(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 September 2018 (13.09.2018)

WIPO | PCT

(10) International Publication Number
WO 2018/164577 A1

(51) International Patent Classification:
C12Q 1/6886 (2018.01)

(21) International Application Number:
PCT/NL2018/050143

(22) International Filing Date:
09 March 2018 (09.03.2018)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
17160346.7 10 March 2017 (10.03.2017) EP

(71) Applicant: **SELF-SCREEN B.V.** [NL/NL]; Biothof 15 1, 1098 RX Amsterdam (NL).

(72) Inventors: **MEIJER, Christophorus Johannes Lambertus Maria**; c/o Biothof 15 1, 1098 RX Amsterdam (NL). **STEENBERGEN, Renske Daniëla Maria**; c/o Biothof 15 1, 1098 RX Amsterdam (NL). **SNIJDERS, Petrus Josephus Ferdinandus**; c/o Biothof 15 1, 1098 RX Amsterdam (NL). **HEIDEMAN, Daniëlle Anne Marie**; c/o Biothof 15 1, 1098 RX Amsterdam (NL).

(74) Agent: **JANSEN, C.M.**; V.O., Carnegieplein 5, 2517 KJ Den Haag (NL).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*

WO 2018/164577 A1

(54) Title: A METHYLATION CLASSIFIER FOR DETECTION OF HPV-INDUCED INVASIVE CANCERS, NONHPV- INDUCED GYNAECOLOGICAL AND ANOGENITAL CANCERS AND THEIR HIGH-GRADE PRECURSOR LESIONS

(57) Abstract: The invention relates to method for detecting HPV-induced high-grade precancerous lesions, HPV-induced invasive cancers and nonHPV-induced gynaecological and anogenital cancers, said method comprising detection of a methylation classifier consisting of the genes LHX8, ASCL1 and ST6GALNAC5 and their regulatory sequence in a cell whereby such hypermethylation indicates the presence of HPV- induced precursor lesions with invasive potential, HPV-induced invasive cancers and nonHPV-induced gynaecological and anogenital cancers. The invention further comprises the use of the methylation classifier consisting of genes LHX8, ASCL1 and ST6GALNAC5 and their regulatory sequence in such a method and a test kit for the detection of LHX8. ASCL1 and/or ST6GALNAC5 methylation.

A methylation classifier for detection of HPV-induced invasive cancers, nonHPV-induced gynaecological and anogenital cancers and their high-grade precursor lesions

FIELD OF THE INVENTION

The invention relates to the field of cancer prevention and medical diagnostics; and is concerned with a molecular diagnostic assay for cancers, especially human papillomavirus (HPV)-induced invasive cancers and high-grade precursor lesions thereof, such as invasive cervical cancer and premalignant cervical lesions, nonHPV-induced gynaecological and anogenital cancers. In particular, the present invention relates to the use of a methylation classifier based on the genes LHX8, ASCL1 andST6GALNAC5 and their regulatory sequences as marker for hrHPV-induced invasive cancers, nonHPV-induced gynaecological and anogenital cancers and their premalignant lesions with invasive potential.

BACKGROUND OF THE INVENTION

Cancer of the uterine cervix is the fourth most common cancer in women world-wide and is responsible for approximately 250.000 cancer deaths a year.

Cervical squamous cell carcinoma development is characterized by a sequence of premalignant lesions, so-called cervical intraepithelial neoplasia (CIN), which are graded 1 to 3, referring to mild dysplasia (CIN 1), moderate dysplasia (CIN 2) and severe dysplasia/carcinoma *in situ* (CIN 3), respectively. CIN 1 is also referred to as low grade squamous intraepithelial lesion (LSIL) and CIN 2 and CIN 3 together as high grade squamous intraepithelial lesion (HSIL). For cervical adenocarcinoma, adenocarcinoma in situ (ACIS) is an established precursor lesion. In principle, these premalignant lesions are reversible, although the more severe the lesion, the lower the chance of spontaneous regression. Cervical cancer is considered a preventable disease because the premalignant stages can be detected by exfoliative cytology and treated relatively easily when necessary, with only minor side effects. Cervical screening is aimed to early diagnose the high-grade premalignant (i.e., CIN 2/3 and adenocarcinoma in situ) and treatable cancerous lesions, thereby reducing the mortality of invasive cervical cancer. General medical practice comprises the treatment of all women with morphologically confirmed CIN 2, CIN 3 and adenocarcinoma in situ, in order to prevent the development of cervical cancer.

Over the past decade it has been well established that cervical carcinogenesis is initiated by an infection with high-risk human papillomavirus (hrHPV). Expression of the viral oncogenes E6 and E7, which disturb the p53 and Rb tumor suppressor pathways, respectively, has been shown to be essential for both the onset of oncogenesis and the maintenance of a malignant phenotype. Therefore, testing for hrHPV appeared as an attractive, primary screening tool. However, consistent with a multistep process of carcinogenesis, additional alterations in the host cell genome are required for progression of an hrHPV infected cell to invasive cancer cell. Only a small proportion of women infected with high-risk HPV will develop high-grade premalignant cervical lesions (CIN 2/3) and, if left untreated, cervical cancer. In most women with premalignant cervical lesions the lesions regress spontaneously. Of the women who participate in population based screening, about 5-6% have a positive hrHPV test. However, only at maximum 20% of them (1% of the participating women) have ≥CIN 2/3. Therefore, primary screening by hrHPV testing will be accompanied with a substantial number of redundant follow-up procedures and unnecessary anxiety amongst women, unless markers can be applied to the cervical smears that allow stratification of hrHPV positive women for risk of ≥CIN 2/3 and ≥adenocarcinoma in situ.

A major challenge is to reduce the percentage of HPV test positive women to those that have clinically meaningful lesions. One mode is to use cytology as a secondary (so-called triage) test for hrHPV positive women. Still, this leaves a substantial number of hrHPV positive women with normal cytology (3.5% of the women in the screening population), of which still 10% have or acquire ≥CIN 3. Moreover, cytology is not an option for self-sampled cervico-vaginal specimens that can be taken at home, since these are not representative for the cytological status of the cervix. Another mode is to use HPV16/18 genotyping. This however leaves women with non-HPV16/18 types who are, although to a lesser extent, also at risk of ≥CIN 2/3 and ≥adenocarcinoma in situ. Therefore, there is a need for supplementary or alternative triage tools to stratify hrHPV positive women into those with and without ≥CIN 2/3 and ≥adenocarcinoma in situ.

Primary screening for cervical cancer using disease markers based on host cell changes in cancer genes provides a promising alternative provided that specificity and sensitivity is sufficiently high. This option is of particular interest for low and middle income countries, where quality-controlled cytology is absent and implementation of

follow-up algorithms for HPV-positive women is complicated. In these countries self-sampling has shown to facilitate access to cervical screening (Laczano-Ponce et al., Lancet. 2011; 378: 1868-1873). In this sense it is extremely useful to have markers that also prevail in self-samples. It appears that there is a huge difference in the 'behaviour' of markers on vaginal smears that have been obtained by medical skilled personnel like doctors or nurses and on vaginal swabs that have been collected by the woman herself. It has appeared that many markers that would be suitable for doctor-provided samples are not useful in self-samples. The necessity to work with self-samples instead of doctor samples is high in low and middle income countries since in those countries there is less medical personnel per capita and often there is a cultural problem by letting other persons taking a vaginal sample. Further, even in highly developed countries self-sampling is an ideal way to reduce the costs of obtaining vaginal samples for large scale cervical cancer screening. An even higher acceptance rate can be expected when urine samples can be used for early cancer detection.

Endometrial cancer is the most common gynaecologic malignancy in many developed countries (Siegel et al., CA Cancer J. Clin., 64 (2014), pp. 9–29). Early stage endometrial cancer has a very good prognosis. Therefore, early detection will increase the chance of cure and avoid or reduce cumbersome and costly therapeutic intervention. As shedding of tumor cells in cervical scrapes has been demonstrated, early detection of endometrial cancer by non-invasive sampling is feasible. However, conventional cytology on cervical scrapes has a very low sensitivity for detection of endometrial cancer. Testing for molecular alterations, such as DNA methylation, associated with endometrial cancer, may provide a promising approach for early detection of endometrial cancer (de Strooper et al., J Clin Pathol. 2014;67(12):1067-71, Bakkum-Gamez et al. Gynecol Oncol. 2015;137(1):14-22). Similarly, testing for cancer specific DNA alterations in cervical scrapes for the detection of ovarian cancer has recently been proposed (Kindle et al., Sci Transl Med. 2013; 5(167):1-21).

SUMMARY OF THE INVENTION

The inventors now have found a method for detecting HPV-induced invasive cancers, nonHPV-induced gynaecological and anogenital cancers, and their high-grade precancerous lesions wherein said method comprises the detection of hypermethylation in the gene sequence of LHX8, ASCL1 and ST6GALNAC5 or in the regulatory sequence of these genes in a cell whereby such hypermethylation indicates the

presence of HPV-induced precursor lesions with invasive potential, HPV-induced invasive cancers and/or nonHPV-induced gynaecological and anogenital cancers. Preferably, in such a method said HPV-induced high-grade precancerous lesion or HPV-induced invasive carcinoma is a high-grade premalignant cervical lesion or invasive cervical cancer, more preferably a high-risk HPV-induced invasive cancer. Preferably, said nonHPV-induced gynaecological cancer is a endometrial cancer. Preferably said nonHPV-induced anogenital cancer is a vulvar or penile cancer.

Also comprised in the present invention is a method for detecting vulva cancer and vulvar HPV-induced precursor lesions with invasive potential, said method comprising detection of a methylation classifier consisting of ASCL1 and one or more of ZIC1, ZNF-582 and LHX8 sequences in a cell whereby such hypermethylation indicates the presence of vulva cancer and vulvar HPV-induced precursor lesions with invasive potential.

Further comprised in the invention is a method for detecting anal cancer and anal HPV-induced precursor lesions with invasive potential, said method comprising detection of a methylation classifier consisting of ASCL1 and one or more of ZIC1 and ZNF-582 sequences in a cell whereby such hypermethylation indicates the presence of anal cancer and anal HPV-induced precursor lesions with invasive potential.

Also part of the invention is a method for the detection of ovarian cancer and ovarian precursor lesions, said method comprising detection of a methylation classifier consisting of ASCL1, and one or more of LHX8 and ZIC1 sequences in a cell whereby such hypermethylation indicates the presence of ovarian cancer and ovarian precursor lesions.

In a preferred embodiment of the invention the hypermethylation is detected in the sequences as indicated in Figures 1, 2, 3, 7 and 8. In a further preferred embodiment said hypermethylation is an increased methylation of the methylation classifier LHX8, ASCL1, ZNF-582, ZIC1 and ST6GALNAC5 CpG rich promoter and/or gene sequences, such as regulatory sequences, in a test cell as compared to a comparable normal cell

In a preferred embodiment of the invention the detection of (hyper)methylation is performed by using a methylation sensitive restriction endonuclease, chosen from the group consisting of BssHII, MspI, NotI and HpaII. In another preferred embodiment of the invention the detection of (hyper)methylation is performed using nanotechnology.

In an alternative preferred embodiment of the invention, the detection of (hyper)methylation is performed via a methylation specific PCR, which is based on bisulfite modification of DNA, followed by specific PCR reactions that target CpG rich sequences. Preferably in such a method the reagent is a nucleic acid probe or primer that binds to the nucleic acid as indicated in Figure 1, 2 or 3, and more preferably said nucleic acid probe or primer has a detectable label.

In another embodiment of the invention the nucleic acid probe has a nucleotide sequence selected from the group consisting of:

a) a polynucleotide sequence capable of hybridizing under stringent conditions to the sequence ASCL1 as set forth in Figure 1 or to the sequence LHX8 as set forth in Figure 2 or to the sequence ST6GALNAC5 as set forth in Figure 3, ; or to the sequence ZIC1 as set forth in Figure 8 or to the sequence ZNF-582 as set forth in Figure 7;

b) a polynucleotide having at least 70% identity to the polynucleotide of a);

c) a polynucleotide complementary to the polynucleotide of a); and

d) a polynucleotide comprising at least 15 bases of a nucleotide of a) or b).

Further preferred in the present method of the invention the methylation of the gene sequence of LHX8, ASCL1, ZNF-582, ZIC1 and ST6GALNAC5 or the regulatory sequence of these genes is determined.

Also part of the invention is the use of LHX8, ASCL1, ZNF-582, ZIC1 and ST6GALNAC5 as a molecular diagnostic marker for the detection of HPV-induced invasive carcinoma, nonHPV-induced gynaecological or anogenital carcinoma, and their high-grade precancerous lesion. Further part of the invention is the use of the methylation classifier LHX8, ASCL1, and ST6GALNAC5 as a molecular diagnostic marker for the detection of HPV-induced high-grade precancerous lesion or HPV-induced invasive carcinoma or nonHPV-induced gynaecological or anogenital cancer, preferably wherein the methylation of said marker is predictive for the occurrence of said lesion, carcinoma or cancer. Also part of the invention is the use of the methylation classifier ASCL1, and one of ZIC1, ZNF-582 and LHX8 as a molecular diagnostic marker for the detection of vulva cancer or vulvar HPV-induced high-grade precancerous lesion, preferably wherein the methylation of said marker is predictive for the occurrence of said lesion or cancer. Further part of the invention is the use of the methylation classifier ASCL1, and one of ZIC1 and ZNF-582 as a molecular

diagnostic marker for the detection of anal cancer or anal HPV-induced high-grade precancerous lesion, preferably wherein the methylation of said marker is predictive for the occurrence of said lesion or cancer. Next, part of the invention is formed by the use of the methylation classifier ASCL1, and one of LHX8 and ZIC1 as a molecular diagnostic marker for the detection of ovarian cancer, preferably wherein the methylation of said marker is predictive for the occurrence of said cancer.

