeijer AJ, Hollema H, Boezen HM, Sidransky D, van der Zee AG. Detecting cervical cancer by quantitative promoter hypermethylation assay on cervical scrapings: a feasibility study. *Molecular Cancer Research*. 2004; 2: 289-295.
86. Rijkaart DC, Berkhof J, Rozendaal L, van Kemenade FJ, Bulkman NW, Heideman DA, Kenter GG, Cuzick J, Snijders PJ, Meijer CJ. Human papillomavirus testing for the detection of high-grade cervical intraepithelial neoplasia and cancer: final results of the POBASCAM randomised controlled trial. *The Lancet Oncology*. 2012; 13: 78-88.
87. Wang R, Wang G, Zhang N, Li X, Liu Y. Clinical evaluation and cost-effectiveness analysis of serum tumor markers in lung cancer. *BioMed Research International*. 2013; 2013: 195692.
88. YODEN WJ. Index for rating diagnostic tests. *Cancer*. 1950; 3: 32-35.
89. Fluss R, Faraggi D, Reiser B. Estimation of the Youden Index and its associated cutoff point. *Biometrical Journal*. 2005; 47: 458-472.