The invention also comprises a kit of parts for use in a method of detecting HPV-induced invasive carcinoma, nonHPV-induced gynaecological or anogenital carcinoma and their high- grade precancerous lesion said kit comprising

-means for the detection of LHX8, ASCL1 and ST6GALNAC5 methylation wherein said means comprise probes and/or primers specific for the ASCL1 nucleotide sequence of Figure 1 and the LHX8 nucleotide sequence of Figure 2 and the ST6GALNAC5 nucleotide sequence of Figure 3;

-means for the detection of ASCL1 and ST6GALNAC5 methylation wherein said means comprise probes and/or primers specific for the ASCL1 nucleotide sequence of Figure 1 and the ST6GALNAC5 nucleotide sequence of Figure 3; or

-means for the detection of ASLCL1 and LHX8 methylation wherein said means comprise probes and/or primers specific for the ASCL1 nucleotide sequence of Figure 1 and the LHX8 nucleotide sequence of Figure 2, or

- means for the detection of ASCL1 and at least one of ZIC1, ZNF-582 and LHX8 wherein said means comprise probes and/or primers specific for the ASCL1 nucleotide sequence of Figure 1 and either of the ZNF-582 nucleotide sequence of Figure 7, the ZIC1 nucleotide sequence of Figure 8 and/or the LHX8 nucleotide sequence of Figure 2;

The kit preferably further comprises -means for the detection of HPV infection, wherein said means comprise probes and primers specific for HPV.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the ASLC1 promoter region and CpG rich sequence exon 1; the coding sequence is given in upper case.

Figure 2 shows the LHX8 5' regulatory region and coding sequence (exon 1). The coding sequence is in upper case.

Figure 3 shows the ST6GALNAC5 5' regulatory region, coding sequence (exon1 and 2) and intron 1. The coding sequence is in upper case.

Figure 4 shows boxplots of the methylation levels of LHX8, ASCL1 and ST6GALNAC5 in hrHPV-positive self-collected cervico-vaginal specimens as measured by quantitative methylation specific PCR (qMSP) in lavage samples (A) and brush self samples (B). In qMSP analysis the methylation values of the targets are normalised to reference gene *ACTB* using the comparative Ct method ($2^{-\Delta Ct}$). On the y-axes levels of methylated DNA are presented; on the x-axes groups represent controls (i.e. women with ≤CIN1), women with CIN2, CIN3, and squamous cervical cancer. The levels of methylation of LHX8, ASCL1 and ST6GALNAC5 increases with the severity of the disease and are significantly increased in CIN3 and SCC compared to controls. Significant differences in methylation levels between groups are indicated, * indicates <0.05, ** indicates <0.005, *** indicates <0.0005. C shows Receiver Operating Characteristics (ROC) curves.

Figure 5 A shows boxplots of the methylation levels of LHX8, ASCL1 and ST6GALNAC5 cervical scrapes and urine of controls (i.e. women with ≤CIN1) and women with cervical cancer as measured by quantitative methylation specific PCR (qMSP). In qMSP analysis the methylation values of the targets are normalised to reference gene *ACTB* using the comparative Ct method ($2^{-\Delta Ct}$). On the y-axes levels of methylated DNA are presented; on the x-axes groups represent controls (green; women with ≤CIN1) and women with cervical cancer (red). The levels of methylation of LHX8, ASCL1 and ST6GALNAC5 are significantly increased in both scrapes and urine of women with cervical cancer compared to controls. Significant differences in methylation levels between groups are indicated, * indicates <0.05, ** indicates <0.005, *** indicates <0.0005. B. shows Receiver Operating Characteristics (ROC) curve of any marker combination in hrHPV+ cervical scrapes, showing that a combination of markers enables the detection of all cancers at highest specificity. **C.** shows Receiver Operating Characteristics (ROC) curve of individual markers (left) and any marker combination in urine, showing that a combination of 3 markers enables the detection of all cancers at highest specificity.

Figure 6 shows the boxplots of the methylation levels of LHX8, ASCL1 and ST6GALNAC5 in cervical scrapes of women with endometrial cancer as measured by quantitative methylation specific PCR. In qMSP analysis the methylation values of the

targets are normalised to reference gene *ACTB* using the comparative Ct method ($2^{-\Delta Ct}$). On the y-axes levels of methylated DNA are presented; on the x-axes groups represent control women and women with endometrial cancer. The levels of methylation of LHX8, ASCL1 and ST6GALNAC5 are significantly increased in women with endometrial cancer compared to controls

Figure 7 shows the ZNF-582 5' regulatory region and coding sequence (exon 1). The coding sequence is in upper case.

Figure 8 shows the ZIC1 5' regulatory region and coding sequence (exon 1). The coding sequence is in upper case.

Figure 9 shows the boxplots of the methylation levels of (**A**) *ASCL1*, (**B**) *ST6GALNAC5*, (**C**) *LHX8*, (**D**) *ZNF-582* and (**E**) *ZIC1* in vulvar tissue specimens. Log_2-transformed DNA methylation levels relative to reference gene *ACTB* (y-axis) are depicted for normal vulvar epithelium (green), VIN without (yellow) and with (orange) VSCC and VSCCs (red). DNA methylation levels are significantly increased in VIN at risk of cancer (VIN with VSCC; orange) and VSCC (red) compared to VIN without cancer (yellow) and normal vulvar epithelium (green). VIN: vulvar intraepithelial neoplasia; VSCC: vulvar squamous cell carcinoma.

Figure 10 shows the boxplots of the methylation levels of : (**A**) *ASCL1*, (**B**) *ST6GALNAC5*, (**C**) *ZNF-582* and (**D**) *ZIC1 in anal tissue specimens.*. Log_2-transformed DNA methylation levels relative to reference gene *ACTB* (y-axis) are depicted for the different histological categories of anal tissue specimens from HIV-positive men (X-axis). Normal and AIN1 (green), AIN2 and AIN3 (orange) and cancer (SCC; red) * and ∘: extreme and mild outlier samples, respectively.. Methylation levels of all 4 marker increase significantly with severity of anal disease. **: $p < 0.01$, ***: $p < 0.001$, NS: not significant.

DETAILED DESCRIPTION OF THE INVENTION

The term "HPV-induced invasive cancer" refers to a carcinoma induced by high-risk HPV, which invades surrounding tissue. This includes all HPV-induced carcinoma histotypes, i.e., squamous cell carcinomas, adenocarcinomas, adenosquamous carcinomas and neuroendocrine carcinomas in relevant organs such as cervix, oral cavity, oropharynx, anus, rectum, penis, vulva, vagina, etc. It especially

includes head and neck squamous cell carcinomas (HNSCC), cervical squamous cell carcinomas and cervical adenocarcinomas.

The term "invasive cervical cancer" refers to a cervical carcinoma invading surrounding tissue. This includes all carcinoma histotypes, i.e., squamous cell carcinomas, adenocarcinomas, adenosquamous cell carcinomas and neuroendocrine carcinomas.

The term "nonHPV-induced gynaecological or anogenital cancer" refers to endometrial cancer, ovarian cancer, vulvar cancer, vaginal cancer, anal cancer and penile cancer that are HPV-negative.

The terms "premalignant lesion" and "precursor lesion" refer to a stage in the multistep cellular evolution to cancer with a strongly increased chance to progress to a carcinoma. With classical morphology the pathologist is unable to predict in the individual patient which of these lesions will progress or regress. The current patent refers to a method, which can predict invasive cancer or a high-grade precursor lesion thereof.

The term "high-grade premalignant cervical lesion" refers to a stage in the multistep cellular evolution to cervical cancer with a strongly increased chance to progress to a cervical carcinoma. The term "capable of specifically hybridizing to" refers to a nucleic acid sequence capable of specific base-pairing with a complementary nucleic acid sequence and binding thereto to form a nucleic acid duplex.

A "complement" or "complementary sequence" is a sequence of nucleotides which forms a hydrogen-bonded duplex with another sequence of nucleotides according to Watson-Crick base-paring rules. For example, the complementary base sequence for 5'-AAGGCT-3' is 3'-TTCCGA-5'.

The term "stringent hybridization conditions" refers to hybridization conditions that affect the stability of hybrids, e.g., temperature, salt concentration, pH, formamide concentration and the like. These conditions are empirically optimised to maximize specific binding and minimize non-specific binding of the primer or the probe to its target nucleic acid sequence. The terms as used include reference to conditions under which a probe or primer will hybridise to its target sequence, to a detectably greater degree than other sequences (e.g. at least 2-fold over background). Stringent conditions are sequence dependent and will be different in different circumstances. Longer sequences hybridise specifically at higher temperatures. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for

the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of a complementary target sequence hybridises to a perfectly matched probe or primer. Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes or primers (e.g. 10 to 50 nucleotides) and at least about 60°C for long probes or primers (e.g. greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. Exemplary low stringent conditions or "conditions of reduced stringency" include hybridization with a buffer solution of 30% formamide, 1 M NaCl, 1% SDS at 37°C and a wash in 2x SSC at 40°C. Exemplary high stringency conditions include hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.1x SSC at 60°C. Hybridization procedures are well known in the art and are described in e.g. Ausubel et al, Current Protocols in Molecular Biology, John Wiley & Sons Inc., 1994.

The term "oligonucleotide" refers to a short sequence of nucleotide monomers (usually 6 to 100 nucleotides) joined by phosphorous linkages (e.g., phosphodiester, alkyl and aryl-phosphate, phosphorothioate), or non-phosphorous linkages (e.g., peptide, sulfamate and others). An oligonucleotide may contain modified nucleotides having modified bases (e.g., 5-methyl cytosine) and modified sugar groups (e.g., 2'-O-methyl ribosyl, 2'-O-methoxyethyl ribosyl, 2'-fluoro ribosyl, 2'-amino ribosyl, and the like). Oligonucleotides may be naturally-occurring or synthetic molecules of double- and single-stranded DNA and double- and single-stranded RNA with circular, branched or linear shapes and optionally including domains capable of forming stable secondary structures (e.g., stem-and-loop and loop-stem-loop structures).

The term "primer" as used herein refers to an oligonucleotide which is capable of annealing to the amplification target allowing a DNA polymerase to attach thereby serving as a point of initiation of DNA synthesis when placed under conditions in which synthesis of primer extension product which is complementary to a nucleic acid strand is induced, i.e., in the presence of nucleotides and an agent for polymerization such as DNA polymerase and at a suitable temperature and pH. The (amplification) primer is preferably single stranded for maximum efficiency in amplification. Preferably, the primer is an oligodeoxy ribonucleotide. The primer must be sufficiently long to prime the synthesis of extension products in the presence of the

agent for polymerization. The exact lengths of the primers will depend on many factors, including temperature and source of primer. A "pair of bi-directional primers" as used herein refers to one forward and one reverse primer as commonly used in the art of DNA amplification such as in polymerase chain reaction (PCR) amplification.

The term "probe" refers to a single-stranded oligonucleotide sequence that will recognize and form a hydrogen-bonded duplex with a complementary sequence in a target nucleic acid sequence analyte or its cDNA derivative.

DNA methylation is a biochemical process that is important for normal development in higher organisms. It involves the addition of a methyl group to the 5 position of the cytosine pyrimidine ring or the number 6 nitrogen of the adenine purine ring. DNA methylation at the 5 position of cytosine has the specific effect of reducing gene expression and has been found in every vertebrate examined. In adult somatic tissues, DNA methylation typically occurs in a CpG dinucleotide context.

Using a genome wide DNA methylation screen on self-collected cervico-vaginal specimens and extensive evaluation of a large series of genes that are targeted by methylation in self-collected cervico-vaginal specimens of women with CIN3 and cervical cancer it has now been found that a methylation classifier consisting of the genes encoding achaete-scute family bHLH transcription factor 1 (further referred to as *ASCL1;* Genbank Accession NM_004316), the gene encoding LIM homeobox 8 (further referred to as *LHX8;* Genbank Accession NM_001001933), and the gene encoding ST6 N-Acetylgalactosaminide Alpha-2,6-Sialyltransferase 5 (further referred to as *ST6GALNAC5*; Genbank Accession NM_030965) and their regulatory sequences are important determinants of hr-HPV induced carcinogenesis. The genes *LHX8, ASCL1 and ST6GALNAC5* and their regulatory sequences thus provide valuable markers to diagnose invasive cervical cancer and the high-grade precursor lesions thereof in different sampling types, i.e. cervical scrapes, self-collected cervico-vaginal specimens, as well as urine specimens. Additionally, the present invention is suited to diagnose non-cervical hrHPV-associated invasive cancers and their high-grade precursor lesions. Moreover, their methylation is additionally suited to diagnose endometrial and other nonHPV- induced gynaecological and anogenital cancers.

Further genes that have been found useful for detection of these kinds of hrHPV associated cancers and precancerous lesions are ZIC1 and ZNF-582.

The ZIC1 gene encodes a 48 kDa protein that functions as a transcription factor and is a member of the ZIC family of C2H2-type zinc finger proteins. Members of this family are important during development. ZIC1 is involved in neurogenesis. It plays important roles in the early stage of organogenesis of the CNS, as well as during dorsal spinal cord development and maturation of the cerebellum (reviewed by Grinberg and Millen, Clin Genet. 2005, 67(4):290-6). ZIC1 hypermethylation has been described in colorectal, gastric, ovarian and hepatocellular cancer (Gan et al., PLoS One. 2011, 6(2):e16916; Wang et al., Biochem Biophys Res Commun. 2009,379(4):959-63; Huang et al., Epigenetics. 2013, 8(6):624-34; Wang et al., Tumour Biol. 2014,35(8):7429-33). It has been mentioned as one of the factors that was found in a screen of hypermethylated genes in carcinoma in situ and cancer in a cervical swab (Wang et al., Cancer Med. 2015, 4(1):43-55) but only as one out of more than 2200 genes. Further, in our earlier application WO 2017/034407 it has been described that *ZIC1* and *GHSR* promoter methylation are important determinants of hr-HPV induced carcinogenesis. The *ZIC1* and *GHSR* genomic and regulatory sequences thus provide valuable markers to diagnose invasive cervical cancer and the high-grade precursor lesions thereof.

ZNF-582 is a member of the KRAB-ZNF family and these family members function as corepressors of gene transcription in biological processes related to DNA damage response, proliferation, cell cycle control and neoplastic transformation (Urrutia 2003; Genome Biol. 4: 231). This marker is found to be methylated in acute myeloid leukemia (Figueroa et al. 2010; Cancer Cell. 17: 13-27) and has also been described in cervical cancer (Huang et al. 2012; PLOS one. 7: e41060).

Cervical cancer is almost exclusively associated with human papillomavirus (HPV) infection. Human papillomaviruses, constitute a group of more than 150 types of viruses, as identified by variations in DNA sequence. The various HPVs cause a variety of cutaneous and mucosal diseases. HPVs are broadly classified into low-risk and high-risk types, based on their ability to induce malignant changes in infected cells. Low risk HPV types such as 1, 2, 4, 6, 11, 13 and 32 are primarily associated with benign lesions or common warts, while the high risk types, such as 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, and 68 are primarily associated with premalignant and malignant epithelial lesions. The high-risk HPV types have been found to cause

invasive carcinoma of the uterine cervix, as well as invasive carcinoma elsewhere in the anogenital tract and/or head-neck region. Therefore, the present invention is not only suited to detect invasive cervical cancer and precursor stages thereof , but also other invasive cancers and corresponding precursor stages that are induced by HPV, particularly of the high-risk type. Thus, the present invention provides a method for the risk assessment of any HPV-induced high-grade premalignant lesion or invasive cancer.

Very suitable HPV-induced precursor lesions and invasive cancers in the context of the present invention are cervical precancerous lesions and invasive cervical cancers, but also precursor lesions and invasive cancers induced by high-risk HPV in other tissues such as oral cavity, oropharynx, anus, rectum, penis, vulva, vagina, etc.

As stated above, the methylation classifier consisting of methylation markers ASCL1, LHX8, and ST6GALNAC5 are also capable of detecting nonHPV-induced precursor lesions and invasive cancers. In the context of the present invention such cancers preferably are endometrial cancer, ovarian cancer, vulvar cancer, vaginal cancer, anal cancer and penile cancer .

A test cell may be a (pre)neoplastic cell, a proliferating cervical cell, or any other cell wherein the presence of an HPV-induced precursor lesion with invasive potential, HPV-induced invasive cancer, nonHPV- induced gynaecological and anogenital cancer is to be detected.

The ASCL1 marker is a proneural transcription factor and functions as a main regulator at the onset of differentiation in neurogenesis (Vasconcelos et al. 2014; Front. Cell. Neurosci. 8: 412). ASCL1 is found to be methylated in oral and colorectal cancer (Jin et al. 2009; Cancer Res. 69:7412–21; Li et al. 2015; Epigenetics 10:229–36). The LHX8 marker functions as a highly conserved transcription factor to regulate cell fate in neurogenesis, tooth morphogenesis and oogenesis (Zhou et al.2015; FASEB J.; 29:4083–91). Methylation of LHX8 has been described for cervical, breast, and colorectal cancer (Tommasi et al. 2009; Breast cancer Res. 11: R14; Øster, B. et al. 2011; Int. J. Cancer 129:2855–66; Farkas et al. 2013; Epigenetics 8:1213–25; Boers et al. 2016; Clin. Epigenetics 8:29).

The ST6GALNAC5 marker is a transmembrane sialyltransferase involved in the biosynthesis of gangliosides on the cell surface. (Drolez et al. 2016; Int. J. Mol. Sci. 17:1309). Methylation of ST6GALNAC5 has been described for cervical, breast, and colorectal cancer (Tommasi et al. 2009; Breast cancer Res. 11: R14; Øster, B. et al.

2011; Int. J. Cancer 129:2855–66; Farkas et al. 2013; Epigenetics 8:1213–25; Boers et al. 2016; Clin. Epigenetics 8:29).

Although individual markers have been described as methylated in cervical cancer, their diagnostic value to detect cervical cancer and CIN3 upon testing cervical material obtained by any sampling method, collected using any self-sampling device, urine and cervical scrapes is novel. More importantly, the three genes surprisingly are complementary to each other in terms of cervical cancer and CIN3 detection across all sample types. It is found that the ST6GALNAC5 marker which has an intermediate diagnostic performance enables the detection of individual lesions that are not detected by the markers LHX8 and/or ASLC1. Also the latter two show complementarity to each other as found upon logistic regression analysis.

The present inventors have now established that detection of a methylation classifier based a combination of *LHX8, ASCL1* and *ST6GALNAC5* promoter methylation is a frequent event in cervical carcinomas of both squamous cell carcinoma, adeno-squamous carcinoma, adenocarcinoma and neuroendocrine carcinoma histotypes, and their high-grade precursor lesions. Most interestingly, the present inventors have shown that methylation classifier detecting hypermethylation of the genes LHX8, ASCL1 and ST6GALNAC5 and their regulatory sequences can be detected in both lavage-collected and brush collected cervico-vaginal self-samples but also in cervical scrapes collected by physicians and that this feature is able to predict the presence of a high-grade CIN lesion or invasive carcinoma. In addition, hypermethylation of the genes LHX8, ASCL1 and ST6GALNAC5 and their regulatory sequences can also be detected in urine specimens collected by self-sampling and LHX8, ASCL1 and ST6GALNAC5 methylation was found to be associated with the presence of an underlying high-grade CIN lesion or invasive cervical cancer.

Hypermethylation of the genes LHX8, ASCL1 and ST6GALNAC5 and their regulatory sequences can also be detected in HPV-positive and HPV-negative vulvar and anal cancers and their high-risk precursor lesions.

Moreover, detection of hypermethylation of the genes LHX8, ASCL1 and ST6GALNAC5 and their regulatory sequencesis suited to diagnose endometrial, ovarian and other nonHPV- induced gynaecological and anogenital cancers.

Accordingly, the present invention provides a method for detecting HPV-induced high-grade precancerous lesions and HPV-induced invasive cancers, nonHPV-

induced gynaecological and anogenital cancers such as endometrial and ovarian carcinoma, said method comprising detection of hypermethylation of the genes LHX8, ASCL1 and ST6GALNAC5 and their regulatory sequences in a cell whereby such hypermethylation indicates the presence of HPV-induced precursor lesions with invasive potential and HPV-induced invasive cancers and nonHPV- induced gynaecological and anogenital cancers such as endometrial carcinoma.

The test cell of the subject may comprise a cell from a sample of mucosal cells, such as cervical cells, and also other tissue such as oral cavity, oropharynx, penis, vulva, anus, rectum, endometrium, ovarium and other tissues wherein a precursor lesion or cancer associated with HPV or nonHPV- induced gynaecological, anogenital cancer or oropharyngeal cancer is to be detected. All such samples may be used as a sample in a method of the present invention. Preferably, a sample of a patient's cells comprise cervical cells or other epithelial cells of the anogenital or oropharyngeal tract as test cells. The cervical cells may e.g. be presented as a histological or cytological specimen. Cytological specimens comprise conventional cervical smears as well as thin layer preparations of cervical specimens and cervico-vaginal or vaginal specimens collected by self-sampling. Alternatively, cells may be presented in urine samples. A test cell wherein the present invention is especially advantageous over other known methods of detecting cancers in the cervix and adjacent tissues is a test cell obtained from a self-sample.

A method of the present invention is particularly suited for the detection of high-grade precancerous lesions and invasive cancers associated with LHX8, ASCL1 and ST6GALNAC5 that are induced by high-risk HPVs or derived from the (female) anogenital tract. A method of detecting HPV-induced high-grade precancerous lesions with invasive potential, HPV-induced invasive cancers and nonHPV- induced gynaecological and anogenital cancer may comprise measuring LHX8, ASCL1 and ST6GALNAC5 promoter.

Figure 1 shows the CpG-rich promoter region of the ASCL1 gene as well as part of exon 1 of the coding sequence.

Figure 2 shows the CpG-rich promoter region of the LHX8 gene as well as part of exon 1 the coding sequence.

Figure 3 shows the CpG-rich promoter region of the ST6GALNAC5 gene as well as exon 1, intron 1 and part of exon 2 of the coding sequence.

The invention also comprises a method for detecting vulva cancer and vulvar HPV-induced precursor lesions with invasive potential, said method comprising detection of a methylation classifier consisting of ASCL1 and one or more of ZIC1, ZNF-582 and LHX8 sequences in a cell whereby such hypermethylation indicates the presence of vulva cancer and vulvar HPV-induced precursor lesions with invasive potential. As has been shown in the experimental part of the present invention the combination of ASCL1 with one or more of the mentioned gens forms a very useful methylation test panel for detecting these kind or cancers/lesions. Thus, also included is the use of the methylation classifier ASCL1, and one of ZIC1, ZNF-582 and LHX8 as a molecular diagnostic marker for the detection of vulva cancer or vulvar HPV-induced high-grade precancerous lesion, preferably wherein the methylation of said marker is predictive for the occurrence of said lesion or cancer.

Further, it has been demonstrated in the experimental part of the present invention that detection of a methylation classifier consisting of ASCL1 and one or more of ZIC1 and ZNF-582 sequences in a cell can be used to detect anal cancer and anal HPV-induced precursor lesions with invasive potential. In such a case, hypermethylation indicates the presence of anal cancer and anal HPV-induced precursor lesions with invasive potential. Thgus also part of the invention is formed by the use of the methylation classifier ASCL1, and one of ZIC1 and ZNF-582 as a molecular diagnostic marker for the detection of anal cancer or anal HPV-induced high-grade precancerous lesion, preferably wherein the methylation of said marker is predictive for the occurrence of said lesion or cancer.

Next, it has been shown that detection of a methylation classifier consisting of ASCL1, and one or more of LHX8 and ZIC1 sequences in a cell is very suitable for the detection of ovarian cancer and ovarian precursor lesions, whereby hypermethylation of the mentioned genes indicates the presence of ovarian cancer and ovarian precursor lesions. Accordingly, also the use of the methylation classifier ASCL1, and one of LHX8 and ZIC1 as a molecular diagnostic marker for the detection of ovarian cancer, preferably wherein the methylation of said marker is predictive for the occurrence of said cancer, is envisaged in the present invention.

Detection of methylation is performed on nucleic acid, such as DNA. The reagents that are used are typically a nucleic acid (DNA) probe or (PCR) primer or a

restriction endonuclease, preferably a methylation sensitive restriction endonuclease for the detection of the presence of methyl groups on the test cell DNA.

The test cell component may be detected directly *in situ* or it may be isolated from other cell components by common methods known to those of skill in the art before contacting with the reagent (see for example, "Current Protocols in Molecular Biology", Ausubel et al. 1995. 4th edition, John Wiley and Sons; "A Laboratory Guide to RNA: Isolation, analysis, and synthesis", Krieg (ed.), 1996, Wiley-Liss; "Molecular Cloning: A laboratory manual", J. Sambrook, E.F. Fritsch. 1989. 3 Vols, 2nd edition, Cold Spring Harbor Laboratory Press)

Since examples presented show frequent methylation of the ASCL1 gene and regulatory sequence, it is desirable to directly determine whether the *ASCL1* gene and regulatory sequenceis hypermethylated. Similarly, it is also desirable to directly determine whether the LHX8 and ST6GALNAC5 gene and regulatory sequenceis hypermethylated. In particular, the cytosine rich areas termed "CpG islands", which are primarily situated in the 5' regulatory regions of genes are normally unmethylated. The term "hypermethylation" includes any methylation of cytosine at a position that is normally unmethylated in the LHX8, ASCL1 and ST6GALNAC5 gene and regulatory sequence (e. g. the LHX8 and ASCL1 regulatory sequence and first exon and ST6GALNAC5 regulatory sequence, first exon, first intron and part of second exon , see Figures 1, 2 and 3, respectively). DNA methylation can be detected by the following assays currently used in scientific research:

- Methylation-Specific PCR (MSP), which is based on a chemical reaction of sodium bisulfite with DNA that converts unmethylated cytosines of CpG dinucleotides to uracil or UpG, followed by traditional PCR. However, methylated cytosines will not be converted in this process, and primers are designed to overlap the CpG site of interest, which allows one to determine methylation status as methylated or unmethylated.
- Whole genome bisulfite sequencing, also known as BS-Seq, which is a high-throughput genome-wide analysis of DNA methylation. It is based on aforementioned sodium bisulfite conversion of genomic DNA, which is then sequenced on a Next-generation sequencing platform. The sequences obtained are then re-aligned to the reference genome to determine methylation states of CpG dinucleotides based on mismatches resulting from the conversion of unmethylated cytosines into uracil.

- The HELP assay, which is based on restriction enzymes' differential ability to recognize and cleave methylated and unmethylated CpG DNA sites.
- ChIP-on-chip assays, which is based on the ability of commercially prepared antibodies to bind to DNA methylation-associated proteins like MeCP2.
- Restriction landmark genomic scanning, a complicated and now rarely-used assay based upon restriction enzymes' differential recognition of methylated and unmethylated CpG sites; the assay is similar in concept to the HELP assay.
- Methylated DNA immunoprecipitation (MeDIP), analogous to chromatin immunoprecipitation, immunoprecipitation is used to isolate methylated DNA fragments for input into DNA detection methods such as DNA microarrays (MeDIP-chip) or DNA sequencing (MeDIP-seq).
- Pyrosequencing of bisulfite treated DNA. This is sequencing of an amplicon made by a normal forward primer but a biatenylated reverse primer to PCR the gene of choice. The Pyrosequencer then analyses the sample by denaturing the DNA and adding one nucleotide at a time to the mix according to a sequence given by the user. If there is a mis-match, it is recorded and the percentage of DNA for which the mis-match is present is noted. This gives the user a percentage methylation per CpG island.
- Molecular break light assay for DNA adenine methyltransferase activity – an assay that relies on the specificity of the restriction enzyme DpnI for fully methylated (adenine methylation) GATC sites in an oligonucleotide labeled with a fluorophore and quencher. The adenine methyltransferase methylates the oligonucleotide making it a substrate for DpnI. Cutting of the oligonucleotide by DpnI gives rise to a fluorescence increase.
- Methyl Sensitive Southern Blotting is similar to the HELP assay, although uses Southern blotting techniques to probe gene-specific differences in methylation using restriction digests. This technique is used to evaluate local methylation near the binding site for the probe.
- Quantum-dot based methylation assay – an assay as described in Bailey, V. et al. (Genome Res. 19:1455-1461, 2009) in which the high specificity of MSP and the high sensitivity and simplicity of the quantum dot FRET (QD-FRET) technology (Zhang, C. et al., 2005, Nat. Mater. 4:826-831) is combined.
- DNA methylation detection using nanochip technology. This technique is able to detect DNA methylation at high sensitivity and specificity in minimal

amounts of clinical material, without the need for bisulfite conversion and PCR amplification. Methods using solid states nanopores have been described by Shim, J. et al. (Sci. Rep. 3:1389, 2013). A device for lab on a chip technology is described in patent publication WO2009104367 (A1).

Hypermethylation preferably can be detected by methylation specific PCR, which is based on bisulfite modification of DNA, followed by specific PCR reactions that target CpG rich sequences.

An alternative preferred means to test for methylated sequences is by next generation sequencing of bisulfite modified DNA.

A third preferred means to detect methylated sequences and to discriminate between methylated and unmethylated DNA is based on nanotechnology.

For purposes of the invention a nucleic acid probe specific for LHX8, ASCL1 and/or ST6GALNAC5 may be used to detect the presence of LHX8, ASCL1 and/or ST6GALNAC5 polynucleotide (using nucleic acid probe) in biological fluids or tissues. Oligonucleotide primers based on any coding sequence region and regulatory sequence region in the LHX8, ASCL1 and/or ST6GALNAC5 sequence are useful for amplifying DNA, for example by PCR.

When using PCR primers, nucleic acid probes or restriction endonucleases, the 5' regulatory region and coding sequence of the LHX8 and, ASCL1 genes or the 5' regulatory region, coding sequence and first intron of ST6GALNAC5 sequence (as specified in Figures 1, 2 and 3 respectively) is analysed.

Any specimen containing a detectable amount of LHX8, ASCL1 and ST6GALNAC5 polynucleotide can be used. Preferred samples for testing according to methods of the invention include such specimens as (cervical or vaginal) scrapes, cervico-vaginal lavages or swabs, urine, blood and/or (cervical) biopsies and the like. Although the subject can be any mammal, preferably the subject is human.

Diagnostic methods for the detection of disorders, include methods wherein a sample for testing is provided, which sample comprises a cell preparation from cervical or other tissue. Preferably such samples are provided as smears or other cytological samples. Additional suitable samples include urine and blood.

A cell or tissue sample obtained from a mammal, preferably a human, is suitably pretreated to allow contact between the cellular DNA of a test cell comprised in said sample with a reagent that detects LHX8, ASCL1 and ST6GALNAC5 and

detects an alteration in the methylation of the genes LHX8, ASCL1 and ST6GALNAC5 or their regulatory sequence as compared to that of a comparable normal cell. Samples may be mounted on a suitable support to allow observation of individual cells. Examples of well-known support materials include glass, polystyrene, polypropylene, polyethylene, polycarbonate, polyurethane, optionally provided with layers to improve cell adhesion and immobilization of the sample, such as layers of poly-L-lysine or silane. Cervical smears or biopsies may for instance be prepared as for the Papanicolaou (Pap) test or any suitable modification thereof as known by the skilled person, and may be fixed by procedures that allow proper access of the reagent to the target component. In certain embodiments of the invention the cytological specimens are provided as conventional smear samples or thin layer preparations of cervical cells or liquid based cytology samples or any other kind of preparation known to those of skill in the art. If storage is required, routine procedures use buffered formalin for fixation followed by paraffin embedding, which provides for a well-preserved tissue infrastructure.

In one embodiment of a method of the invention an increased methylation of the genes *LHX8, ASCL1* and *ST6GALNAC5* or their regulatory sequence in the test cell is detected as compared to the comparable normal cell.

The present invention also provides a kit of parts as defined in the claims, for use in a method of detecting HPV-induced precursor lesions with invasive potential, HPV-induced invasive cancers and nonHPV- induced gynaecological and anogenital cancers. Such a kit may suitably comprise a brush or spatula to take a (cervical) scrape either or not together with a container filled with collection medium to collect test cells. Alternatively, a sampling device consisting of an irrigation syringe, a disposable female urine catheter and a container with irrigation fluid will be included to collect cervical cells by cervico-vaginal lavage. Additionally or alternatively,, a container to collect urine is suitable, preferably to be used to collect first-void urine. A kit according to the present invention further comprises primers and probes for the detection of LHX8, ASCL1 and ST6GALNAC5 gene or regulatory sequences .

A kit of parts according to the invention further comprises means for the detection of methylation of LHX8, ASCL1 and ST6GALNAC5 gene or regulatory sequence, such as, methylation-sensitive restriction enzymes, or probes or primers capable of hybridising to the nucleotide sequence of Fig. 1 and/or Fig. 2 and/or Figure 3.

In yet another alternative embodiment of a kit of the invention the means for the detection of methylation of LHX8, ASCL1 and ST6GALNAC5 gene or regulatory sequence may be combined with means for the detection of HPV infection, preferably for the detection of HPV infection of the high-risk type. Such means may comprise HPV-specific primers or probes, protein markers for HPV infection or even surrogate markers for HPV infection as are known in the art.

The present invention will now be illustrated by way of the following, non limiting examples.

EXAMPLES

Example 1. Discovery of LHX8, ASCL1 and ST6GALNAC5 as optimal methylation marker panel to detect cervical cancer and precancer in self-samples

A comprehensive analysis of genome-wide DNA methylation changes detectable in self-samples and associated with cervical carcinogenesis has been conducted by means of the Infinium 450K BeadChip array on hrHPV-positive self-samples. The sample series included consisted of 68 hrHPV-positive self-samples of 39 women with high-grade cervical intraepithelial neoplasia grade 3 (CIN3) and 29 women with low-grade cervical intraepithelial neoplasia grade 0 or 1 (≤CIN1) The Infinium HumanMethylation450 BeadChip array analyses over 485,000 methylation sites per sample at single nucleotide resolution of the human genome (Illumina, San Diego, CA, USA). By this approach we identified 12 methylation targets that were specifically associated with the presence of CIN3. Next, the 12 most discriminative methylated genes were analysed using multiplex qMSP assays in two separate large series consisting of either hrHPV-positive lavage self-samples (n=245) or hrHPV-positive brush self-samples (n=246) from women with and without CIN3. In both lavage and brush self-samples, 11 of 12 selected genes showed significantly increased methylation levels ($p < 0.0005$) in self-samples from women with CIN3 compared to hrHPV-positive controls (see also Figure 4).

To subsequently build an optimal classifier of methylation markers which is universally applicable to any self-sample type, logistic regression analysis was performed on both lavage and brush self-sample qMSP datasets. Following backward elimination, the optimal methylation classifiers in both self-sample types were

compared. The reason for testing both self-sampler types in order to find a methylation classifier that is universally applicable to self-collected cervico-vaginal specimens, independent of the collection device used, is based on previous research findings indicating that the clinical performance of methylation markers depends on the type of self-sampler used. For example it was found that methylation markers MAL and mir124-2 show a good clinical performance for the detection of CIN3+ in lavage self-samples, but not in brush-based self-samples. Surprisingly, in this study an optimal methylation classifier consisting of the genes *ASCL1*, *LHX8* and *ST6GALNAC5* and their regulatory sequence, was found for both self-sample types. Methylation levels of *ASCL1*, *LHX8* and *ST6GALNAC5* as detected in lavage and brush self-samples are shown in Figure 4A and 4B, respectively. Methylation levels of all 3 genes were significantly increased in women with CIN3 and extremely high in women with cervical cancer. The methylation classifier showed an excellent clinical performance for CIN3 detection in both hrHPV-positive lavage (AUC=0.90) and brush (AUC=0.86) self-samples. The methylation classifier scored 83% (25 of 30) of lavage self-samples and 76% (52 of 68) of brush self-samples from women with CIN3 methylation-positive with a corresponding specificity in hrHPV-positive controls of 80%. Importantly, virtually all women with cervical cancer tested positive for the methylation classifier. The clinical performance for detection of CIN3+ (CIN3 and cancer) is superior to previously published methylation marker panels. The methylation classifier has a 79%-89% sensitivity and 75%-77% specificity for CIN3+ in lavage and brush self-samples. In comparison, the best currently available methylation panel tested on large self-sample series is *FAM19A4/miR124-2,* which has a sensitivity of 69-71% and a specificity of 68-76% for CIN3+ detection in both lavage and brush self-samples (De Strooper *et al.* *Gynecol. Oncol.* **141,** 341–347 (2016)). Other marker panels, such as *JAM3/EPB41L3/TERT/C13ORF18*, have only been tested on small selected series of self-samples (Boers. *et al. Br. J. Cancer* **111,** 1095–101 (2014).;Eijsink *et al. Gynecol. Oncol.* **120,** 280–283 (2011).)

Primers and probes used for LHX8, ASCL1 and ST6GALNAC5 qMSP analysis are listed in Table 1. The housekeeping gene ß-actin (ACTB) was chosen as a reference for total DNA input measurement. Quantification was performed using the comparative Ct method (Schmittgen et al., Nat Protoc 2008, 3:1101-1108).

Example 2 Validation of methylation classifier in lavage and brush self-samples

To validate the clinical performance of the methylation classifier, we analysed another, independent, large series of hrHPV-positive lavage self-samples (n=198) and brush self-samples (n=278) using a multiplex qMSP assay. The methylation classifier showed a good and comparable clinical performance for CIN3 detection as observed in the above described sample series, both for hrHPV-positive lavage (AUC=0.88) and brush (AUC=0.90) self-samples (see Figure 4C for ROC curves). In the validation set, 74% (26 of 35) of lavage self-samples and 88% (49 of 56) of brush self-samples from women with CIN3 showed methylation-positivity with a corresponding specificity in hrHPV-positive controls of 79% and 81%, respectively.

Example 3. LHX8, ASCL1 and ST6GALNAC5 methylation as marker for primary screening in self-collected cervico-vaginal specimens

Methylation markers tested so far are not well suitable for use in primary screening due to a too low specificity at an acceptable sensitivity for CIN2/3 and cancer. When evaluating the use the methylation classifier consisting of LHX8, ASCL1 and ST6GALNAC5 as a marker for primary screening it was surprisingly found that upon analysis of HPV-negative self-samples and self-samples of women with CIN3, the classifier also has a very high AUC of 0.895 on lavage self-samples to detect CIN3 as well as on brush self-samples (AUC is 0.828).

The present findings show that a methylation classifier detecting hypermethylation of the genes *LHX8, ASCL1* and *ST6GALNAC5* and their regulatory sequence enables the detection of underlying CIN2+ not only when applied to self-collected cervico-vaginal lavage specimens but self-collected vaginal brush samples as well. The latter is a specimen type in which previous known markers often performed with low clinical sensitivity. Consequently, the LHX8, ASCL1 and ST6GALNAC5 markers can be considered as pan-detection markers showing equal performance independent of the sampling device used. Importantly, the high AUC values and specificity rates enable primary screening by testing the methylation classifier genes, obviating the need for initial hrHPV testing. As described below (Example 6) such application also enables the detection of hrHPV-negative cancers that are missed in current screening settings.

Example 4: Detection of *LHX8, ASCL1* and *ST6GALNAC5* methylation in hrHPV-positive cervical scrapes

From women participating in a population-based screening we studied cervical scrapes of hrHPV positive women in which CIN3 (n=56) was diagnosed, and hrHPV negative and positive women in whom at maximum CIN 1 was diagnosed (n=40 and n=87, respectively). Additionally, cervical scrapes of women diagnosed with squamous cell carcinoma (SCC;n=23) and adenocarcinoma (AdCa;n=3) of the cervix were tested. Cervical scrapes of these women were collected in preservation medium in which nucleic acids are well preserved.

LHX8, ASCL1 and *ST6GALNAC5* methylation was significantly increased in scrapes of women with CIN3, SCC and AdCa compared to HPV negative and HPV positive controls. Receiver Operating Characteristics analysis showed an AUC of 0,890 to detect CIN3 in an HPV-positive population and an AUC of 1 to detect cervical cancer (Figure 5B).

Thereby these genes provide promising triage markers in screening by primary HPV testing but can also be used as primary screening markers for both HPV and non-HPV related carcinomas.

Example 5: LHX8, ASCL1 and ST6GALNAC5 methylation in urine specimens as markers for cervical cancer and their high-grade precursor lesions

A total of 44 urine samples collected from patients with cervical cancer and 47 urine from control women were tested for *LHX8, ASCL1* and *ST6GALNAC5* methylation. Compared to controls *LHX8, ASCL1* and *ST6GALNAC5* methylation levels were significantly increased in urine samples of women with cervical cancer (Figure 5A). Receiver Operating Characteristics analysis showed an AUC of 0,936 to detect cervical cancer (Figure 5C). Importantly one cancer that was not detected by LHX8 or ASCL1 was detected by ST6GALNAC5, confirming the complementarity of the markers within the classifier. These results show that detection of the methylation classifier genes allows for cervical cancer screening using urine, thereby providing the most easy and most patient friendly non-invasive cancer screening methods currently available. Additionally methylation testing can be used to triage hrHPV-positive women in urine-based HPV screening programs.

Example 6: *LHX8, ASCL1* and *ST6GALNAC5* methylation in cervical scrapes of women with endometrial carcinoma

A total of 24 cervical scrapes of women with endometrial carcinoma were tested for *LHX8, ASCL1* and *ST6GALNAC5* methylation. Compared to cervical scrapes of hrHPV negative and positive women in whom at maximum CIN 1 was diagnosed, *LHX8, ASCL1* and *ST6GALNAC5* methylation levels were significantly increased in women with endometrial carcinoma (Figure 6).

Table 1 Primer and probe sequences (5'-3') used for LHX8, ASCL1 and ST6GALNAC5 quantitative MSP analysis

	Forward primer	**Reverse primer**	**Probe**
ASCL1	CGTTTAAGTAAGTTAA GCGATAGCGTTC	CCACGACGACCGACTAC TACTACG	GTTGTAAACGTCGGTTTAA TTTTAGC
LHX8	TAGGCGTCGTGACGGT TGTAC	AAACAAATCAAAAATTC CGAACG	GTTTCGTATTTTGAGGGTT TTTATTAGTTCGTTTTTCG
ST6GALNAC 5	GACGGTGTTTTTTTTG TTTTAGTTGC	GCAACTAAAACAAAAAA AAC	TGCGGCGGGGTTGGGGTTT A

Also further primers and probes can be useful in the present invention (also in view of Examples 7-9). These are listed in Table 2:

Table 2 Primer and probe sequences (5'-3') used for ZIC1, ASCL1 , LHX8 and ZNF-582 quantitative MSP analysis

	Forward primer	**Reverse primer**	**Probe**
ZIC1	GGGCGGGTTAATGAGTTG C	TCACGTACTACCGACGC TAACG	CGCCGCGCCAACGAAAAAC
ZNF-582	TTTAAGGTCGGGTTGTTG TTTTTAC	GCACAAAACACACCGAT ACTACG	ATAATAAAACCGACGCCGCAAT ATCTTCCG
LHX8	CGTCGTGACGGTTGTAC	CTCGACGCGAAAAACG	TCAAAATACGAAACAACG
ASCL1	AGTAAGTTAAGCGATAGC GTTC	ACCGCTAAAATTAAACC GACG	TTCGTTCGAATTGATGCGTT

Example 7: LHX8, ASCL1, ST6GALNAC5, ZNF-582 and ZIC1 methylation for the detection of vulvar cancer and precancer

Vulvar intra-epithelial neoplasia (VIN) is the precursor lesion of vulvar squamous cell carcinoma (VSCC) but only a minority of VIN progresses to cancer. High-grade vulvar intraepithelial neoplasia (VIN) is the precancerous state of vulvar squamous cell carcinomas (VSCC), but only a minority of VINs progress to cancer, indicating a heterogeneous disease. Current clinical and histological classifications are insufficient to predict the cancer risk. Consequently, affected women are treated similarly with mutilating interventions. Hence, there is a clinical need for objective biomarkers reflecting the cancer risk. Therefore, we assessed the potential value of the DNA methylation marker panel LHX8, ASCL1 and ST6GALNAC5 as well as the methylation marker ZNF-582 and ZIC1 for risk stratification of VIN. The marker ZNF-582 was identified in the same genome-wide screen as described in Example 1. Figure 7 shows the ZNF582 5'regulatory region and coding sequence (exon 1). The coding sequence is in upper case and underlined is the CpG rich region.

ZIC1 was identified in a previous genome-wide screen for detection of cervical pre-cancer and cancer (Verlaat et al., Clinical Cancer Research, 2017). Figure 8 shows the ZIC1 5'regulatory region and coding sequence (exon 1). The coding sequence is in upper case and the CpG rich regions are indicated in grey.

Eight normal vulva samples (controls), 39 VIN and 75 VSCC were included in this study. Of the 39 VIN, 25 were associated with VSCC (VIN with VSCC) and 14 not (VIN without VSCC, i.e. no development of VSCC during >10 years of follow-up. This unique series of well-characterized VIN lesions with clinical follow-up enables the identification of markers that distinguish VIN lesions with a high cancer risk in need of aggressive treatment from VIN lesions that have a low cancer risk and would benefit from a more conservative, e.g. non-invasive management.

Testing for LHX8, ASCL1, ST6GALNAC5 and ZNF-582 methylation showed differential methylation between HPV-positive VIN without VSCC and VIN with VSCC of all five genes (Figure 9).

Receiver Operating Characteristic (ROC) curves for the comparison HPV- positive VIN without VSCC vs. HPV-positive VIN with VSCC based on logistic regression, yielded

AUCs of 0.83 (LHX8), 0.80 (ASCL1), 0.69 (ST6GALNAC5), 0.80 (ZNF-582) and 0.78 (ZIC1).

Analysis of marker combinations revealed an increased AUC when combining two markers: ASCL1 plus ZIC1 (AUC=0.82), ASCL1 plus ZNF-582 (AUC=0.82) and ASCL1 plus LHX8 (AUC=0,83) These data indicate that a combination of two of the methylation markers LHX8, ASCL1, ZIC1 and ZNF-582 can be used for cancer risk stratification of HPV-positive VIN and for the detection of VSCC.

Example 8: Methylation markers for the detection of anal cancer and precancer

Anal cancer is caused by high-risk HPV and preceded by precursor lesions: anal intraepithelial neoplasia (AIN; graded 1-3). Since only a minority will eventually progress to cancer, preferably only lesions with high malignant potential are treated. Unfortunately, a predictive cancer progression marker is lacking. Currently all high-grade AIN is treated, resulting in significant overtreatment. Therefore, we assessed the potential value of DNA methylation markers ASCL1, ST6GALNAC5, ZIC1 and ZNF-582 for risk stratification of AIN. The markers ZIC1 and ZNF-582 were identified by genome-wide analysis in HPV-containing samples (see Example 1 and Verlaat et al., Clinical Cancer Research, 2017).

Archival tissue samples of HIV+ men with anal squamous cell carcinoma (SCC; n=26), AIN3 (n=24), AIN2 (n=42) and men without AIN2 or worse (normal+AIN1; n=56) were analysed for DNA methylation of four genes known to display hypermethylation during HPV-induced carcinogenesis using quantitative methylation-specific PCR. Methylation levels of all 4 genes were significantly higher with increasing severity of disease (Figure 10). Logistic regression and AUC analysis was used to determine the performance of the methylation markers for the detection of anal cancer and AIN3.. The AUCs for detecting AIN3 or worse were 0.89 (ASCL1), 0.84 (ST6GALNAC5), 0.87 (ZIC1) and 0.91 (ZNF-582).

Analysis of marker combinations revealed an even further improved detection of AIN3 or worse, with an AUC of 0.90 for ASCL1 plus ZIC1 and 0.92 for ASCL1 plus ZNF-582. Importantly, both marker combinations detected all anal cancers.

These data indicate that the methylation marker combinations ASCL1 plus ZIC1 and ASCL1 plus ZNF-582 can be used to detect anal cancer and high-grade AIN at risk of

progression to cancer, which is particularly important for the clinical management of HIV-positive men having sex with men.

Example 9: *LHX8, ASCL1,* and *ZIC1* methylation in cervical scrapes of women with ovarian cancer

A total of 6 cervical scrapes of women with ovarian cancer, including both serous and clear cell cancers, were tested for *LHX8, ASCL1* and *ZIC1* methylation. Compared to cervical scrapes of women without disease (i.e. hrHPV negative and positive women in whom at maximum CIN 1 was diagnosed), *LHX8, ASCL1* and *ZIC1* methylation levels were significantly increased in women with ovarian cancer. These data indicate that a combination of any of these three markers enables the detection of ovarian cancer in cervical scrapes.

Claims

1. A method for detecting HPV-induced high-grade precancerous lesions and HPV-induced invasive cancers and nonHPV-induced gynaecological and anogenital cancers, said method comprising detection of a methylation classifier consisting of LHX8, ASCL1 and ST6GALNAC5 sequences in a cell whereby such hypermethylation indicates the presence of HPV-induced precursor lesions with invasive potential, HPV-induced invasive cancers and nonHPV-induced gynaecological and anogenital cancers.

2. Method according to claim 1, wherein said HPV-induced high-grade precancerous lesion or HPV-induced invasive carcinoma is a high-grade premalignant cervical lesion or invasive cervical cancer.

3. Method according to claim 1 or 2, wherein said HPV-induced invasive cancer is a high-risk HPV-induced invasive cancer.

4. Method according to claim 1, wherein said nonHPV-induced gynaecological cancers is an endometrial or ovarian cancer.

5. Method for detecting vulva cancer and vulvar HPV-induced precursor lesions with invasive potential, said method comprising detection of a methylation classifier consisting of ASCL1 and one or more of ZIC1, ZNF-582 and LHX8 sequences in a cell whereby such hypermethylation indicates the presence of vulva cancer and vulvar HPV-induced precursor lesions with invasive potential.

6. Method for detecting anal cancer and anal HPV-induced precursor lesions with invasive potential, said method comprising detection of a methylation classifier consisting of ASCL1 and one or more of ZIC1 and ZNF-582 sequences in a cell whereby such hypermethylation indicates the presence of anal cancer and anal HPV-induced precursor lesions with invasive potential.

7. Method for the detection of ovarian cancer and ovarian precursor lesions, said method comprising detection of a methylation classifier consisting of ASCL1, and one or more of LHX8 and ZIC1 sequences in a cell whereby such hypermethylation indicates the presence of ovarian cancer and ovarian precursor lesions.

8. Method according to any one of the preceding claims, wherein said hypermethylation is detected in the CpG rich sequences as indicated in Figures 1, 2, 3, 7 and 8.

9. A method of detecting HPV-induced high-grade precancerous lesion and HPV-induced invasive cancers and nonHPV-induced gynaecological or anogenital

cancer according to any of claims 1-4, or detection of vulva cancer and vulvar HPV-induced precursor lesions with invasive potential according to claim 5 and/or of detecting anal cancer and anal HPV-induced precursor lesions with invasive potential according to claim 6, and/or a method for the detection of ovarian cancer and ovarian precursor lesions, wherein said hypermethylation is an increased methylation of the methylation classifier LHX8, ASCL1, ZNF-582, ZIC1 and/or ST6GALNAC5 CpG rich promoter and/or gene sequences in a test cell as compared to a comparable normal cell.

10. Method according to any of claims 1-9, wherein the reagent is a methylation sensitive restriction endonuclease, chosen from the group consisting of BssHII, MspI, NotI and HpaII.

11. Method according to any of claims 1-9 wherein the method involves nanotechnology, preferably lab on a chip technology.

12. Method according to claim any of claims 1-9, wherein a methylation specific PCR is performed, which is based on bisulfite modification of DNA, followed by specific PCR reactions that target CpG rich sequences

13. Method according to claim 12, wherein the reagent is a nucleic acid probe or primer that binds to the nucleic acid as indicated in Figure 1 or 2.

14. Method according to claim 13, wherein said nucleic acid probe or primer has a detectable label.

15 Method according to claim 12, 13 or 14, wherein the nucleic acid probe has a nucleotide sequence selected from the group consisting of:

a) a polynucleotide sequence capable of hybridizing under stringent conditions to the sequence ASCL1 as set forth in Figure 1 or to the sequence LHX8 as set forth in Figure 2 or to the sequence ST6GALNAC5 as set forth in Figure 3; or to the sequence ZIC1 as set forth in Figure 8 or to the sequence ZNF-582 as set forth in Figure 7;

b) a polynucleotide having at least 70% identity to one of the polynucleotides of a);

c) a polynucleotide complementary to one of the polynucleotides of a); and

d) a polynucleotide comprising at least 15 bases of one of the polynucleotides of a) or b).

16. Method according to any of claims 1 -13, wherein the methylation of the methylation classifier comprising of the LHX8, ASCL1, ZNF-582, ZIC1 and/or ST6GALNAC5 gene is determined

17. Use of the methylation classifier LHX8, ASCL1, and ST6GALNAC5 as a molecular diagnostic marker for the detection of HPV-induced high-grade precancerous lesion or HPV-induced invasive carcinoma or nonHPV-induced gynaecological or anogenital cancer, preferably wherein the methylation of said marker is predictive for the occurrence of said lesion, carcinoma or cancer.

18. Use of the methylation classifier ASCL1, and one of ZIC1, ZNF-582 and LHX8 as a molecular diagnostic marker for the detection of vulva cancer or vulvar HPV-induced high-grade precancerous lesion, preferably wherein the methylation of said marker is predictive for the occurrence of said lesion or cancer.

19. Use of the methylation classifier ASCL1, and one of ZIC1 and ZNF-582 as a molecular diagnostic marker for the detection of anal cancer or anal HPV-induced high-grade precancerous lesion, preferably wherein the methylation of said marker is predictive for the occurrence of said lesion or cancer.

20. Use of the methylation classifier ASCL1, and one of LHX8 and ZIC1 as a molecular diagnostic marker for the detection of ovarian cancer, preferably wherein the methylation of said marker is predictive for the occurrence of said cancer.

21. Kit of parts for use in a method of detecting HPV-induced high- grade precancerous lesion or HPV-induced invasive carcinoma or nonHPV-induced gynaecological or anogenital cancer, said kit comprising

- means for the detection of LHX8, ASCL1 and ST6GALNAC5 methylation wherein said means comprise probes and/or primers specific for the ASCL1 nucleotide sequence of Figure 1 and the LHX8 nucleotide sequence of Figure 2 and the ST6GALNAC5 nucleotide sequence of Figure 3;

-means for the detection of ASCL1 and ST6GALNAC5 methylation wherein said means comprise probes and/or primers specific for the ASCL1 nucleotide sequence of Figure 1 and the ST6GALNAC5 nucleotide sequence of Figure 3; or

-means for the detection of ASCL1 and LHX8 methylation wherein said means comprise probes and/or primers specific for the ASCL1 nucleotide sequence of Figure 1 and the LHX8 nucleotide sequence of Figure 2; or

- means for the detection of ASCL1 and at least one of ZIC1, ZNF-582 and LHX8 wherein said means comprise probes and/or primers specific for the ASCL1 nucleotide sequence of Figure 1 and either of the ZNF-582 nucleotide sequence of Figure 7, the ZIC1 nucleotide sequence of Figure 8 and/or the LHX8 nucleotide sequence of Figure 2;

preferably further comprising:

-means for the detection of HPV infection, wherein said means comprise probes and primers specific for HPV.

5

Figure 1: ASCL1 sequence (regulatory region and exon1) derived from Refseq NM_004316 9hg19). The coding sequence is in upper case and underlined is the CpG rich region.

>hg19_refGene_NM_004316
ttctttatgatatccgctaagctggtccggaaataatctttatggggagg
gggtggcggtaggggggcgatggtacaggggggccagaggtcatcctagggg
gacgtccctgccatatacacccacctacaggacggctcacaaccactcct
cggtgtcgcttccccgcggcccccccacacacttgctcagttatggggagc
acatcctagtttttagagctgaatgggacattagagaccatattctgtgg
ctgcagacgaggaagcgaaggctcagagaggatgccacttcgaggagcca
cagagcattgagaggacgccttgggactagaacccacgttttcacatagt
ccagcacttttttttcactgttctggacggagtccctcccccaaccatgtt
tctaaacttcaatcgtaatttgctccaatttctagggtcaccgaggaacc
cgaagagaataacagtgaggagagagagaaaacaggaaaagtcgagcccc
actccctcctcacctccacaccgttcctgtgccatttttttctgcccaaac
ccttccctgcgctttgcttcaagttcttagtagaatccaagagagcttca
ccccaagtctttccacctatacacctcaattcctagagccatttgtccct
cctgtgacgcccccccacccccttcctaaagccacccccggcagcagcccg
ccccgagcgcgccgcctgtttattcagccgggagtccggcacgcgccagg
cgcacgcactgcaacaacaaacccagctgaatggagagtttgcaaggagc
gggagaaaggaacgggagggggggagaggagaggaggagggggagtttag
ggagtgggtgggaggaagaggtaagaggaggggggggagtgggggctgca
gccgctcgctgcagcagcgggggagtgggggggcgaggcgggggccagggctg
cgcgtgggggctgggtgtcccattgaaaaggcggacgcactccggcagccc
AGCACTCTCTCACTTCTGGCCAGGGAACGTGGAAGGCGCACCGACAGGGA
TCCGGCCAGGGAGGGCGAGTGAAAGAAGGAAATCAGAAAGGAAGGGAGTT
AACAAAATAATAAAAACAGCCTGAGCCACGGCTGGAGAGACCGAGACCCG
GCGCAAGAGAGCGCAGCCTTAGTAGGAGAGGAACGCGAGACGCGGCAGAG
CGCGTTCAGCACTGACTTTTGCTGCTGCTTCTGCTTTTTTTTTTCTTAGA
AACAAGAAGGCGCCAGCGGCAGCCTCACACGCGAGCGCCACGCGAGGCTC
CCGAAGCCAACCCGCGAAGGGAGGAGGGGAGGGAGGAGGAGGCGGCGTGC
AGGGAGGAGAAAAAGCATTTTCACTTTTTTTGCTCCCACTCTAAGAAGTC
TCCCGGGGATTTTGTATATATTTTTTAACTTCCGTCAGGGCTCCCGCTTC
ATATTTCCTTTTCTTTCCCTCTCTGTTCCTGCACCCAAGTTCTCTCTGTG
TCCCCCTCGCGGGCCCCGCACCTCGCGTCCCGGATCGCTCTGATTCCGCG
ACTCCTTGGCCGCCGCTGCGCATGGAAAGCTCTGCCAAGATGGAGAGCGG
CGGCGCCGGCCAGCAGCCCCAGCCGCAGCCCCAGCAGCCCTTCCTGCCGC
CCGCAGCCTGTTTCTTTGCCACGGCCGCAGCCGCGGCGGCCGCAGCCGCC
GCAGCGGCAGCGCAGAGCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA
GCAGCAGGCGCCGCAGCTGAGACCGGCGGCCGACGGCCAGCCCTCAGGGG
GCGGTCACAAGTCAGCGCCCAAGCAAGTCAAGCGACAGCGCTCGTCTTCG
CCCGAACTGATGCGCTGCAAACGCCGGCTCAACTTCAGCGGCTTTGGCTA
CAGCCTGCCGCAGCAGCAGCCGGCCGCCGTGGCGCGCCGCAACGAGCGCG
AGCGCAACCGCGTCAAGTTGGTCAACCTGGGCTTTGCCACCCTTCGGGAG
CACGTCCCCAACGGCGCGGCCAACAAGAAGATGAGTAAGGTGGAGACACT
GCGCTCGGCGGTCGAGTACATCCGCGCGCTGCAGCAGCTGCTGGACGAGC
ATGACGCGGTGAGCGCCGCCTTCCAGGCAGGCGTCCTGTCGCCCACCATC
TCCCCCAACTACTCCAACGACTTGAACTCCATGGCCGGCTCGCCGGTCTC
ATCCTACTCGTCGGACGAGGGCTCTTACGACCCGCTCAGCCCCGAGGAGC
AGGAGCTTCTCGACTTCACCAACTGGTTCTGAGGGGCTCGGCCTGGTCAG
GCCCTGGTGCGAATGGACTTTGGAAGCAG

Figure 2: LHX8 sequence (regulatory region and exon1) derived from Refseq NM_001001933 (hg19). The coding sequence is in upper case and underlined is the CpG rich region.

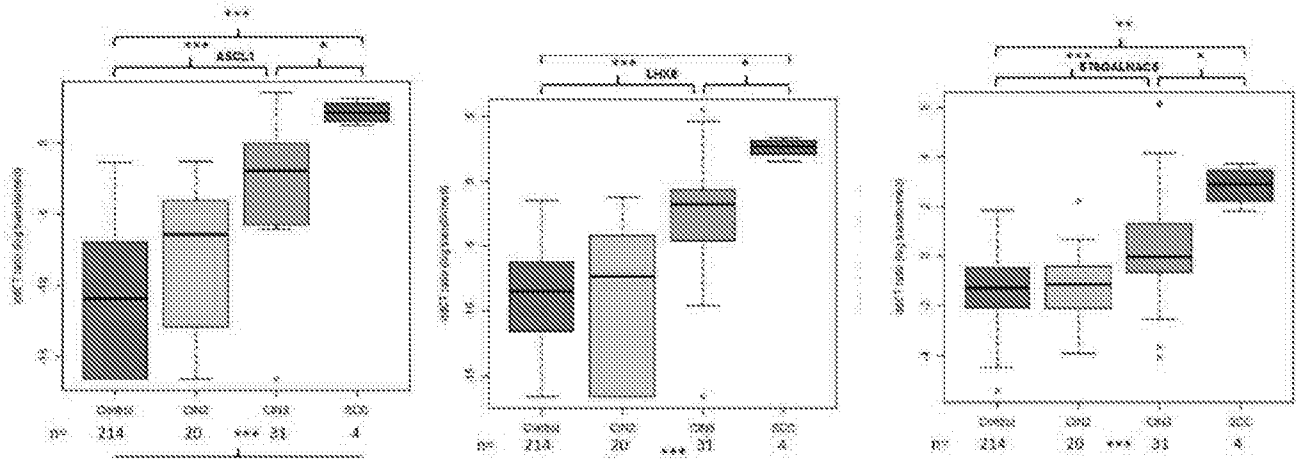
>hg19_refGene_NM_001001933
acatgtagagaaggcgggttttcctgaaagaggcgaagcaatttctccag
gaaagactttcccccacacgcccccttcctttatattaggttccaccctc
tggaaaacaaaacttttttttttttctttttcctggagggactcaggaaaa
gctcagtgctcacttcactcagagctcagtgaagctgggaaaggaatttt
aagaacggttcatcagaaagtggtcaggccacagcggcctctttggacga
agacacacttgtagcattatccttctcggcatcagcttttattagtggat
cggggcgggggagggggagatcggcagacacggacagcctctgaccctc
tggagttggtatgtgataagcagccctagcagtgccatgtattggaagaa
cgatcagatgtttgtgtgtaaactagtagcaaaggacgtgccggagctgg
cagttcccccctgagaaggtgagcgagccgacgcctggccagaccagctga
atcgcagtgtccttgaaactcgagttgtttgggctcctaaacaaggttca
gaaactacctgtaacggcctcatcttcagacctggcaatttttttttttt
attacgtagtagcgttagtcagtaatcattgcactcccctccccaaaagct
cacttaacctgagacatggagactgtaatttgggagatgaagcctcgagc
ctaaggcgctctcaggtccatgtgagtcgtgcttttgttctatttgctgt
ggtgatcgtggcggtccgggagagtgggtcgggaggactgggcggctgtc
gggtggaaccggagacctggctcgtttcgttcggctgcccgcggacactt
gcctgccctgaggcacacctcccctctgcggggctccagaaaggcctccgg
gatccctgggcttgccgggagagccattttacctaaaatgccaaaagaaa
accgcagggcacaaaacgagtcacgcggtttgtctgagtagactggggaa
gctgacaagctctcacctcactcgggaggcaggacctgtgggtatttggt
gttttaagtttgaaatcatctgcagcctgtctgagggcctcagcaggacg
cgcgcgggcgcccctgactatttctttgctccgaagccggttgggggaca
ccaggcgaaagcggagctgttctgtttaaacgccctttgtgtgtgaggac
gcggcctcagtgcagctacgctgggtcccgcgtagggagcgcaccccaga
aagggagggtctcggccgacggtccgggaggcccccaggttggtatcttg
cctggtgcttccagagggagggagaaggggaggcgaacagtcagaaatgc
agaagtgcgagccgctcgagttcgggggtacccacccgcccccgagctaga
gtttttggaaacgaaagcggtctaagcccagggacgctgcgttttcgttt
aagaacaaaactccagcaaatacttggaaacacttgagtaacgcgcgttg
tttctttaaaacattttgtcagaggttgggtgagttcttttaatggcaca
aattaaattcagctggatattttccattattctacccgcttttgcgctg
gttaaaaataatcctttagtagctcaaatgttgacttactgagcaataag
tggcaatttacaccttaaagaaacgagtgtaaatcactcgccactctaat
tgctgtaagtttacgaaagagctgctggctcggtttaatctgaagcattt
tcattcaaattgtcatcggaacaaacaccaggcttcttaaatcccgctgt
aattagctttcaagtatcctattaaacctcttcacctttcgggctatcct
attcaaaaagcccccttttttaaaatatatatgtgtcttgctcttttaatt
gaggttattaaagaaatctagtttcttctcccaccccctcctcaatcctc
ttttcttctctcgcactccacctcccactgccccgcaccccctttagccc
ggtgttccccgctccgctatgattgacgtctggaaagaaagagctttgtg
agggatgattgttattaacttgttatccccggcgggggggcgggggaacc
gacgtgcccgggtgagcgccggagacccggagccccgggggagcgcgggacg
agctaccagcgctcgggtggcggccgccagcggccagcgaaggaggctgc
gcgccagccccgccccgcggcgcccgggctcaggcgccgtgacggctgcacg
cgctgccccgcactctgagggccttcattagctcgctccccgcgccgaggg
ctgggggcgggcagcacgctcggaacttctgatctgtttctccatactttc
tccccctcctactccgcagTGTCAGGGGCTCATGTCAGAGGAGTGCGGGC
GGACTACAGCCCTGGCGGCCGGGAGGACTCGCAAAGGCGCCGGGGAAGAG
GGACTG

Figure 3: ST6GALNAC5 sequence (regulatory region, exon1, intron 1 and exon 2) derived from Refseq NM_030965 (hg19). The coding sequence is in upper case and underlined is the CpG rich region.

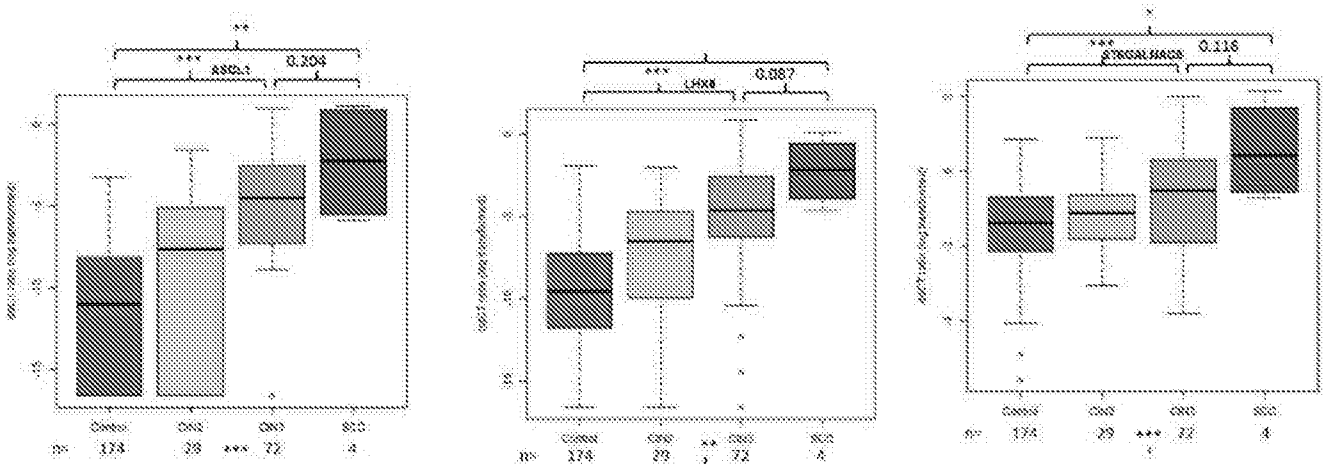
>hg19_refGene_NM_030965
agatgttaagtaatacattaagacaactattttcactgagcaatttcaca
agaaatatgtttttagctaaattagacatttaacaaatgccatttgtaat
tcaatgcagaggagacagcatccttaaaacaagctattgatgtgactttc
ccaacataaaacaatacttatgtttttttaattttaattgaagtcctgaa
tgtgaaactatcgccgtcagtgcagtagaccaacggaaagcaccaaatac
cttaggattatgaaataagccatatgcagttttatctctagcagacagaa
attcttcatccaaattaggaagaatctacaggtcaagttaatcatgtagc
tgtgacgtgctgaatgtttttaaaactttaagatcatttagatgaacact
aggaaactcagaggtcaggcagaaatgtaaattgtaatattcaatgaaaa
taataaatgaaataattgagatgaaatattatctaaagaatgcatgagaa
taactgaatctgataaagtaaaattaaaattaagagagaagagaaaataa
gaaagtgagcaattgaaaaacggaataataaataatttttttaagaaaaat
aatttcatcatgtatttctctatgctacgtacataaacgcaaacgctata
aatagagttattgacatttgggaggttgatttgtttttatcaacatcgca
aaacagaaaatttaggttgcagattttcacatggctatcacgttcagaat
ccaatcaggttctccttctttactacttgatgacttctttaaagataaat
aagcccgcggaccaagaagtgggtacactggctcggttaactctctctcc
ccagaaatttcactactgaaaagattattatttgggggcgggggaagggga
tgtagaggtctttaggacccagcaggcggcggcaggcggcagttgtgtag
atcgctgagagactacgagggtccggttcagttttaattctgtctctaat
CTCTGCAACAGCCGCGCTTCCCGGGTCCCGCGGCTCCCGCGCGCGATCTG
CCGCGGCCGGCTGCTGGGCAAAAATCAGAGCCGCCTCCGCCCCATTACCC
ATCATGGAAACCCTCCAGGAAAAAGTGGCCCCGGACGCGCGAGCCTGAGG
ATTCTGCACAAAAGAGGTGCCCAAAATGAAGACCCTGATGgtgagtcagt
tgtggcaactccaccgggcaaagaggggatccccgggctcagggtccac
gggacgcaccgtggagactccgagacgcctaaccctgggccgcgaggtcg
cctgttacaaagggacaactttctacccgctccgcgttccctcccgattc
tccagctctgcctggctcggaatcccagagccaggatgggaactcggggt
tgcctcgcctcctagatctccggcgagaggtccgaggggtggcggagag
ctgcaggagcgatggaggagtgggcagattgctcaagggatgggggtgcc
caaaagcaacagcctgccaaaaactaagagggacggggagggggacctt
tgcagactttcttcgttttcttagatttcaaacttgcaaggatcgcaagg
atccagggccccaggaaaggagggtgtgaaggactcaaaattccagcag
cttggctggggtggctgcgccagacgggccccttccccaaagtgcaaaccc
accctgtcctcggccccggcgcgctccctccctcagccccgggccgtac
accacctgccctctaccgagagatctgggcggcggcggccgaaagcagcg
acgcgccccggagcatcccttgcgatacgctaggggacggtgctttctctg
tcccagttgcgtgcggcggggctggggcccaggccgccccaaatctcccc
cactagagtgaccaccgcacagttgtccccgctgggcgcgctcctccggt
gtctgcgctcagccgctctcctcttctctctcccgcccgcccgcagCGCC
ATGGTCTGGCAGTGTGTTTAGCGCTCACCACCATGTGCACCAGCTTGTTG
CTAGTGTACAGCAGCCTCGGCGGCCAGAAGGAGCGGCCCCCGCAGCAGCA
GCAGCAGCAGCAGCAACAGCAGCAGCAGGCGTCGGCCACCGGCAGCTCGC
AGCCGGCGGCGGAGAGCAGCACCCAGCAGCGCCCCGGGGTCCCCGCGGGA
CCGCGGCCACTGGACGGATACCTCGGAGTGGCGGACCACAAG

Figure 4

A. hrHPV-positive lavage self-samples



B. hrHPV-positive brush self-samples



C.

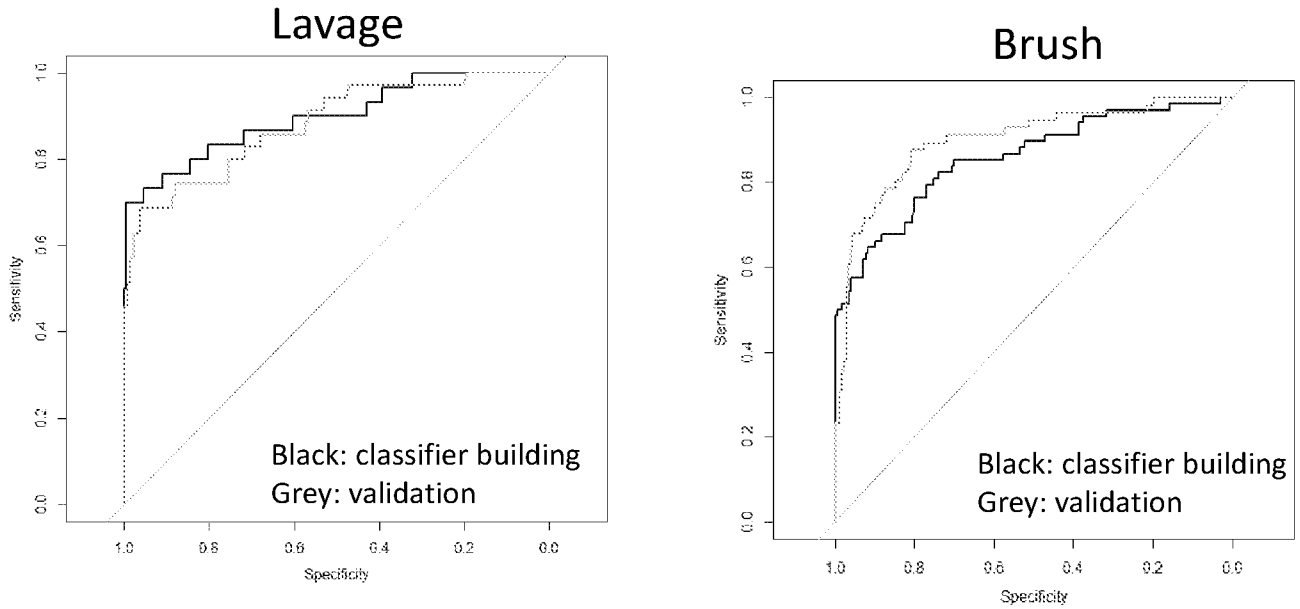
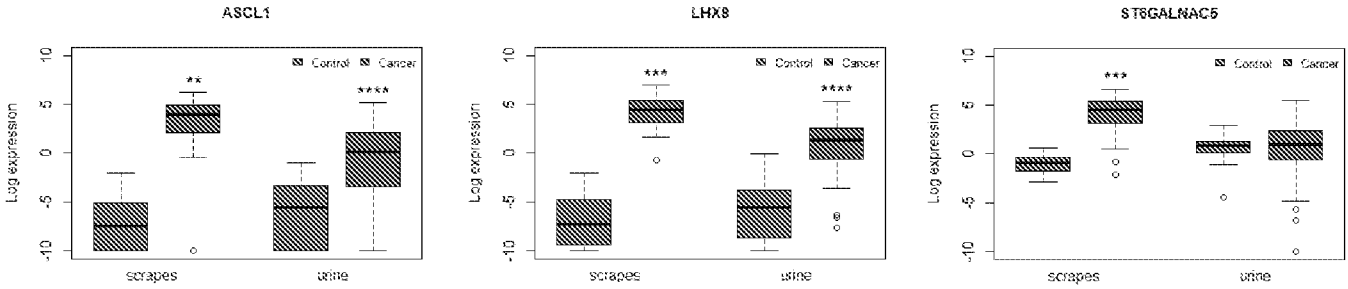
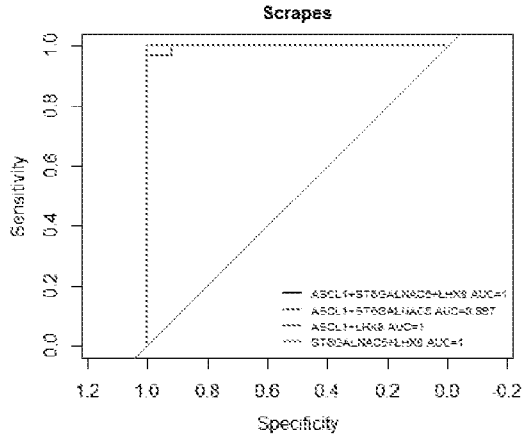


Figure 5

A. qMSP results in scrapes and urine



B. Receiver Operating Characteristics (ROC) curves in HPV+ scrapes



C. Receiver Operating Characteristics (ROC) curves in urine

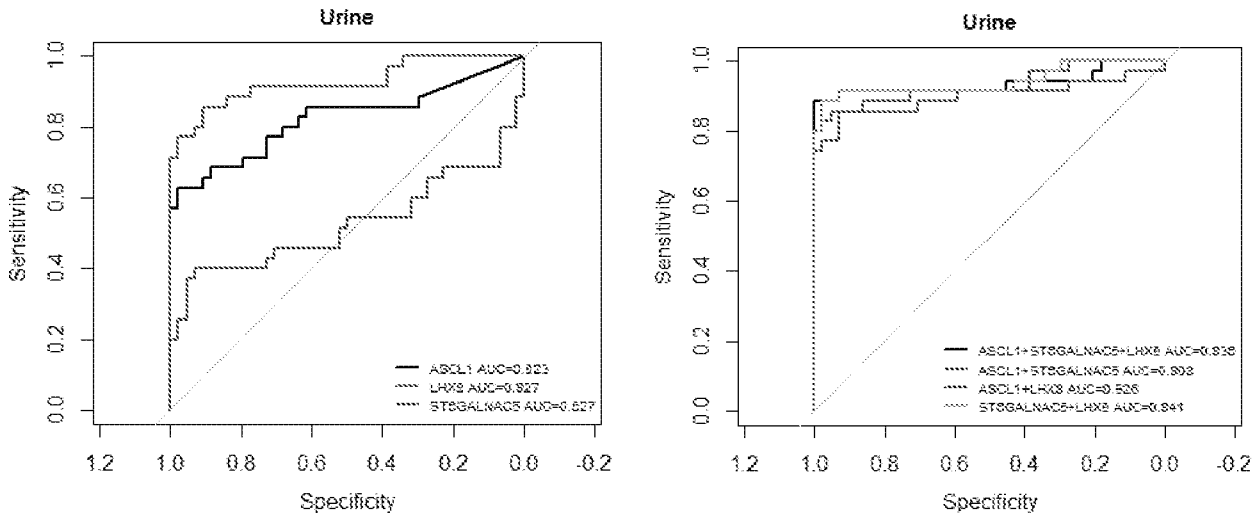
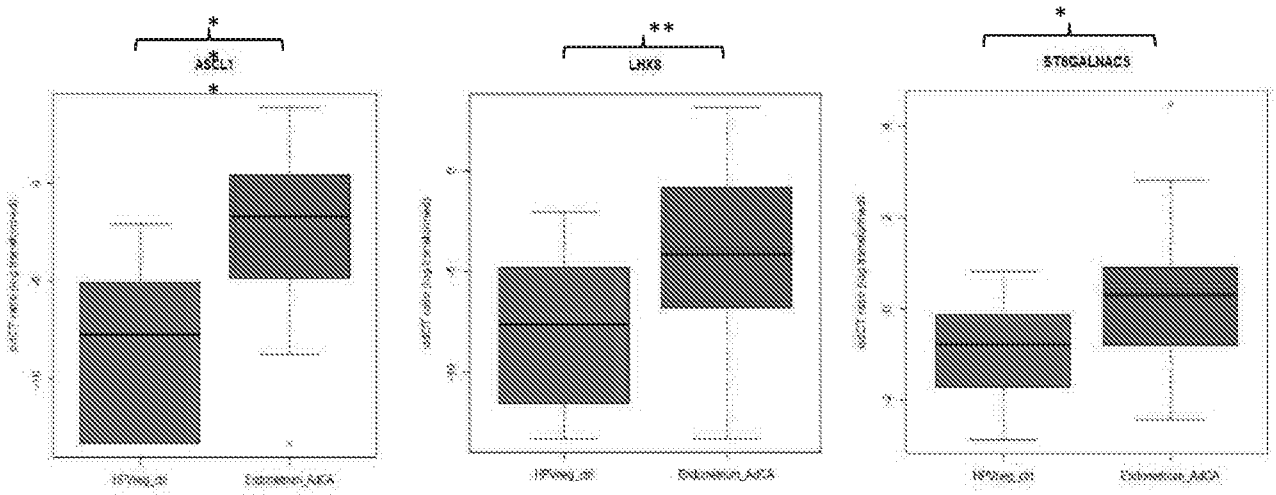


Figure 6

qMSP results in cervical scrapes of patients with endometrial cancer



n= 27 24 n= 27 24 n= 27 24

Figure 7: ZNF-582 sequence (regulatory region and exon1) derived from Refseq NM_144690 (hg19). The coding sequence is in upper case and underlined is the CpG rich region.

>hg19_refGene_NM_144690
agttcaggcattctggctccagtcttagagatggttaagggttcacactc
ttaaccatttattacaccatagagctcaccaggtttgagggaaacaggat
caaatcaaaagagtcactcaggactccagtcctcactcaaggacaaactg
ttccacctcggacagggagagtttccgcattctgagacccagcataacag
gtcctgaccggcatctggcactcggactcccaatcatactggatcacact
ggctcgggatgtgtaaagtccagggcttctcacatttgatgacaccaaag
ccgcctaaaaacaagagagaattaacaactacctacggcggtctgatatt
tgcccaagagatgccgccccataaaactcctttacatctttataacgttt
ttattttgcgttctccttcataacccacatttaactcaccatagatgtaa
tgtttaaaattagttaccagataaactcttacgcttccaaactttaaggt
tccttcgaaaccttctggtaaaactgttgttccacggaaatgggaacgta
acggatgaggcaatcttccacagccgcacacagttgtgtatccaccgcta
aacggtcccagtcatacattcaacgacccacgcggagtcagaagctacca
ccacacactgtcaaaatcacgcacacacagtgacggccccttgcccactc
ggtcactcgcccacaatctctcgctagagaatcacacgcagatagcacac
ccagcaccacagaccccaggaagcaacccagggactcgaacacacgaaca
gcactcctccgcgcactgcgcaggcacgcctgcgtccggctcaccctgaa
acatcgcgagatccggcttcaaggccgggctgctgcctttacgcctaaag
actatgtttcccggaagacactgcggcgccggccctatcatggcgcagca
tcggtgtgctttgtgcgtctgcgccatcttccggctgcgcacggcgaatc
CACCGGTACCGTGGTGGAAGCGCGCCCTGGGCTGCCGGGGGCGCGGCCGC
GGTGGCACTTGGACCCGAGGAGGCGGCAG

Figure 8: ZIC1 promoter region and coding sequence (derived from Refseq: NM_003412 (hg19)); CpG islands in grey and coding sequence in UPPER CASE



Figure 9

qMSP results in vulvar specimens

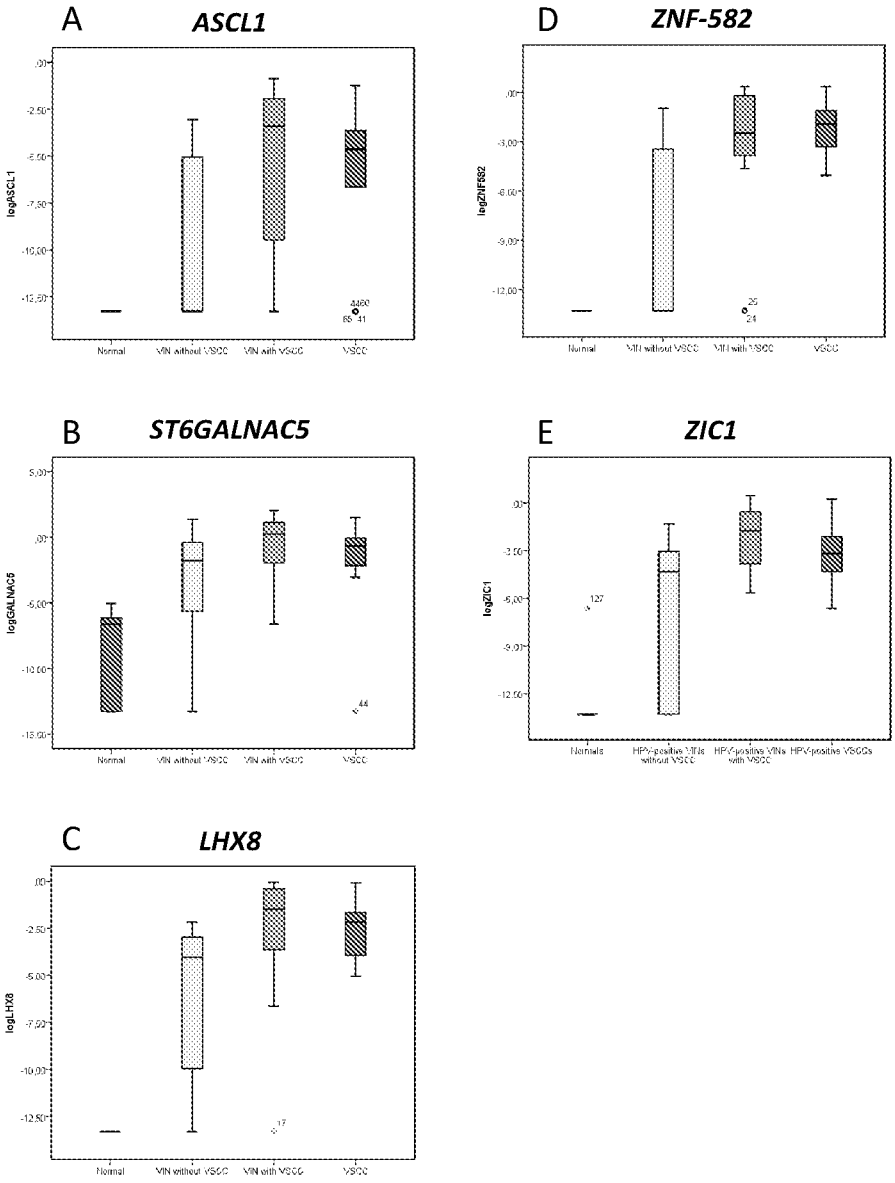
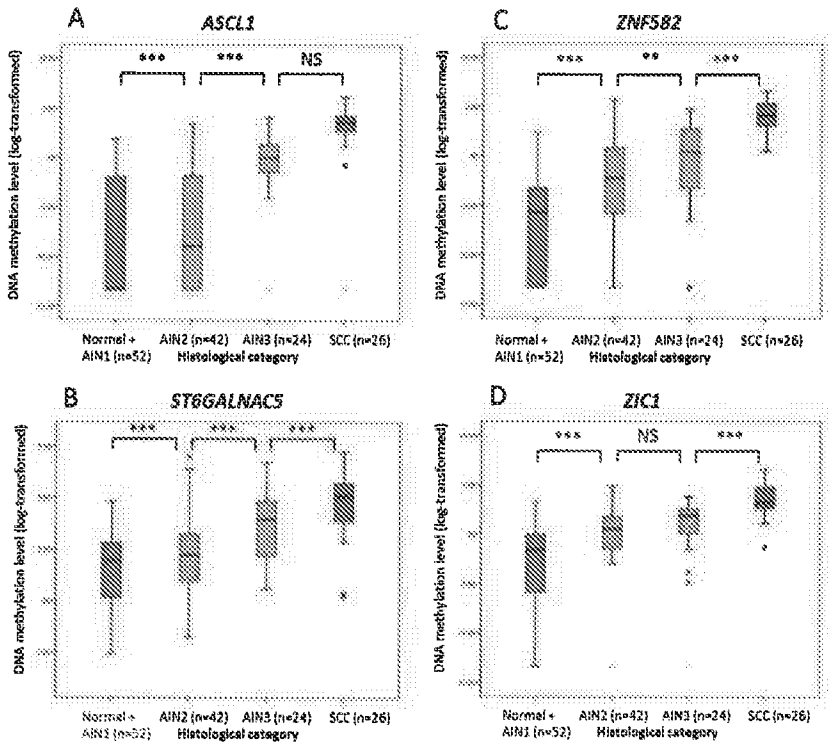


Figure 10

qMSP results in anal specimens



INTERNATIONAL SEARCH REPORT

International application No
PCT/NL2018/050143

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12Q1/6886
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, BIOSIS, Sequence Search, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2016/048138 A1 (UNIV GRONINGEN [NL]; GRONINGEN ACAD ZIEKENHUIS [NL]) 31 March 2016 (2016-03-31) claims 1, 2, 14	1-4, 8-17,21
A	SANJA A FARKAS ET AL: "Genome-wide DNA methylation assay reveals novel candidate biomarker genes in cervical cancer", EPIGENETICS, vol. 8, no. 11, 1 November 2013 (2013-11-01), pages 1213-1225, XP055203588, ISSN: 1559-2294, DOI: 10.4161/epi.26346 table 4	1-4,8-17

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
3 May 2018	14/05/2018

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Santagati, Fabio

2

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MICHAEL P. TRIMARCHI ET AL: "Identification of endometrial cancer methylation features using combined methylation analysis methods", PLOS ONE, vol. 12, no. 3, 9 March 2017 (2017-03-09), page e0173242, XP055402191, DOI: 10.1371/journal.pone.0173242 tables 1, 2	1-4,8-17
A	WO 2017/034407 A1 (SELF-SCREEN B V [NL]) 2 March 2017 (2017-03-02) cited in the application claim 1; example 7	5,6, 8-16,18, 19,21
A	WO 2016/115354 A1 (UNIV TAIPEI MEDICAL [CN]; YEN YUN [US]) 21 July 2016 (2016-07-21) page 25, line 1 - line 11; example 11; table 13	7-16,20, 21
X	Anonymous: "Infinium HumanMethylation450 BeadChip", , 9 March 2012 (2012-03-09), XP055401052, Retrieved from the Internet: URL:https://support.illumina.com/content/dam/illumina-marketing/documents/products/datasheets/datasheet_humanmethylation450.pdf [retrieved on 2017-08-24] the whole document	21

INTERNATIONAL SEARCH REPORT

International application No.
PCT/NL2018/050143

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☒ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/NL2018/050143

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2016048138 A1	31-03-2016	CN 107002136 A	01-08-2017
		EP 3198034 A1	02-08-2017
		JP 2017529843 A	12-10-2017
		US 2017240974 A1	24-08-2017
		WO 2016048138 A1	31-03-2016
WO 2017034407 A1	02-03-2017	AU 2016312177 A1	22-03-2018
		EP 3135768 A1	01-03-2017
		KR 20180038050 A	13-04-2018
		WO 2017034407 A1	02-03-2017
WO 2016115354 A1	21-07-2016	TW 201632629 A	16-09-2016
		WO 2016115354 A1	21-07-2016

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows: