Innovation, Sciences et Développement économique Canada
Office de la Propriété Intellectuelle du Canada

Innovation, Science and Economic Development Canada
Canadian Intellectual Property Office

CA 2718092 C 2019/03/19

(11)(21) **2 718 092**

(12) **BREVET CANADIEN**
CANADIAN PATENT

(13) **C**

(86) **Date de dépôt PCT/PCT Filing Date:** 2009/03/23
(87) **Date publication PCT/PCT Publication Date:** 2009/09/24
(45) **Date de délivrance/Issue Date:** 2019/03/19
(85) **Entrée phase nationale/National Entry:** 2010/09/09
(86) **N° demande PCT/PCT Application No.:** EP 2009/053386
(87) **N° publication PCT/PCT Publication No.:** 2009/115615
(30) **Priorité/Priority:** 2008/03/21 (US61/038,549)

(51) **Cl.Int./Int.Cl.** ***C12Q 1/68*** (2018.01),
C12Q 1/6809 (2018.01), ***C12Q 1/6844*** (2018.01),
C12Q 1/6886 (2018.01)

(72) **Inventeurs/Inventors:**
VAN CRIEKINGE, WIM, BE;
DEREGOWSKI, VALERIE, BE;
DEHASPE, LUC, BE;
WISMAN, G. BEA A., NL;
VAN DER ZEE, ATE G. J., NL;
SCHUURING, E. M. D., NL

(73) **Propriétaire/Owner:**
MDXHEALTH S.A., BE

(74) **Agent:** FASKEN MARTINEAU DUMOULIN LLP

(54) **Titre : DETECTION ET PRONOSTIC DU CANCER DU COL DE L'UTERUS**
(54) **Title: DETECTION AND PROGNOSIS OF CERVICAL CANCER**

(57) **Abrégé/Abstract:**
The present invention relates to methods and kits for identifying, diagnosing, prognosing, and monitoring cervical cancer. These methods include determining the methylation status or the expression levels of particular genes, or a combination thereof.

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
24 September 2009 (24.09.2009)

PCT

(10) International Publication Number
WO 2009/115615 A3

(51) International Patent Classification:
C12Q 1/68 (2006.01)

(21) International Application Number:
PCT/EP2009/053386

(22) International Filing Date:
23 March 2009 (23.03.2009)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/038,549 21 March 2008 (21.03.2008) US

(71) Applicant *(for all designated States except US)*: **ONCOMETHYLOME SCIENCES S.A.** [BE/BE]; Avenue de l'Hôpital 11, CHU, Tour 5 GIGA, niveau 3, B-4000 Sart-Tilman (Liège) (BE).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **VAN CRIEKINGE, Wim** [BE/BE]; ONCOMETHYLOME SCIENCES S.A., Avenue de l'Hôpital 11, CHU, Tour 5 GIGA, niveau 3, B-4000 Sart-Tilman (Liège) (BE). **DEREGOWSKI, Valerie** [BE/BE]; ONCOMETHYLOME SCIENCES S.A., Avenue de l'Hôpital 11, CHU, Tour 5 GIGA, niveau 3, B-4000 Sart-Tilman (Liège) (BE). **DEHASPE, Luc** [BE/BE]; ONCOMETHYLOME SCIENCES S.A., Avenue de l'Hôpital 11, CHU, Tour 5 GIGA, niveau 3, B-4000 Sart-Tilman (Liège) (BE). **WISMAN, G. Bea A.** [NL/NL]; Department of Gynaecological Oncology, University Medical Centre Groningen, University of Groningen, NL-9700 RB Groningen (NL). **VAN DER ZEE, Ate G. J.** [NL/NL]; Department of Gynaecological Oncology, University Medical Centre Groningen, University of Groningen, NL-9700 RB Groningen (NL). **SCHUURING, E.m.d.** [NL/NL]; Department of Gynaecological Oncology, University Medical Centre Groningen, University of Groningen, NL-9700 RB Groningen (NL).

(74) Agent: **CAMPABADAL i MONFÀ, Gemma**; Koningin Astridlaan 33, B-9100 Sint-Niklaas (BE).

(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, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, 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, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (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, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *as to the identity of the inventor (Rule 4.17(i))*

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

— *of inventorship (Rule 4.17(iv))*

Published:

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

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

— *with sequence listing part of description (Rule 5.2(a))*

(88) Date of publication of the international search report:
12 November 2009

(54) Title: DETECTION AND PROGNOSIS OF CERVICAL CANCER

(57) Abstract: The present invention relates to methods and kits for identifying, diagnosing, prognosing, and monitoring cervical cancer. These methods include determining the methylation status or the expression levels of particular genes, or a combination thereof.

WO 2009/115615 A3

DETECTION AND PROGNOSIS OF CERVICAL CANCER

FIELD OF THE INVENTION

The present invention relates to the area of cancer diagnostics and therapeutics. In particular, it relates to methods and kits for identifying, diagnosing, prognosing, and monitoring cervical cancer. These methods include determining the methylation status or the expression levels of particular genes, or a combination thereof.

BACKGROUND TO THE INVENTION

Cervical cancer is the fifth most deadly cancer in women. Worldwide, approximately 500,000 cases of cervical cancer are diagnosed and about 250,000 women die from this disease annually.

Most (80-90%) invasive cervical cancer develops in flat, scaly surface cells that line the cervix (called squamous cell carcinomas, SCC). Approximately 10-15% of cases develop in glandular surface cells (called adenocarcinomas, AdC). Less commonly, cervical cancers have features of both SCC and AdC. These are called adenosquamous carcinomas or mixed carcinomas.

During the process of cervical cancer development, normal cervical cells gradually develop pre-cancerous changes that turn into cancer. Cervical cancer evolves from pre-existing noninvasive premalignant lesions referred to as cervical intraepithelial neoplasias (CINs), ranging from CIN I (mild dysplasia) to CIN II (moderate dysplasia) to CIN III (severe dysplasia/carcinoma in situ). This process usually takes several years but sometimes can happen in less than a year. For most women, pre-cancerous cells will remain unchanged and disappear without any treatment.

Screening for malignant and premalignant disorders of the cervix is usually performed according to the Papanicolaou (PAP) system. The cervical smears are examined by light microscopy and the specimens containing morphologically abnormal cells are classified into PAP I to V, at a scale of increasing severity of the lesion. But, present PAP test has some limitations and is not completely ideal for screening as it suffers from suboptimal single-test sensitivity, limited reproducibility, and many equivocal.

There is a strong association between certain subtypes of the Human Papillomavirus (HPV) and cervical cancer. Studies have shown that only high-risk HPV types are

involved in the progression from cytological normal cervix cells to high grade squamous intraepithelial lesions. Around 15 high-risk (cancer-causing) HPV types have been identified. Although it has been suggested that high-risk HPV testing may improve cervical cancer screening, the specificity for high grade cervical neoplasia of high risk HPV testing is relatively low. This low specificity of HPV testing leads to a higher number of unnecessarily follow-up diagnostic workups (e.g. colposcopy) and unnecessarily treatment with cryotherapy or loop electrosurgical excision procedure, which permanently alters the cervix and have unknown consequences on fertility and pregnancy.

To improve early detection, the combination of HPV and PAP tests is now approved by the FDA for screening women 30 years of age and older. However, co-testing substantially increases the cost of screening.

In the meanwhile, vaccines for preventing cervical cancer have been developed and one has already been approved by the FDA. But, immunization will only protect against HPV types that are targeted by the vaccine; protection will not be absolute and its longevity is uncertain; as yet, the possibility of genotype replacement cannot be excluded; and older women not covered by vaccination programs will continue to be at risk. Therefore, cervical screening will still be required for control.

Cancer biomarkers have been described in literature and aberrant methylation of genes has been linked to cervical cancer (Virmani *et al,* 2001). In addition, methylation markers may serve for predictive purposes as they often reflect the sensitivity to therapy or duration of patient survival.

DNA methylation is a chemical modification of DNA performed by enzymes called methyltransferases, in which a methyl group (m) is added to certain cytosines (C) of DNA. This non-mutational (epigenetic) process (mC) is a critical factor in gene expression regulation. (See J.G. Herman, Seminars in Cancer Biology, 9: 359-67, 1999).

An early diagnosis is critical for the successful treatment of many types of cancer, including cervical cancer. If the exact methylation profiles of cervical tumors are available and drugs targeting the specific genes are obtainable, then the treatment of cervical cancer could be more focused and rational. Therefore, the detection and mapping of novel methylation markers is an essential step towards improvement of cervical cancer prevention, screening, and treatment. Thus, there is a continuing need

in the art to identify methylation markers that can be used for improved assessment of cervical cancer.

SUMMARY OF THE INVENTION

The present invention is based on the finding that several genes are identified as being differentially methylated in cervical cancers. This information is useful for cervical cancer screening, risk-assessment, prognosis, disease identification, disease staging, and identification of therapeutic targets. The identification of new genes that are methylated in cervical cancer allows accurate and effective early diagnostic assays, methylation profiling using multiple genes and identification of new targets for therapeutic intervention.

Accordingly, in a first aspect, the invention provides a method for identifying cervical cancer or its precursor, or predisposition to cervical cancer. Epigenetic modification of at least one gene selected from the group consisting of genes according to Table 1, is detected in a test sample containing cervical cells or nucleic acids from cervical cells. The test sample is identified as containing cells that are neoplastic, precursor to neoplastic, or predisposed to neoplasia, or as containing nucleic acids from cells that are neoplastic, precursor to neoplastic, or predisposed to neoplasia.

Preferably, the at least one gene is selected from a group of genes consisting of JAM3, LMX1A, CDO1, NID2, ALX3, ALX4, AR, ARID4A, ATM, AURKA, B4GALT1, BMP2, BMP6, BNIP3, C13orf18, C16orf48, C9orf19, CALCA, CAMK4, CCNA1, CCND2, CDH1, CDH4, CDK6, CDKN1B, CDKN2B, CLSTN2, CLU, COL1A1, CPT1C, CTDSPL, CYCLIND2, DAPK1, DBC1, DDX19B, DKK2, EGFR, EGR4, EPB41L3, FOS, FOXE1, GADD45A, GATA4, GDAP1L1, GNB4, GPNMB, GREM1, Gst-Pi, HHIP, HIN1, HOOK2, HOXA1, HOXA11, HOXA7, HOXD1, IGSF4, ISYNA1, JPH3, KNDC1, KRAS, LAMA1, LOC285016, LOX, LTB4R, MAL, MTAP, MYO18B, NDRG2, NOL4, NPTX1, NPTX2, OGFOD2, PAK3, PAX1, PDCD4, PHACTR3, POMC, PRKCE, RAD23B, RALY, RARA, RASSF1A, RBP4, RECK, RPRM, SALL4, SEMA3F, SLC5A8, SLIT1, SLIT2, SLIT3, SMPD1, SOCS1, SOX1, SOX17, SPARC, SPN, SST, TAC1, TERT, TFPI-2, TLL1, TNFAIP1, TRMT1, TWIST1, UGT1A1, WIF1, WIT1, WT1, XRCC3, and ZGPAT.

In one embodiment of the present invention, the detection of epigenetic modification comprises detection of methylation of a CpG dinucleotide motif in the gene and/or promoter region of the gene; and/or detection of expression of mRNA of the gene.

The invention also relates to a kit for assessing cervical cancer or its precursor, or predisposition to cervical cancer in a test sample containing cervical cells or nucleic acids from cervical cells. The kit comprises in a package: a reagent that (a) modifies methylated cytosine residues but not non-methylated cytosine residues, or that (b) modifies non-methylated cytosine residues but not methylated cytosine residues; and at least one pair of oligonucleotide primers that specifically hybridizes under amplification conditions to a region of a gene selected from the group consisting of genes according to Table 1 and/or the aforementioned group of genes. The region is preferably within about10 kbp of said gene's transcription start site.

In a further aspect, the invention provides for oligonucleotide primers and/or probes and their sequences for use in the methods and assays of the invention.

The invention also relates to screening protocols for the screening of woman for cervical cancer and the precursors thereof. Such method for cervical cancer screening combines hr-HPV testing and methylation testing, or combines PAP tests with methylation testing. Methylation testing in such screening method preferably detects the epigenetic modification of at least one gene selected from the group consisting of JAM3, LMX1A, CDO1, NID2, ALX3, ALX4, AR, ARID4A, ATM, AURKA, B4GALT1, BMP2, BMP6, BNIP3, C13orf18, C16orf48, C9orf19, CALCA, CAMK4, CCNA1, CCND2, CDH1, CDH4, CDK6, CDKN1B, CDKN2B, CLSTN2, CLU, COL1A1, CPT1C, CTDSPL, CYCLIND2, DAPK1, DBC1, DDX19B, DKK2, EGFR, EGR4, EPB41L3, FOS, FOXE1, GADD45A, GATA4, GDAP1L1, GNB4, GPNMB, GREM1, Gst-Pi, HHIP, HIN1, HOOK2, HOXA1, HOXA11, HOXA7, HOXD1, IGSF4, ISYNA1, JPH3, KNDC1, KRAS, LAMA1, LOC285016, LOX, LTB4R, MAL, MTAP, MYO18B, NDRG2, NOL4, NPTX1, NPTX2, OGFOD2, PAK3, PAX1, PDCD4, PHACTR3, POMC, PRKCE, RAD23B, RALY, RARA, RASSF1A, RBP4, RECK, RPRM, SALL4, SEMA3F, SLC5A8, SLIT1, SLIT2, SLIT3, SMPD1, SOCS1, SOX1, SOX17, SPARC, SPN, SST, TAC1, TERT, TFPI-2, TLL1, TNFAIP1, TRMT1, TWIST1, UGT1A1, WIF1, WIT1, WT1, XRCC3, and ZGPAT. Dependent on the outcome, the women screened for cervical cancer is referred for colposcopy, or referred for hr-HPV and/or PAP testing and/or methylation testing on a more regular basis.

Epigenetic loss of gene function can be rescued by the use of DNA demethylating agents and/or DNA methyltransferase inhibitors. Accordingly, the invention also provides for a method for predicting the likelihood of successful treatment or resistance to treatment of cancer with such agent. If the gene is methylated, the

likelihood of successful treatment is higher than if the gene is unmethylated, or methylated to a lesser degree. Conversely, if the gene is unmethylated, or methylated to a lesser degree, the likelihood of resistance to treatment is higher than if the gene is methylated.

In a related aspect, epigenetic loss of gene function(s) can identify the stage of the disease and from that the need of treatment. Accordingly, the invention provides for a method for predicting suitable treatment comprising determining the methylation status of a gene or a combination of genes. If the gene is methylated, the need of cervical resection is identified; if the gene is unmethylated or methylated to a lesser degree, it is decided that there is no need for cervical resection.

SUMMARY OF THE FIGURES

Figures 1A, 1B, and 1C: The number of probes (w) that is retrieved using parameters x (number of P-calls in primary cancers for probe), y (number of P-calls in untreated cell-lines for probe) and z (number of P-calls in treated cell-lines for probe).

Figure 2: Step-plot to determine optimal number of probes for further analysis. Step-plot of the number of retrieved known markers as a function of the position after relaxation ranking (this is the number of selected probes after ranking). The step plot shows the actual (observed) number of markers. If the markers were randomly distributed, one would expect the profile, marked with 'expected' (details in the text). The trend of the observed markers versus the number of selected probes is indicated with dashed lines.

Figure 3: (Hyper) methylation analysis of the promoter region (-430 to -5 of TSS) of the CCNA1 gene by COBRA and sequence analysis.
A: schematic representation of the restriction enzyme sites (B: BstUI and T: TaqI) in the virtual hypermethylated BSP nucleotide sequence after bisulfite treatment. Vertical bars represent CG site, arrow represents TSS (retrieved from Ensembl).
B: Result of COBRA analysis of the BSP products of 10 tumor samples (T1-T10), *in vitro* methylated DNA as a positive control (IV) and leukocyte DNA as a negative (unmethylated) control (L).
C: Schematic representation of the sequencing results. From each tumor, the BSP-products were cloned into TOPO®-pCR4 (Invitrogen) and sequencing (BaseClear) was performed on M13-PCR products of 7-9 independent clones. Circles represent CG dinucleotides: the darker, the more clones at this site were methylated.

Figure 4: Representative COBRA on 3 gene promoters (SST, AUTS2 and SYCP3).
A: schematic representation of the restriction enzyme sites in the virtual hypermethylated BSP nucleotide sequence after bisulfite treatment. (B: BstUI, T: TaqI and H: HinfI). Bars represent CG site and arrow is TSS (retrieved from Ensembl).
B: Result of COBRA analysis of BSP products of tumor samples (T1-T10) and 5 normal cervices (N1-N5), *in vitro* methylated DNA as a positive control (IV) and leukocyte DNA as a negative (unmethylated) control (L); lane B is water blank.

Figure 5:
A: Position of the different primers relative to the TSS (transcription start site). Multiple primer designs are displayed by blue boxes and red boxes (=final primer pairs retained for the assays). The exon of ALX4 is indicated in green. The number of CpG count is spotted in blue over a region of 20Kb.
B: List of sequences for the different primer sets, converted and unconverted amplicon sequences used in Figure 5 A.

Figure 6: Ranked methylation table from the LightCycler® platform. 27 methylation profiles from cervical cancer samples (left) are compared against 20 normal tissue samples (right). Samples are shown along the X-axis where each vertical column represents the methylation profile of one individual sample across the 63 different assays (Y-axis). Assays demonstrating the best methylation discriminators between the 2 groups are displayed at the top, with discrimination effect decreasing towards the bottom. The black boxes indicate the methylated results; grey boxes indicate the unmethylated results; white boxes indicate invalid results. (NA: not applicable; NT: not tested)

Figure 7: Amplification plot for the standard curve for *TAC1_56187*

Figure 8: Amplification plot for standard curve and samples for *TAC1_56187*

Figure 9: Linear regression of standard curve for *TAC1_56187*

Figure 10: Decision tree for ratio determination

Figure 11: Performance of the individual markers on cervical tissue samples using qMSP.

DETAILED DESCRIPTION OF THE INVENTION

We describe a new sorting methodology to enrich for genes which are silenced by promoter methylation in human cervical cancer. The pharmacological unmasking expression microarray approach is an elegant method to enrich for genes that are

silenced and re-expressed during functional reversal of DNA methylation upon treatment with demethylating agents. However, such experiments are performed in *in vitro* (cancer) cell lines mostly with poor relevance when extrapolating to primary cancers. To overcome this problem, we incorporated data from primary cancer samples in the experimental design. A pharmacological unmasking microarray approach was combined with microarray expression data of primary cancer samples. For the integration of data from both cell lines and primary cancers, we developed a novel ranking strategy, which combines reactivation in cell lines and no expression in primary cancer tissue.

We also used a Genome-wide Promoter Alignment approach with the capacity to define a further substantial fraction of the cancer gene promoter CpG island DNA methylome. Markers clustering with known methylation markers might indicate towards common mechanisms underlying the methylation event and thus identify novel genes that are more methylation-prone.

Studies of the genes defined by the different approaches will contribute to understanding the molecular pathways driving tumorigenesis, provide useful new DNA methylation biomarkers to monitor cancer risk assessment, early diagnosis, and prognosis, and permit better monitoring of gene re-expression during cancer prevention and/or therapy strategies.

Using the aforementioned techniques, we have identified cytosines within CpG dinucleotides of DNA from particular genes isolated from a test sample, which are differentially methylated in human cervical cancer tissue samples and normal cervical tissue control samples. The cancer tissues samples are hypermethylated or hypomethylated with respect to the normal samples (collectively termed epigenetic modification). The differential methylation has been found in genomic DNA of at least one gene selected from the group consisting of JAM3, LMX1A, CDO1, NID2, ALX3, ALX4, AR, ARID4A, ATM, AURKA, B4GALT1, BMP2, BMP6, BNIP3, C13orf18, C16orf48, C9orf19, CALCA, CAMK4, CCNA1, CCND2, CDH1, CDH4, CDK6, CDKN1B, CDKN2B, CLSTN2, CLU, COL1A1, CPT1C, CTDSPL, CYCLIND2, DAPK1, DBC1, DDX19B, DKK2, EGFR, EGR4, EPB41L3, FOS, FOXE1, GADD45A, GATA4, GDAP1L1, GNB4, GPNMB, GREM1, Gst-Pi, HHIP, HIN1, HOOK2, HOXA1, HOXA11, HOXA7, HOXD1, IGSF4, ISYNA1, JPH3, KNDC1, KRAS, LAMA1, LOC285016, LOX, LTB4R, MAL, MTAP, MYO18B, NDRG2, NOL4, NPTX1, NPTX2, OGFOD2, PAK3, PAX1, PDCD4, PHACTR3, POMC, PRKCE, RAD23B, RALY, RARA, RASSF1A, RBP4, RECK, RPRM,

SALL4, SEMA3F, SLC5A8, SLIT1, SLIT2, SLIT3, SMPD1, SOCS1, SOX1, SOX17, SPARC, SPN, SST, TAC1, TERT, TFPI-2, TLL1, TNFAIP1, TRMT1, TWIST1, UGT1A1, WIF1, WIT1, WT1, XRCC3, and ZGPAT.

Accordingly, in a first aspect, the invention provides a method for identifying cervical cancer or its precursor, or predisposition to cervical cancer. Epigenetic modification of at least one gene selected from the group consisting of genes according to Table 1, is detected in a test sample containing cervical cells or nucleic acids from cervical cells. The test sample is identified as containing cells that are neoplastic, precursor to neoplastic, or predisposed to neoplasia, or as containing nucleic acids from cells that are neoplastic, precursor to neoplastic, or predisposed to neoplasia.

Preferably, the at least one gene is selected from a group of genes consisting of JAM3, LMX1A, CDO1, NID2, ALX3, ALX4, AR, ARID4A, ATM, AURKA, B4GALT1, BMP2, BMP6, BNIP3, C13orf18, C16orf48, C9orf19, CALCA, CAMK4, CCNA1, CCND2, CDH1, CDH4, CDK6, CDKN1B, CDKN2B, CLSTN2, CLU, COL1A1, CPT1C, CTDSPL, CYCLIND2, DAPK1, DBC1, DDX19B, DKK2, EGFR, EGR4, EPB41L3, FOS, FOXE1, GADD45A, GATA4, GDAP1L1, GNB4, GPNMB, GREM1, Gst-Pi, HHIP, HIN1, HOOK2, HOXA1, HOXA11, HOXA7, HOXD1, IGSF4, ISYNA1, JPH3, KNDC1, KRAS, LAMA1, LOC285016, LOX, LTB4R, MAL, MTAP, MYO18B, NDRG2, NOL4, NPTX1, NPTX2, OGFOD2, PAK3, PAX1, PDCD4, PHACTR3, POMC, PRKCE, RAD23B, RALY, RARA, RASSF1A, RBP4, RECK, RPRM, SALL4, SEMA3F, SLC5A8, SLIT1, SLIT2, SLIT3, SMPD1, SOCS1, SOX1, SOX17, SPARC, SPN, SST, TAC1, TERT, TFPI-2, TLL1, TNFAIP1, TRMT1, TWIST1, UGT1A1, WIF1, WIT1, WT1, XRCC3, and ZGPAT.

Preferably, at least one gene is selected from the group consisting of JAM3, LMX1A, CDO1, NID2, CCNA1, HOXA11, GREM1 and TAC1. Preferably, epigenetic silencing of a gene combination is detected and preferably selected from the group of gene combinations consisting of:

- NID2 and HOXA11;
- JAM3, CDO1, HOXA11, and CCNA1;
- JAM3 and HOXA11;
- JAM3, HOXA11 and GREM1;
- JAM3, NID2, HOXA11 and CDO1;
- JAM3, TAC1, HOXA11, and CDO1;
- JAM3, HOXA11, and CDO1;
- JAM3 and CDO1;

- JAM3 and NID2;
- NID2 and CDO1;
- JAM3 and LMX1A
- NID2 and LMX1A, and
- JAM3, CDO1 and NID2

"Identifying" a disease or predisposition of disease is defined herein to include detecting by way of routine examination, screening for a disease or pre-stadia of a disease, monitoring staging and the state and/or progression of the disease, checking for recurrence of disease following treatment and monitoring the success of a particular treatment. The identification can also have prognostic value, and the prognostic value of the tests can be used as a marker of potential susceptibility to cancer.

The term "Epigenetic modification" can be described as a stable alteration in gene expression potential that takes place during development and cell proliferation, mediated by mechanisms other than alterations in the primary nucleotide sequence of a gene. Three related mechanisms that cause alteration in gene expression are recognized: DNA methylation, histone code changes and RNA interference.

Epigenetic modification of a gene can be determined by any method known in the art. One method is to determine that a gene which is expressed in normal cells or other control cells is less expressed or not expressed in tumor cells. Diminished gene expression can be assessed in terms of DNA methylation status or in terms of expression levels as determined by their methylation status, generally manifested as hypermethylation. Conversely, a gene can be more highly expressed in tumor cells than in control cells in the case of hypomethylation. This method does not, on its own, however, indicate that the silencing or activation is epigenetic, as the mechanism of the silencing or activation could be genetic, for example, by somatic mutation. One method to determine that silencing is epigenetic is to treat with a reagent, such as DAC (5'-deazacytidine), or with a reagent which changes the histone acetylation status of cellular DNA or any other treatment affecting epigenetic mechanisms present in cells, and observe that the silencing is reversed, *i.e.,* that the expression of the gene is reactivated or restored.

Another means to determine epigenetic modification is to determine the presence of methylated CpG dinucleotide motifs in the silenced gene or the absence of methylation CpG dinucleotide motifs in the activated gene. In one embodiment, epigenetic

modification of a CpG dinucleotide motif in the promoter region of the at least one gene selected from a group of genes according to Table 1 is determined. Methylation of a CpG island at a promoter usually prevents expression of the gene. The islands can surround the 5' region of the coding region of the gene as well as the 3' region of the coding region. Thus, CpG islands can be found in multiple regions of a nucleic acid sequence. The term "region" when used in reference to a gene includes sequences upstream of coding sequences in a regulatory region including a promoter region, in the coding regions (e.g., exons), downstream of coding regions in, for example, enhancer regions, and in introns. All of these regions can be assessed to determine their methylation status. When the CpG distribution in the promoter region is rather scarce, levels of methylation are assessed in the intron and/or exon regions. The region of assessment can be a region that comprises both intron and exon sequences and thus overlaps both regions. Typically these reside near the transcription start site (TSS), for example, within about 10 kbp, within about 5 kbp, within about 3 kbp, within about 1 kbp, within about 750 bp, within about 500 bp, within 200 bp or within 100 bp. Once a gene has been identified as the target of epigenetic modification in tumor cells, determination of reduced or enhanced expression can be used as an indicator of epigenetic modification.

Expression of a gene can be assessed using any means known in the art. Typically expression is assessed and compared in test samples and control samples which may be normal, non-malignant cells. Either mRNA or protein can be measured. Methods employing hybridization to nucleic acid probes can be employed for measuring specific mRNAs. Such methods include using nucleic acid probe arrays (e.g. microarray technology, in situ hybridization, Northern blots). Messenger RNA can also be assessed using amplification techniques, such as RT-PCR. Sequencing-based methods are an alternative; these methods started with the use of expressed sequence tags (ESTs), and now include methods based on short tags, such as serial analysis of gene expression (SAGE) and massively parallel signature sequencing (MPSS). Differential display techniques provide another means of analyzing gene expression; this family of techniques is based on random amplification of cDNA fragments generated by restriction digestion, and bands that differ between two tissues identify cDNAs of interest. Specific proteins can be assessed using any convenient method including immunoassays and immuno-cytochemistry but are not limited to that. Most such methods will employ antibodies, or engineered equivalents thereof, which are specific for the particular protein or protein fragments. The sequences of the mRNA

(cDNA) and proteins of the markers of the present invention are known in the art and publicly available.

Alternatively, methylation-sensitive restriction endonucleases can be used to detect methylated CpG dinucleotide motifs. Such endonucleases may either preferentially cleave methylated recognition sites relative to non-methylated recognition sites or preferentially cleave non-methylated relative to methylated recognition sites. Non limiting examples of the former are Aat II, Acc III, Ad I, AcI I, Age I, AIu I, Asc I, Ase 1, AsiS I, Ban I, Bbe I, BsaA I, BsaH I, BsiE I, BsiW I, BsrV I, BssK 1, BstB I, BstN I, Bs I, CIa I, Eae I, Eag I, Fau I, Fse I, Hha I, mPl I, HinC II, Hpa 11, Npy99 I, HpyCAIV, Kas I, Mbo I, MIu I, MapA l I. Msp I, Nae I, Nar I, Not 1, PmI I, Pst I, Pvu I, Rsr II, Sac II, Sap I, Sau3A I, Sfl I, Sfo I, SgrA I, Sma I SnaB I, Tsc I, Xma I, and Zra I. Non limiting examples of the latter are Acc II, Ava I, BssH II, BstU I, Hpa II, Not I, and Mho I.

Alternatively, chemical reagents can be used that selectively modify either the methylated or non-methylated form of CpG dinucleotide motifs. Modified products can be detected directly, or after a further reaction which creates products that are easily distinguishable. Means which detect altered size and/or charge can be used to detect modified products, including but not limited to electrophoresis, chromatography, and mass spectrometry. Examples of such chemical reagents for selective modification include hydrazine and bisulfite ions. Hydrazine-modified DNA can be treated with piperidine to cleave it. Bisulfite ion-treated DNA can be treated with alkali. Other means for detection that are reliant on specific sequences can be used, including but not limited to hybridization, amplification, sequencing, and ligase chain reaction. Combinations of such techniques can be used as is desired.

The principle behind electrophoresis is the separation of nucleic acids via their size and charge. Many assays exist for detecting methylation and most rely on determining the presence or absence of a specific nucleic acid product. Gel electrophoresis is commonly used in a laboratory for this purpose.

One may use MALDI mass spectrometry in combination with a methylation detection assay to observe the size of a nucleic acid product. The principle behind mass spectrometry is the ionizing of nucleic acids and separating them according to their mass to charge ratio. Similar to electrophoresis, one can use mass spectrometry to detect a specific nucleic acid that was created in an experiment to determine methylation (Tost, J. *et al.* 2003).

One form of chromatography, high performance liquid chromatography, is used to separate components of a mixture based on a variety of chemical interactions between a substance being analyzed and a chromatography column. DNA is first treated with sodium bisulfite, which converts an unmethylated cytosine to uracil, while methylated cytosine residues remain unaffected. One may amplify the region containing potential methylation sites via PCR and separate the products via denaturing high performance liquid chromatography (DHPLC). DHPLC has the resolution capabilities to distinguish between methylated (containing cytosine) and unmethylated (containing uracil) DNA sequences. Deng, D. et al. describes simultaneous detection of CpG methylation and single nucleotide polymorphism by denaturing high performance liquid chromatography.

Hybridization is a technique for detecting specific nucleic acid sequences that is based on the annealing of two complementary nucleic acid strands to form a double-stranded molecule. One example of the use of hybridization is a microarray assay to determine the methylation status of DNA. After sodium bisulfite treatment of DNA, which converts an unmethylated cytosine to uracil while methylated cytosine residues remain unaffected, oligonucleotides complementary to potential methylation sites can hybridize to the bisulfite-treated DNA. The oligonucleotides are designed to be complimentary to either sequence containing uracil (thymine) or sequence containing cytosine, representing unmethylated and methylated DNA, respectively. Computer-based microarray technology can determine which oligonucleotides hybridize with the DNA sequence and one can deduce the methylation status of the DNA. Similarly primers can be designed to be complimentary to either sequence containing uracil (thymine) or sequence containing cytosine. Primers and probes that recognize the converted methylated form of DNA are dubbed methylation-specific primers or probes (MSP).

An additional method of determining the results after sodium bisulfite treatment involves sequencing the DNA to directly observe any bisulfite-modifications. Pyrosequencing technology is a method of sequencing-by-synthesis in real time. It is based on an indirect bioluminometric assay of the pyrophosphate (PPi) that is released from each deoxynucleotide (dNTP) upon DNA-chain elongation. This method presents a DNA template-primer complex with a dNTP in the presence of an exonuclease-deficient Klenow DNA polymerase. The four nucleotides are sequentially added to the reaction mix in a predetermined order. If the nucleotide is complementary to the template base and thus incorporated, PPi is released. The PPi

and other reagents are used as a substrate in a luciferase reaction producing visible light that is detected by either a luminometer or a charge-coupled device. The light produced is proportional to the number of nucleotides added to the DNA primer and results in a peak indicating the number and type of nucleotide present in the form of a pyrogram. Pyrosequencing can exploit the sequence differences that arise following sodium bisulfite-conversion of DNA.

A variety of amplification techniques may be used in a reaction for creating distinguishable products. Some of these techniques employ PCR. Other suitable amplification methods include the ligase chain reaction (LCR) (Barringer *et al*, 1990), transcription amplification (Kwoh *et al*. 1989; WO88/10315), selective amplification of target polynucleotide sequences (US Patent No. 6,410,276), consensus sequence primed polymerase chain reaction (US Patent No 4,437,975), arbitrarily primed polymerase chain reaction (WO90/06995), nucleic acid based sequence amplification (NASBA) (US Patent Nos. 5,409,818; 5,554,517; 6,063,603), microsatellite length polymorphism (MLP), and nick displacement amplification (WO2004/067726).

Sequence variation that reflects the methylation status at CpG dinucleotides in the original genomic DNA offers two approaches to PCR primer design. In the first approach, the primers do not themselves cover or hybridize to any potential sites of DNA methylation; sequence variation at sites of differential methylation are located between the two primers. Such primers are used in bisulfite genomic sequencing, COBRA, Ms-SNuPE. In the second approach, the primers are designed to anneal specifically with either the methylated or unmethylated version of the converted sequence. If there is a sufficient region of complementarity, *e.g.,* 12, 15, 18, or 20 nucleotides, to the target, then the primer may also contain additional nucleotide residues that do not interfere with hybridization but may be useful for other manipulations. Exemplary of such other residues may be sites for restriction endonuclease cleavage, for ligand binding or for factor binding or linkers or repeats. The oligonucleotide primers may or may not be such that they are specific for modified methylated residues.

One way to distinguish between modified and unmodified DNA is to hybridize oligonucleotide primers which specifically bind to one form or the other of the DNA. After primer hybridization, an amplification reaction can be performed. The presence of an amplification product indicates that a sample hybridized to the primer. The specificity of the primer indicates whether the DNA had been modified or not, which in turn indicates whether the DNA had been methylated or not. For example, bisulfite

ions convert non-methylated cytosine bases to uracil bases. Uracil bases hybridize to adenine bases under hybridization conditions. Thus an oligonucleotide primer which comprises adenine bases in place of guanine bases would hybridize to the bisulfite-modified DNA, whereas an oligonucleotide primer containing the guanine bases would hybridize to the non-converted (initial methylated) cytosine residues in the modified DNA. Amplification using a DNA polymerase and a second primer yield amplification products which can be readily observed. This method is known as MSP (Methylation Specific PCR; Patent Nos 5,786,146; 6,017,704; 6,200,756). Primers are designed to anneal specifically with the converted sequence representing either the methylated or the unmethylated version of the DNA. Preferred primers and primer sets for assessing the methylation status of the concerned gene by way of MSP will specifically hybridize to a converted sequence provided in Table 2, or to its complement sequence. Most preferred primers and primer sets are provided in Table 1 and are represented by SEQ ID NO. 1 to 264. Sense primers comprise or consist essentially of SEQ ID NO. 1 to 132, antisense primers consist essentially of SEQ ID NO. 133 to 264. The amplification products can be optionally hybridized to specific oligonucleotide probes which may also be specific for certain products. Alternatively, oligonucleotide probes can be used which will hybridize to amplification products from both modified and non-modified DNA.

Thus, present invention provides for a method for identifying cervical cancer or its precursor, or predisposition to cervical cancer in a test sample containing cervical cells or nucleic acids from cervical cells comprising: contacting a methylated CpG-containing nucleic acid of at least one gene selected from the group consisting of genes according to Table 1 with bisulfite to convert unmethylated cytosine to uracil; detecting the methylated CpGs in the nucleic acid by contacting the converted nucleic acid with oligonucleotide primers whose sequence discriminates between the bisulfite-treated methylated and unmethylated version of the converted nucleic acid; and identifying the test sample as containing cells that are neoplastic, precursor to neoplastic, or predisposed to neoplasia, or as containing nucleic acids from cells that are neoplastic, precursor to neoplastic, or predisposed to neoplasia.

Modified and non-modified DNA can be distinguished with use of oligonucleotide probes which may also be specific for certain products. Such probes can be hybridized directly to modified DNA or to amplification products of modified DNA. Probes for assessing the methylation status of the concerned gene will specifically hybridize to the converted sequence but not to the corresponding non converted sequence. Probes

are designed to anneal specifically with the converted sequence representing either the methylated or unmethylated version of the DNA. Preferred converted sequences are provided in Table 2. Preferred probes anneal specifically with the converted sequence representing the methylated version of the DNA, or to the complement sequence thereof. Oligonucleotide probes can be labeled using detection systems known in the art. These include but are not limited to fluorescent moieties, radioisotope labeled moieties, bioluminescent moieties, luminescent moieties, chemiluminescent moieties, enzymes, substrates, receptors, or ligands.

Another way for the identification of methylated CpG dinucleotides utilizes the ability of the MBD domain of the McCP2 protein to selectively bind to methylated DNA sequences (Cross et al, 1994; Shiraishi et al, 1999). Restriction endonuclease digested genomic DNA is loaded onto expressed His-tagged methyl-CpG binding domain that is immobilized to a solid matrix and used for preparative column chromatography to isolate highly methylated DNA sequences. Variants of this method have been described and may be used in present methods of the invention.

Real time chemistry allows for the detection of PCR amplification during the early phases of the reactions, and makes quantitation of DNA and RNA easier and more precise. A few variants of real-time PCR are well known. They include Taqman® (Roche Molecular Systems), Molecular Beacons®, Amplifluor® (Chemicon International) and Scorpion® DzyNA®, Plexor™ (Promega) etc. The TaqMan® system and Molecular Beacon® system have separate probes labeled with a fluorophore and a fuorescence quencher. In the Scorpion® system the labeled probe in the form of a hairpin structure is linked to the primer.

Quantitation in real time format may be on an absolute basis, or it may be relative to a methylated DNA standard or relative to an unmethylated DNA standard. The absolute copy number of the methylated marker gene can be determined; or the methylation status may be determined by using the ratio between the signal of the marker under investigation and the signal of a reference gene with a known methylation (e.g. β-actin), or by using the ratio between the methylated marker and the sum of the methylated and the non-methylated marker.

Real-Time PCR detects the accumulation of amplicon during the reaction, but alternatively end-point PCR fluorescence detection techniques may be used. Confirming the presence of target DNA at the end point stage may indeed be sufficient and it can use the same approaches as widely used for real time PCR.

DNA methylation analysis has been performed successfully with a number of techniques which are also applicable in present methods of the invention. These include the MALDI-TOFF, MassARRAY (Ehrich, M. *et al.* 2005), MethyLight (Trinh B. *et al.* 2001), Quantitative Analysis of Methylated Alleles (Zeschnigk M. *et al.* 2004), Enzymatic Regional Methylation Assay (Galm *et al.*, 2002), HeavyMethyl (Cottrell, SE *et al.*, 2004), QBSUPT, MS-SNuPE (Gonzalgo and Jones, 1997), MethylQuant (Thomassin H. *et al.* 2004), Quantitative PCR sequencing, and Oligonucleotide-based microarray systems (Gitan RS *et al.* ,2006).

The number of genes whose modification is tested and/or detected can vary: one, two, three, four, five, six, seven, eight, nine or more genes according to Table 1 can be tested and/or detected. Detection of epigenetic modification of at least one, two, three, four, five, six, seven, eight, nine or more genes according to Table 1 can be used as an indication of cancer or pre-cancer or risk of developing cancer. The genes are preferably selected from the group of JAM3, LMX1A, CDO1, NID2, CCNA1, HOXA11, GREM1 and TAC1. Preferred gene combinations include

- NID2 and HOXA11;
- JAM3, CDO1, HOXA11, and CCNA1;
- JAM3 and HOXA11;
- JAM3, HOXA11 and GREM1;
- JAM3, NID2, HOXA11 and CDO1;
- JAM3, TAC1, HOXA11, and CDO1;
- JAM3, HOXA11, and CDO1;
- JAM3 and CDO1;
- JAM3 and NID2;
- NID2 and CDO1;
- JAM3 and LMX1A
- NID2 and LMX1A, and
- JAM3, CDO1 and NID2.

Of course, as appropriate, the skilled person would appreciate that functionally relevant variants of each of the gene sequences may also be detected according to the methods of the invention. For example, the methylation status of a number of splice variants may be determined according to the methods of the invention. Variant sequences preferably have at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%,

or at least 99% nucleotide sequence identity with the nucleotide sequences in the database entries. Computer programs for determining percentage nucleotide sequence identity are available in the art, including the Basic Local Alignment Search Tool (BLAST) available from the National Center for Biotechnology Information.

It is possible for the methods of the invention to be used in order to detect more than one gene of interest in the same reaction. Through the use of several specific sets of primers, amplification of several nucleic acid targets can be performed in the same reaction mixture. This may be termed "multiplexing". Multiplexing can also be utilized in the context of detecting both the gene of interest and a reference gene in the same reaction.

The term "test sample" refers to biological material obtained from a mammalian subject, preferably a human subject, and may be any tissue sample, body fluid, body fluid precipitate, or lavage specimen. Test samples for diagnostic, prognostic, or personalized medicine uses can be obtained from cytological samples, from surgical samples, such as biopsies, cervical conization or hysterectomy, from (formalin fixed) paraffin embedded cervix or other organ tissues, from frozen tumor tissue samples, from fresh tumor tissue samples, from a fresh or frozen body fluid such as blood, serum, lymph, or from cervical scrapings, cervical smears, cervical washings and vaginal excretions. Such sources are not meant to be exhaustive, but rather exemplary. A test sample obtainable from such specimens or fluids includes detached tumor cells and/or free nucleic acids that are released from dead or damaged tumor cells. Nucleic acids include RNA, genomic DNA, mitochondrial DNA, single or double stranded, and protein-associated nucleic acids. Any nucleic acid specimen in purified or non-purified form obtained from such specimen cell can be utilized as the starting nucleic acid or acids. The test samples may contain cancer cells or pre-cancer cells or nucleic acids from them. Preferably, the test sample contains squamous cell carcinomas cells or nucleic acids from squamous cell carcinomas, adenocarcinoma cells or nucleic acids of adenocarcinoma cells, adenosquamous carcinoma cells or nucleic acids thereof. Samples may contain mixtures of different types and stages of cervical cancer cells.

Present invention also relates to screening protocols for the screening of woman for cervical cancer and the precursors thereof. Traditionally the Pap Smear has been the primary screening method for the detection of abnormality of the cervix, but its performance is suboptimal. Human Papillomavirus has been associated with the development of cervical cancer. Five high-risk types, 16, 18, 31, 45, and 58, and in particular HPV types 16 and 18 account for approximately 70% of all cervical

carcinomas. A small percentage of women showing persistent infection progress from Low-grade to High-grade lesions. The introduction of methylation markers now adds a new dimension to the screening for and treatment of cervical lesions. Method for cervical cancer screening may combine high-risk human papillomavirus (hr-HPV) testing and methylation testing; or cytological evaluation and methylation testing; or hr-HPV testing and cytological evaluation and methylation testing.

Thus, a further embodiment of the present invention relates to a method for cervical cancer detection or screening comprising the steps of:

a) providing a test sample comprising cervical cells or nucleic acids from cervical cells;
b) assaying the test sample of step a) for high-risk human papillomavirus (hr-HPV);
c) if b) is positive for the presence of hr-HPV, assaying the methylation status of at least one gene selected from the group consisting of genes according to Table 1;
d) if the gene of c) is methylated, refer the woman for colposcopy;
e) if the gene of c) is unmethylated, refer the woman to a more regular screening for the presence of hr-HPV.

The present invention relates further to a method for cervical cancer detection or screening comprising the steps of:

a) providing a test sample comprising cervical cells or nucleic acids from cervical cells;
b) assaying the test sample of step a) for hr-HPV ;
c) if b) is positive for the presence of hr-HPV, assaying the methylation status of at least one gene selected from the group consisting of genes according to Table 1, and/or typing the hr-HPV for the presence of HPV16 and/or HPV18;
d) if the gene of c) is methylated, and/or HPV16 and/or HPV18 positive, refer the woman for colposcopy;
e) if the gene of c) is unmethylated, refer the woman to a more regular screening for the presence of hr-HPV.

In a related embodiment, the invention provides for a method for cervical cancer detection or screening comprising the steps of:

a) performing cytology evaluation on a test sample comprising cervical cells or nucleic acids from cervical cells;
b) if a) is positive, assaying the methylation status of at least one gene selected from the group consisting of genes according to Table 1;
c) if the at least one gene of b) is methylated, refer the woman for colposcopy;

d) if the at least one gene of b) is unmethylated, refer the woman to cytology testing on a more regular basis.

In a related embodiment, the invention provides for a method for cervical cancer detection or screening comprising the steps of:

a) assaying the methylation status of at least one gene selected from the group consisting of genes according to Table 1;
b) if the at least one gene of b) is methylated, perform cytology testing;
c) if b) is tested positive, refer the woman for colposcopy;
d) if b) is negative, refer the woman to methylation testing on a more regular basis.

In all aspects of the invention, the test sample is preferably a cervical, cervicovaginal or vaginal sample of a woman.

The phrase "cervical cancer screening" refers to organized periodic procedures performed on groups of people for the purpose of detecting cervical cancer.

The phrase "assaying for hr-HPV" refers to testing for the presence of hr-HPV. There are various PCR based assays commercially available to measure hr-HPV copy number or viral load in clinical samples. Many testing methods have been used to detect the presence of HPV in cervicovaginal specimens, including viral load quantification, Southern blot, polymerase chain reaction (PCR), ViraPap (Life Technologies, Gaithersburg, MD), Hybrid Capture tube testing, Hybrid Capture microtiter plate assays, and CISH. For instance, assaying for hr-HPV may be performed with the FDA approved Hybrid Capture II assay (Digene Corp., Silver Spring, Md.) with a probe cocktail for 13 carcinogenic types.

The so-called "high risk" HPV types are those strains of HPV more likely to lead to the development of cancer, while "low-risk" viruses rarely develop into cancer. The list of strains considered high risk is being adapted with the time and the increase in epidemiological knowledge. As such, those hr-HPV types comprise, without being limited to, strains 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, and 69. Preferred "high risk" HPV types are HPV16 and HPV18.

The phrase "HPV16 testing" refers to testing for the presence of hr-HPV type 16. Similarly, "HPV18 testing" refers to testing for the presence of hr-HPV type 18. The various methods allowing type-specific HPV testing are well known to the person skilled in the art and are applicable in the methods of present invention. For instance,

testing for the presence of hr-HPV-16 may be accomplished by PCR amplification using primers specific for HPV type 16, which are known by the skilled in the art.

The phrase "performing cytological evaluation" refers to the cytomorphological assessment of cervical samples, which is usually performed by especially trained medical staff. The various methods allowing cytological testing are well known to the person skilled in the art and are applicable in the methods of present invention. Cytological evaluation may be performed with the known Papanicolaou (PAP) smear test. Alternative means for cytological evaluation include liquid based cytology with for example the ThinPrep technique (Cytyc Corporation, Marlborough, Massachusetts, USA).

The term "triaging" refers to sorting out or classifying patients in order to establish priority of treatment's necessity, priority of proper place of treatment, or any other priority in terms of patient management.

The test sample will most of the time be obtained from a subject suspected of being tumorigenic or from a subject undergoing routine examination and not necessarily being suspected of having a disease. Alternatively the sample is obtained from a subject undergoing treatment, or from patients being checked for recurrence of disease.

Testing can be performed diagnostically or in conjunction with a therapeutic regimen. Testing can be used to monitor efficacy of a therapeutic regimen, whether a chemotherapeutic agent or a biological agent, such as a polynucleotide. Epigenetic loss of function of at least one gene selected from the group consisting of genes according to Table 1 can be rescued by the use of DNA demethylating agents and/or DNA methyltransferase inhibitors. Testing can also be used to determine what therapeutic or preventive regimen to employ on a patient. Moreover, testing can be used to stratify patients into groups for testing agents and determining their efficacy on various groups of patients.

Demethylating agents can be contacted with cells *in vitro* or *in vivo* for the purpose of restoring normal gene expression to the cell. Suitable demethylating agents include, but are not limited to 5-aza-2'-deoxycytidine, 5-aza-cytidine, Zebularine, procaine, and L-ethionine. This reaction may be used for diagnosis, for determining predisposition, and for determining suitable therapeutic regimes. Accordingly, the invention also provides for a method for predicting the likelihood of successful treatment or resistance to treatment of cancer with such agent. If the gene is methylated, the

likelihood of successful treatment is higher than if the gene is unmethylated, or methylated to a lesser degree. Conversely, if the gene is unmethylated, or methylated to a lesser degree, the likelihood of resistance to treatment is higher than if the gene is methylated.

In a related aspect, epigenetic loss of gene function(s) can identify the stage of the disease and from that the need of treatment. Accordingly, the invention provides for a method for predicting suitable treatment comprising determining the methylation status of a gene or a combination of genes. If the gene is methylated, the need of cervical resection is identified; if the gene is unmethylated or methylated to a lesser degree, it is decided that there is no need for cervical resection. In cases of early stage (CIN) and carcinoma in situ, abnormal tissue is removed by cryosurgery, laser surgery, conization, or simple hysterectomy (removal of the uterus). Invasive cervical cancer is treated with radical hysterectomy (removal of the uterus, fallopian tubes, ovaries, adjacent lymph nodes, and part of the vagina).

To attain high rates of tumor detection, it may be necessary to combine the methods of the invention with established methods and/or markers for cervical cancer identification (Malinowski D, 2007), such as morphology-based detection methods, HPV methylation testing (Badal *et al.* 2004, Kalantari *et al.* 2004), KRAS and BRAF mutation detection (Kang *et al.* 2007), chromosomal amplification (Rao *et al.* 2004), protein expression (Keating *et al.* 2001)and HPV detection methods (Brink *et al.* 2007): several HPV detection kits are known in the art and commercially available, for example kits such as Digene® HPV Test (Qiagen), AMPLICOR HPV Test (Roche), HPV High-Risk Molecular Assay (Third Wave Technologies), LINEAR ARRAY HPV Genotyping Test (Roche), INNO-LiPA HPV Genotyping (Innogenetics), PapilloCheck (Greiner Bio-One GmbH), PreTect HPV-Proofer (Norchip), NucliSENS EasyQ HPV (BioMérieux), F-HPV typing™ (molGENTIX, S.L.) may be utilized. Such examples are not meant to be exhaustive, but rather exemplary.

Another aspect of the invention is a kit for assessing methylation in a test sample. Kits according to the present invention are assemblages of reagents for testing methylation. They are typically in a package which contains all elements, optionally including instructions. The package may be divided so that components are not mixed until desired. Components may be in different physical states. For example, some components may be lyophilized and some in aqueous solution. Some may be frozen. Individual components may be separately packaged within the kit. The kit may contain reagents, as described above for differentially modifying methylated and non-

methylated cytosine residues. Desirably the kit will contain oligonucleotide primers which specifically hybridize to regions within about 10 kbp, within about 5 kbp, within about 3 kbp, within about 1 kbp, within about 750 bp, within about 500 bp, within 200 bp or within 100 bp kb of the transcription start sites of the genes/markers listed in Table 1.

Typically the kit will contain both a forward and a reverse primer for a single gene or marker. If there is a sufficient region of complementarity, *e.g.,* 12, 15, 18, or 20 nucleotides, then the primer may also contain additional nucleotide residues that do not interfere with hybridization but may be useful for other manipulations. Exemplary of such other residues may be sites for restriction endonuclease cleavage, for ligand binding or for factor binding or linkers or repeats. The oligonucleotide primers may or may not be such that they are specific for modified methylated residues. The kit may optionally contain oligonucleotide probes. The probes may be specific for sequences containing modified methylated residues or for sequences containing non-methylated residues. The kit may optionally contain reagents for modifying methylated cytosine residues. The kit may also contain components for performing amplification, such as a DNA polymerase and deoxyribonucleotides. Means of detection may also be provided in the kit, including detectable labels on primers or probes. Kits may also contain reagents for detecting gene expression for one of the markers of the present invention. Such reagents may include probes, primers, or antibodies, for example. In the case of enzymes or ligands, substrates or binding partners may be sued to assess the presence of the marker. Kits may contain 1, 2, 3, 4, or more of the primers or primer pairs of the invention. Kits that contain probes may have them as separate molecules or covalently linked to a primer for amplifying the region to which the probes hybridize. Other useful tools for performing the methods of the invention or associated testing, therapy, or calibration may also be included in the kits, including buffers, enzymes, gels, plates, detectable labels, vessels, etc.

According to a further aspect, the invention also employs or relies upon or utilizes oligonucleotide primers and/or probes to determine the methylation status of at least one gene selected from a group of genes consisting of JAM3, LMX1A, CDO1, NID2, ALX3, ALX4, AR, ARID4A, ATM, AURKA, B4GALT1, BMP2, BMP6, BNIP3, C13orf18, C16orf48, C9orf19, CALCA, CAMK4, CCNA1, CCND2, CDH1, CDH4, CDK6, CDKN1B, CDKN2B, CLSTN2, CLU, COL1A1, CPT1C, CTDSPL, CYCLIND2, DAPK1, DBC1, DDX19B, DKK2, EGFR, EGR4, EPB41L3, FOS, FOXE1, GADD45A, GATA4, GDAP1L1, GNB4, GPNMB, GREM1, Gst-Pi, HHIP,

HIN1, HOOK2, HOXA1, HOXA11, HOXA7, HOXD1, IGSF4, ISYNA1, JPH3, KNDC1, KRAS, LAMA1, LOC285016, LOX, LTB4R, MAL, MTAP, MYO18B, NDRG2, NOL4, NPTX1, NPTX2, OGFOD2, PAK3, PAX1, PDCD4, PHACTR3, POMC, PRKCE, RAD23B, RALY, RARA, RASSF1A, RBP4, RECK, RPRM, SALL4, SEMA3F, SLC5A8, SLIT1, SLIT2, SLIT3, SMPD1, SOCS1, SOX1, SOX17, SPARC, SPN, SST, TAC1, TERT, TFPI-2, TLL1, TNFAIP1, TRMT1, TWIST1, UGT1A1, WIF1, WIT1, WT1, XRCC3, and ZGPAT. Preferred probes and their sequences bind to at least one of the polynucleotide sequences listed in Table 2, Figure 5B or to the complement sequence thereof. Preferred primers and probes are selected from the primers and probes comprising or consisting essentially of the nucleotide sequences set forth in Table 1. Related to this, the invention also provides for an isolated polynucleotide which consists of a nucleotide sequence listed in Table 1, Table 2 and Figure 5B.

A more complete understanding can be obtained by reference to the following specific examples which are provided herein for purposes of illustration only, and are not intended to limit the scope of the invention.

Table 1: MSP assays and primer design

Row Nr	Assay Name	Gene ID	Official Gene Symbol	Refseq	Sense Primer sequence (5' – 3') SEQ ID NO's 1-132	Antisense Primer sequence (5' – 3') SEQ ID NO's 133-264
1	ALX3_25178	257	ALX3	NM_006492	GTTTGGTTCGGGTTAGCGT	CCTACTTATCTCTCCCGCTCG
2	ALX3_25180	257	ALX3	NM_006492	TTGCGTTTTATTTGTATTTCGC	CTTAACGAACGACTTAACCGACT
3	ALX4_25062	60529	ALX4	NM_021926	TTTTATTGCGAGTCGTCGGTC	TATACCGAACTTATCGCCTCCG
4	AR_24818	367	AR	NM_000044, NM_001011645	TGTATAGGAGTCGAAGGGACGTA	AAACAACTCCGAACGACGA
5	ARID4A_24110	5926	ARID4A	NM_002892	GTTAGGTAAGTGGTACGGCGA	AAAAACGACTACAACTACGACGA
6	ARID4A_24112	5926	ARID4A	NM_002892	ATTTAATGAGGACGGTAGGTAGC	AACAAACTCGCTTCTACACGAA
7	ATM_9746	472	ATM	XM_940791, NM_000051, NM_138292	TTTAATATAAGTCGGGTTACGTTCG	ATACGACGCAAAAACTATCGC
8	AURKA_24802	6790	AURKA	NR_001587, NM_003600	TTAGGGAGTAAGTGCGTTTGC	AAAAACCGATTAACCTACGCTC
9	B4GALT1_1	2683	B4GALT1	NM_001497	TAGACGGTTACGAGTAGGCGGTA	CCTTCTTAAAACGACGACGAA
10	B4GalT1_3	2683	B4GALT1	NM_001497	TTTTTCGTATTTTAGGAAGTGGC	TTCCTCCCGAACCTTTACGA
11	BMP2_17901	650	BMP2	NM_001200	TTTGGGGTTCGATTATATTTC	CGAAAACTCCGAAACCGAT
12	BMP6_24310	654	BMP6	NM_001718	GTTATTTTTCGGCGGGTTC	CTAATAATCGCCCCTTCGC
13	BNIP3	664	BNIP3	NM_004052	TACGCGTAGGTTTTAAGTCGC	TCCCGAACTAAACGAAACCCCG

14	C13orf18_19885	80183	C13orf18	NM_025113	TTTGATTTTTGAAAGC GTCGT	ACACCACGCACCTATA CGC
15	C13orf18_Gron	80183	C13orf18	NM_025113	TTTTTAGGGAAGTAA AGCGTCG	ACGTAATACTAAACCC GAACGC
16	C16orf48_22922	84080	C16orf48	NM_032140	TAGTTTGGTAGTTAG CGGGTC	AAACCTCCGAAATAAC CGTC
17	C9orf19_19865	152007	C9orf19	NM_022343	ATAGGGGGAGTTCG GTACG	ACAATTTACCCCGCTC GACT
18	CALCA_2	796	CALCA	NM_001033952 NM_001033953 NM_001741	CGTTTTTATAGGGTTT TGGTTGGAC	AAATCTCGAAACTCAC CTAACGA
19	CAMK4_27356	814	CAMK4	NM_001744	TAGTTGTATCGGTTT AGGCGTTT	CTACCTTCGTACCCTT CGATT
20	CCNA1_gron	8900	CCNA1	NM_003914	GTTATGGCGATGCGG TTTC	CCAACCTAAAAAACGA CCGA
21	CCND2_25209	894	CCND2	NM_001759	GAAGGTAGCGTTTTT CGATG	AAATAAACCCGATCCG CAA
22	CDH1_17968	999	CDH1	NM_004360	AATTTTAGGTTAGAG GGTTATCGC	ACCAATCAACAACGCG AAC
23	CDH1_23527	999	CDH1	NM_004360	GAGGGGGTAGGAAA GTCGC	CGAAACGACCTAAAAA CCTCG
24	CDH4_24735	1002	CDH4	NM_001794	GGGACGATTTTTCGT TGTTC	TTCTACTACTCTCGCTC TCCGAC
25	CDK6_9703	1021	CDK6	NM_001259	AATTTCGTTTGTAGA GTCGTCGT	TCTATATTAAAAACTTC GCTTCG
26	CDKN1B_23172	1027	CDKN1B	NM_004064	GTCGGTAAGGTTTGG AGAGC	AAAATAACAAAACCCG TCCG
27	CDKN2B_27345	1030	CDKN2B	NM_004936	TTAGAAGTAATTTAG GCGCGTTC	AAACCCCGTACAATAA CCGA
28	CDO1_55928	1036	CDO1	NM_001801	AATTTGATTTGTGTGT GTATCGC	GAAACGTAAAAATATC GTCGCA
29	CDO1_55929	1036	CDO1	NM_001801	GTTTACGCGATTTTT GGGAC	AAAAACCCTACGAACA CGACT
30	CLSTN2_19850	64084	CLSTN2	NM_022131	AGGGTTTTTCGGAGT CGTT	TTCCTCAACCGTCTCC ACG
31	CLU_13810	1191	CLU	NM_001831, NM_203339	AGGCGTCGTATTTAT AGCGTTT	TCCCCTTACTTTCCGC GAC
32	CLU_19838	1191	CLU	NM_001831, NM_203339	GTGGGGGTCGGTGT AGTATC	TCCCTACTAAAAACGC CGAA
33	COL1A1_23253	1277	COL1A1	NM_000088	TATAAAAGGGGTTCG GGTTAGTC	AAATTAACGTCCGCTC ATACG
34	CPT1C_23912	126129	CPT1C	NM_152359	AGGAAGTATTTATTG CGTATGTTTC	CCATCACTTATCCTCG ACGC
35	CTDSPL_23795	10217	CTDSPL	NM_001008392	TAATTTTAAGGAGGA CGAGGGTC	ATAAACTCCAACGACG CGAAA
36	CTDSPL_23804	10217	CTDSPL	NM_005808	GTTTTGGGAGAGGC GGTTC	TCATAATAACGAAACG ACGACC
37	CYCLIND2_1 (official full gene name for CCND2)	894	CCND2	NM_001759	GGTGTAGCGTTTAGG GTCGTC	CGAATTTTTCCTACGTA ACCG
38	DAPK1	1612	DAPK1	NM_004938	GGATAGTCGGATCGA GTTAACGTC	CCCTCCCAAACGCCGA
39	DBC1_23879	1620	DBC1	NM_014618	AGGATAGGTATGAAT TTCGGTTTC	AAACGAACGAACAACA ACGA
40	DDX19B_22963	11269	DDX19B	NM_007242, NM_001014449, NM_001015047	CGGGTTTGAGGGTAA TAGAATCG	CGCCACAATAACGTCG AAA
41	DKK2_23970	27123	DKK2	NM_014421	GTGCGGGGTAAGAA GGAAC	AAAAACAATCAAATAC GAAACGC
42	DKK2_23973	27123	DKK2	NM_014421	GAGAGAGAAAGCGG GAGTTC	TCACAATTACCCCGAA ACG
43	EGFR_23302	1956	EGFR	NM_201283, NM_005228	TAGGAGCGTTGTTTC GGTC	CACGACCCCCTAACTC CGT
44	EGR4_24277	1961	EGR4	NM_001965	TTTAGGTGGGAAGCG TATTTATC	AAACGCTAAAACCGCG AAT
45	EPB41L3_19071	23136	EPB41L3	NM_012307	GGGATAGTGGGGTT GACGC	ATAAAAATCCCGACGA ACGA
46	EPB41L3_19072	23136	EPB41L3	NM_012307	GCGTGGGTTTTCGTC GTAG	CCCAAAACTACTCGCC GCT

47	FOS_22338	2353	FOS	NM_005252	CGGGTTGTAGTTAATATCGAGG	CTCTCTCATTCTACGCCGTTC
48	FOXE1_13314	2304	FOXE1	NM_004473	TTTGTTCGTTTTTCGATTGTTC	TAACGCTATAAAACTCCTACCGC
49	GADD45A_24463	1647	GADD45A	NM_001924	CGTTAATCGGATAAGAGTGCG	AAAACCACGCGAAAAACGA
50	GATA4	2626	GATA4	NM_002052	GTATAGTTTCGTAGTTTGCGTTTAGC	AACTCGCGACTCGAATCCCCG
51	GATA4_13295	2626	GATA4	NM_002052	GGTATTGTTATTTTGCGTTTTC	CCCGAAACAAACTACACGAC
52	GDAP1L1_19773	78997	GDAP1L1	NM_024034	GATTTCGGGTTGTTATGGC	CTAACTTAACCGCATCGCTC
53	GDAP1L1_19775	78997	GDAP1L1	NM_024034	GAAAGAAGGAGGTTTCGGC	CCCGATAAATAATAACATTCACGA
54	GNB4	59345	GNB4	NM_021629	GTTGTGAGTTGCGTTTTTTACGTC	CGCTACCGATATCCGCTAAACG
55	GPNMB_52607	10457	GPNMB	NM_001005340	GGTCGTAGTCGTAGTCGGG	CCGCAAAAACCTAAAACGTAA
56	GREM1_29777	26585	GREM1	NM_013372	GAATTTGGTACGATTTTACGGAG	ATCTAAACTTTCCCTATCGACCG
57	Gst-Pi_New3	2950	GSTP1	NM_000852	ATTTAGTATTGGGGCGGAGC	TAACGAAAACTACGACGACGA
58	HHIP_23319	64399	HHIP	NM_022475	AGTAGTAGGAATAGAAACGGCGA	AAAACTACAACCGCCGACA
59	HIN1_3	92304	SCGB3A1	NM_052863	GAAGTTGGTTAGGGTACGGTC	AACTTCTTATACCCGATCCTCG
60	HOOK2_19741	29911	HOOK2	NM_013312	GGATCGTTGGATTTTGGTTC	TATATCCTCGCCCCACGTAA
61	HOXA1_27316	3198	HOXA1	NM_153620	TTTTTAGAGTAAATAGCGGGAGC	ATACGCCTTACAACCCCTACG
62	HOXA11_23844	3207	HOXA11	NM_005523	TTTTATTTATTCGGGGAGTTGC	ACAAAATCCTCGTTCTCGAAT
63	HOXA7_2	3204	HOXA7	NM_006896	TCGTAGGGTTCGTAGTCGTTT	TCCAAATCTTTTTCCGCGA
64	HOXD1(2)	3231	HOXD1	NM_024501	GTCGGTTGACGTTTTGAGATAAGTC	ACCGTCTTCTCGAACGACG
65	IGSF4_18987	23705	CADM1	NM_014333	TCGGATTTCGTTTTTAGCGTAT	GAACACCTACCTCAAACTAACGAC
66	ISYNA1_19726	51477	ISYNA1	NM_016368	TAGGTTGGTTTGGTTTCGGTC	TAAACGACGACCTCCATCG
67	JAM3	83700	JAM3	NM_032801	GGGATTATAAGTCGCGTCGC	CGAACGCAAAACCGAAATCG
68	JPH3_12611	57338	JPH3	NM_020655	TTAGATTTCGTAAACGGTGAAAAC	TCTCCTCCGAAAAACGCTC
69	KNDC1_19691	85442	KNDC1	NM_033404 NM_152643	TGGATGGAGTTTAGGTTATATCGTC	AAAATACTACGAAACCGCCC
70	KRAS_24235	3845	KRAS	NM_033360	AGGAGGGATTGTCGGATTTAC	GCTCCGAATCAAAATTAACGA
71	LAMA1_63431	284217	LAMA1	NM_005559	TTTTTAGATTTATCGAGTGGCG	CGAACTCACCTCTCTACCGAC
72	LMX1A_9513	4009	LMX1A	NM_177398, NM_177399, NM_001033507	CGGTATCGTTGTTTAGGAGGC	CGTATAACTATTACCTCGAAACGCT
73	LOC285016_22940	285016	hCG_1990170	NM_001002919	AGTTGTTTGGTATTCGCGGT	CGACCCCTCCTAACTTTCG
74	LOX_23395	4015	LOX	NM_002317	GTTAGATTGATTTCGTTCGAGG	AACTAAAATACCCGTACTCCGCT
75	LTB4R_31250	1241	LTB4R	NM_181657	TAGTAGATTTTTAGCGGTGAAGACG	AAAACCTTAACGAAACTAAACGAAA
76	MAL	4118	MAL	NM_002371	TTCGGGTTTTTTTGTTTTAATTC	GAAAACCATAACGACGTACTAACG
77	MTAP_24628	4507	MTAP	NM_002451	GTAAGTGAGTTTCGAGTGTCGC	CTCCGAAAACCATACGCCC
78	MYO18B_24620	84700	MYO18B	NM_032608	GAAAGGTCGGATTTGTTTTTC	ACCATCTCATCACGCCTCG
79	NDRG2_56603	57447	NDRG2	NM_201540 NM_201539 NM_201535 NM_201537	AGATTTTGTGGTTTCGTCGTT	ATCCCCCGAACATTACGATT
80	NID2_9091	22795	NID2	NM_007361	GCGGTTTTTAAGGAGTTTTATTTTC	CTACGAAATTCCCTTTACGCT

81	NOL4_19645	8715	NOL4	NM_003787	GAGAGATTCGGGATTCGTG	GTAATCCAAAAATAAAAACTACGCC
82	NPTX1_2	4884	NPTX1	NM_002522	AGTACGTTGTTTCGGAGTTTTTC	CTTCATCTACACCTCGATACCCG
83	NPTX2_57779	4885	NPTX2	NM_002523	GCGTCGTTTTGTATGGGTATC	CCCGATAACCGCTTCGTAT
84	OGFOD2_23131	79676	OGFOD2	NM_024623	CGAGTAGTAGTTGCGTCGGG	ACAAACGACCCTAAAAACGAAC
85	PAK3_1	5063	PAK3	NM_002578	TGTATGATTTTAGTTCGCGGAT	ACGAATTTTACCTCAAACGACC
86	PAK3_3	5063	PAK3	NM_002578	GCGGGATTTATTTGTTACGGA	AACCCGAAACTACGACTACGAC
87	PAX1_27210	5075	PAX1	NM_006192	ATTGCGTCGGGTTTAGTTTC	GCCCCTTACCCATAACGAAC
88	PAX1_27211	5075	PAX1	NM_006192	GTTTAGGGAAAGCGGACGA	GAACGACAAACAAAACTCGAAA
89	PDCD4_11827	27250	PDCD4	NM_145341, NM_014456	GTTCGTAGTTCGGGGCGTT	GCGATCCTATCAAATCCGAA
90	PHACTR3_11692	116154	PHACTR3	NM_080672 NM_183244 NM_183246	TTATTTTGCGAGCGGTTTC	GAATACTCTAATTCCACGCGACT
91	POMC	5443	POMC	NM_000939	GATTTGGGCGTTTTTGGTTTTTCGC	GACTTCTCATACCGCAATCG
92	PRKCE_24134	5581	PRKCE	NM_005400	GTGGGTTTTAAGTTTACGGTTTC	CCTACCCTCGAAACAAACGA
93	RAD23B_1	5887	RAD23B	NM_002874	GGCGGAGTTTGTATAGAGGC	AACCCGAATTACGCAAACG
94	RALY_19607	22913	RALY	NM_007367	TTTTTGGGTTTCGTTGTTTC	CGCCTCAATAATACCGACC
95	RARA_24121	5914	RARA	NM_001024809	TTCGTTTCGTTTAGGTATCGTTT	CCTCTCGATTCCCTACGTTT
96	RARA_24129	5914	RARA	NM_000964, NM_001033603	TTTAGGATTATAGTGAGCGACGG	TAACCGCCTTTAACCCCGA
97	RASSF1A	11186	RASSF1	NM_007182. NM_170712 NM_170714	GCGTTGAAGTCGGGGTTC	CCCGTACTTCGCTAACTTTAAACG
98	RBP4_24106	5950	RBP4	NM_006744	GGTCGTTTCGTTGTTTTATAGC	GCGTTATACAAATACCCCCG
99	RECK_18940	8434	RECK	NM_021111	TTACGGTTAGTAGAAGGAGTAGCGT	CTACGACCAAACTAAATCCGAAC
100	RPRM_2	56475	RPRM	NM_019845	TCGAGGAAGAAGATGTCGAAG	AAAAACCCGAACGAACGTAA
101	SALL4_12833	57167	SALL4	NM_020436	GAGGCGTAAGTAGGCGAAA	CGCATCTACAAACTCCGAAA
102	SEMA3F_23485	6405	SEMA3F	NM_004186	GATTAGAGCGAGCGAACGA	TAACTACTAAACCCGAACCGAAC
103	SLC5A8_24598	160728	SLC5A8	NM_145913	GGTTTGTTGGTCGTTTTTAGC	CGAAACATCGACACCTTCGT
104	SLC5A8_24601	160728	SLC5A8	NM_145913	GTATTTAGGGTAGCGGGTCG	CGAAATAAAAACTAACAATCGCC
105	SLIT1_23651	6585	SLIT1	NM_003061	GCGTTATGGTGTTTTTATAGCGT	TCTTCGATAACTCTACCCCGA
106	SLIT1_23653	6585	SLIT1	NM_003061	TTGTAGGCGGTTTGTAGTCGT	GACAATCATCCATCAATCGAAA
107	SLIT2_23672	9353	SLIT2	NM_004787	GAGGATCGGTTTAGGTTGC	CAATTCTAAAAACGCACGACT
108	SLIT2_23676	9353	SLIT2	NM_004787	AGGGGAAGACGAAGAGCGT	CACGAACTAACGCTACGCAA
109	SLIT2_23681	9353	SLIT2	NM_004787	TAGCGGAGAGGAGATTACGC	GACCCCTACATCTTAACAACCG
110	SLIT3_23619	6586	SLIT3	NM_003062	AGGGGTATTTATAGGCGTTTAGC	TACCTACTCCGCTACCAACGTAA
111	SMPD1_24061	6609	SMPD1	NM_000543	GAAGGGTAATCGGGTGTTTTC	CTAATTCGTCTATCCCGTCC
112	SOCS1_23595	8651	SOCS1	NM_003745	GATAGGGTTTTGTTTTCGGC	ATTTTACCCCGCTACCTCG
113	SOX1_27153	6656	SOX1	NM_005986	TTGTAGTTTTCGAGTTGGAGGTC	AAAACGATACGCTAAACCCG
114	SOX1_27159	6656	SOX1	NM_005986	GTTAGGAGTTCGTCGGTTAGC	CACCCGAATTACAAATACCGA

115	SOX17_66072	64321	SOX17	NM_22454	GAGATGTTTCGAGGGTTGC	CCGCAATATCACTAAACCGA
116	SPARC_Wis	6678	SPARC	NM_003118	TTTCGCGGTTTTTTAGATTGTTC	CATACCTCAATAACAAACAAACAAACG
117	SPN_24052	6693	SPN	NM_003123, NM_001030288	ATCGTAGGTTGGGTTTGGTC	AAAAACAAAACACGCGAAA
118	SST_23808	6750	SST	NM_001048	TGGTTGCGTTGTTTATCGTTT	TTACCTACTTCCCCGCGAC
119	TAC1_56187	6863	TAC1	NM_003182	GGGTATTTATTGCGACGGAT	CCGACGACAACTACCGAAA
120	TERT_23702	7015	TERT	NM_003219, NM_198255	GGTTTCGATAGCGTAGTTGTTTC	CTACACCCTAAAAACGCGAAC
121	TFPI-2	7980	TFPI2	NM_006528	GTTCGTTGGGTAAGGCGTTC	CATAAAACGAACACCCGAACCG
122	TLL1_24051	7092	TLL1	NM_012464	TAAGGAATTTTGTATTCGGAGGC	ACCTAACAAACTACGAACGCCA
123	TNFAIP1_23212	7126	TNFAIP1	NM_021137	GTGGTTAGCGGATTTCGAGT	AACTAAACAACACTCCGAACGA
124	TRMT1_19794	55621	TRMT1	NM_017722	TTTCGTAGGGTTCGGTGTC	CCGAATACTCTCTAAAACCCGAT
125	TWIST1_3	7291	TWIST1	NM_000474	GTTAGGGTTCGGGGGCGTTGTT	CCGTCGCCTTCCTCCGACGAA
126	TWIST1_9329	7291	TWIST1	NM_000474	TTTAGTTCGTTAGTTTCGTCGGT	TACTACTACGCCGCTTACGTCC
127	UGT1A1_22912	54658	UGT1A1	NM_007120	TTTTGTGGTTAGTCGCGGT	ACGTAAAATAAACAATCAACTATCG
128	WIF1_9096	11197	WIF1	NM_007191	GCGTCGTTAGATATTTTGTTGC	TAACACCCAAACCGAAAAACG
129	WIT1_24567	51352	WIT1	NM_015855	GTATGGAGCGTTTTGCGAT	AACGAATCCACATACCCGA
130	WT1_1	7490	WT1	NM_024426, NM_024424	TGTGTTATATCGGTTAGTTGAGAGC	CGCTACTCCTTAAAAACGCC
131	XRCC3_9322	7517	XRCC3	NM_005432	CGTTTGTTTTTATAGGTTCGGG	ACAACGAAATCGAAAATCGTAA
132	ZGPAT_23961	84619	ZGPAT	NM_032527 NM_181484 NM_181485	TGTATGCGGAGAGGTCGTAG	ACCATTCCCGACTCCTCGT

Table 2: Amplicon details (converted sequences issuing from the methylated version of the DNA)

Row Nr	Assay Name	Gene ID	Official Gene Symbol	Refseq	Amplicon Sequence (converted) (5' – 3') SEQ ID NO's 265-396
1	ALX3_25178	257	ALX3	NM_006492	GTTTGGTTCGGGTTAGCGTTAATTCGGTTTTCGTGGAAGTCGTGGCGAAAGGCGAGAGGGGTAAAAAGTTGAGAAATAGGCGAGCGGGAGAGATAAGTAGG
2	ALX3_25180	257	ALX3	NM_006492	TTGCGTTTTATTTGTATTTCGCGTCGTTTCGCGGTTCGCGGTTGATTCGTTTTTCGGTTTGCGGGTTTTTGGAGTTTTATTTTTTAGAGTCGGTTAAGTCGTTCGTTAAG
3	ALX4_25062	60529	ALX4	NM_021926	TTTTATTGCGAGTCGTCGGTCGTTGTTATGGACGTTTATTATAGTTCGGTGTCGTAGAGTCGGGAGGGTTCGTCGTTTTTTAGGGTATTTTTCGGAGGCGATAAGTTCGGTATA
4	AR_24818	367	AR	NM_000044, NM_001011645	TGTATAGGAGTCGAAGGGACGTATTACGTTAGTTTTAGTTCGGTTTTAGCGATAGTTAACGTTTTTTGTAGCGCGGCGGTTTCGAAGTCGTCGTTCGGAGTTGTTT
5	ARID4A_24110	5926	ARID4A	NM_002892	GTTAGGTAAGTGGTACGGCGAGCGTAAGGGAAGGGGTTAGTTATTGATTAGCGGTAGTAATTGTAGGAATCGTCGTCGTAGTTGTAGTCGTTTTT
6	ARID4A_24112	5926	ARID4A	NM_002892	ATTTAATGAGGACGGTAGGTAGCGAGGTTTTATTCGAAGTTTTTCGGCGTTATGAGTAGTTAATAGGAGTTCGTGTAGAAGCGAGTTTGTT
7	ATM_9746	472	ATM	XM_940791, NM_000051, NM_138292	TTTAATATAAGTCGGGTTACGTTCGAGGGTAATAATATGATTAAAATTATAGTAGGAATTATAATAAGGAATAAGATTTAGGTTAAAGTAAATATAGCGATAGTTTTTGCGTCGTAT

8	AURKA_24802	6790	AURKA	NR_001587, NM_003600	TTAGGGAGTAAGTGCGTTTGCGCGCGGTGTGCGTT TTTAAACGCGATTTAAGGCGTCGGGTTTGTTGTTAAT TAATTATAAGGTAGTTTCGTTCGAGCGTAGGTTAATC GGTTTTT
9	B4GALT1_1	2683	B4GALT1	NM_001497	TAGACGGTTACGAGTAGGCGGTAGGTTCGTTGTAG GGACGCGTTTGGTATCGCGGCGTTGTCGTTTAGGA GCGGTTTTCGAAGTTTTATTTTTTCGTCGTCGTTTTA AGAAGG
10	B4GalT1_3	2683	B4GALT1	NM_001497	TTTTTCGTATTTTAGGAAGTGGCGCGGTTTGTCGAG GGTAGCGTGGAGGAGGAAGAGGAGGCGCGGTTTAA CGCGATCGAAGTTTCGTCGTAAAGGTTCGGGAGGA A
11	BMP2_17901	650	BMP2	NM_001200	TTTGGGGTTCGATTATATTTCGGTTAGCGCGTTTTAG GTTTTCGATTTTTTGTAGTAGGTGTTTCGTATCGCGG CGTTAGGGATCGGTTTCGGAGTTTTCG
12	BMP6_24310	654	BMP6	NM_001718	GTTATTTTTCGGCGGGTTCGTTTTTTTTTTTTGGTTTT TAGTTTTTATTTTTTATGGTCGTTCGGGGCGTTTTTA GTTGTTTAGGTTAGAGAGGTGGCGAAGGGGCGATT ATTAG
13	BNIP3	664	BNIP3	NM_004052	TACGCGTAGGTTTTAAGTCGCGGTTAATGGGCGACG CGGTCGTAGATTCGTTCGGTTTCGTTTTGTTTTGTGA GTTTTTTCGGTCGGGTTGCGGGGTTTCGTTTAGTTC GGGA
14	C13orf18_19885	80183	C13orf18	NM_025113	TTTGATTTTTGAAAGCGTCGTTGCGTTTCGCGTCGC GGGTAGGTAGGGCGGGATTTTTAGGAGGATCGGTA GAGGCGCGTATAGGTGCGTGGTGT
15	C13orf18_Gron	80183	C13orf18	NM_025113	**T**TTTTAGGGAAG**T**AAAG**C**G**TC**GTT**TTC**G**TC**G**T**AGGT ATCGAGACGTCGTTTAGATGGAAGAAATTTTGGAGA TGCGCGTTTTTATATCGGTGTCGCGGCGTTCGGGTT TAGTATTACGT
16	C16orf48_22922	84080	C16orf48	NM_032140	TAGTTTGGTAGTTAGCGGGTCGGGGCGTTTAGTTTT ATTTTTTAGAGCGTTGCGGTTTTGTGTTTGAAGGTTA AATAGTTTGACGGTTATTTCGGAGGTTT
17	C9orf19_19865	152007	C9orf19	NM_022343	ATAGGGGGAGTTCGGTACGGCGCGGGCGTTTAGGA GAGAAGGAATAATAAATGGATGAGGGGGATGTTTAG GGTTGTTTTCGGGATAGTCGAGCGGGGTAAATTGT
18	CALCA_2	796	CALCA	NM_001033952 NM_001033953 NM_001741	CGTTTTTATAGGGTTTTGGTTGGACGTCGTCGTCGT CGTTGTTATCGTTTTTGATTTAAGTTATTTTTCGTTAG GTGAGTTTCGAGATTT
19	CAMK4_27356	814	CAMK4	NM_001744	TAGTTGTATCGGTTTAGGCGTTTTGGTGGGGTGGGA AGGATTCGAGTCGTATTTGAATGAAGGTTAGTTTTTT TTTAAGATATTAATTAGGTAGGGAGAAATCGAAGGG TACGAAGGTAG
20	CCNA1_gron	8900	CCNA1	NM_003914	GTTATGGCGATGCGGTTTCGGAGAGCGTACGTTTGT CGCGGTCGGTATGGAAACGTTTTCGTTAGGTTCGG GGGCGTCGTTGATTGGTCGATTTAATAGACGCGGGT GGGTAGTTTAGTCGTATCGTTAAGTTCGGTCGTTTTT TAGGTTGG
21	CCND2_25209	894	CCND2	NM_001759	GAAGGTAGCGTTTTTCGATGGTGAGTAGGTTTTGTA GGACGCGGTCGTTTCGGAGTAGGTTGCGGTTTCGT ACGGTTTTGCGGATCGGGTTTATTT
22	CDH1_17968	999	CDH1	NM_004360	AATTTTAGGTTAGAGGGTTATCGCGTTTATGCGAGG TCGGGTGGGCGGGTCGTTAGTTTCGTTTTGGGGAG GGGTTCGCGTTGTTGATTGGT
23	CDH1_23527	999	CDH1	NM_004360	GAGGGGGTAGGAAAGTCGCGTTCGTTTTTTATTATT TATTTTTTATTTTTATTATTGGGGGGTTCGGAGCGCG CGAGGTTTTTAGGTCGTTTCG
24	CDH4_24735	1002	CDH4	NM_001794	GGGACGATTTTTCGTTGTTCGGGGTTTTCGAACGGC GGGGGCGGGAGGCGGTAATTTATTCGGAGCGCGTC GGAGAGCGAGAGTAGTAGAA
25	CDK6_9703	1021	CDK6	NM_001259	AATTTCGTTTGTAGAGTCGTCGTCGTCGTCGTCGTC GGAGGAGCGAGTCGATTTTTTTTTTTTTTTTTCGAA GCGAAGTTTTTAATATAGA
26	CDKN1B_23172	1027	CDKN1B	NM_004064	GTCGGTAAGGTTTGGAGAGCGGTTGGGTTCGCGGG ATTCGCGGGTTTGTATTCGTTTAGATTCGGACGGGT TTTGTTATTTT
27	CDKN2B_27345	1030	CDKN2B	NM_004936	TTAGAAGTAATTTAGGCGCGTTCGTTGGTTTTTGAG CGTTAGGAAAAGTTCGGAGTTAACGATCGGTCGTTC GGTTATTGTACGGGGTTT

28	CDO1_55928	1036	CDO1	NM_001801	AATTTGATTTGTGTGTGTATCGCGTTTTTAGCGATTT CGGATTTATTGCGTTGTTAGGGGTTTGGGGGTGGGT TTTTTGTTGTTTTTGCGACGATATTTTTACGTTTC
29	CDO1_55929	1036	CDO1	NM_001801	GTTTACGCGATTTTTGGGACGTCGGAGATAACGGG GTTTTTGGGAAGGCGCGGAGTTCGGGGAAGTCGGG GATGTGCGCGTGAGTCGTGTTCGTAGGGTTTTT
30	CLSTN2_19850	64084	CLSTN2	NM_022131	AGGGTTTTTCGGAGTCGTTTATTAGGGTTTTTTGGG GGTTCGGTTTCGATTGGGTAGGGGGATTTGGATAG GGTTTCGGAGCGTGGAGACGGTTGAGGAA
31	CLU_13810	1191	CLU	NM_001831, NM_203339	AGGCGTCGTATTTATAGCGTTTTGTTCGCGTATATAT TTTTTTTGGGGTTGGTTGTAAATTTGTATGATTTACG TTTAAAGAATGTCGCGGAAAGTAAGGGGA
32	CLU_19838	1191	CLU	NM_001831, NM_203339	GTGGGGGTCGGTGTAGTATCGGGTTGGGGGCGTC GGGGGGCGTATTATTATTACGAATAGTTGTGTTGGT TTTAGGAGAGATTTTGAGGTGCGGTCGTTCGGCGTT TTTAGTAGGGA
33	COL1A1_23253	1277	COL1A1	NM_000088	TATAAAAGGGGTTCGGGTTAGTCGTCGGAGTAGAC GGGAGTTTTTTTTCGGGGTCGGAGTAGGAGGTACG CGGAGTGTGAGGTTACGTATGAGCGGACGTTAATTT
34	CPT1C_23912	126129	CPT1C	NM_152359	AGGAAGTATTTATTGCGTATGTTTCGTAGTTTGGGAT GTTGAGGTTGTGAGCGGAGGCGAGCGTCGAGGATA AGTGATGG
35	CTDSPL_23795	10217	CTDSPL	NM_001008392	TAATTTTAAGGAGGACGAGGGTCGGTTGTCGGGCG CGGGCGAGAAAGGTGAGGAGGGGCGTAGGCGGTC GCGGGTTGGGGGCGAGCGTATATTTCGCGTCGTTG GAGTTTAT
36	CTDSPL_23804	10217	CTDSPL	NM_005808	GTTTTGGGAGAGGCGGTTCGGGTTCGCGTTTTAGTT TTCGTCGTCGTCGTCGTTGGGTTCGAGCGGTCGTC GTTTCGTTATTATGA
37	CYCLIND2_1 (official full gene name for CCND2)	894	CCND2	NM_001759	GGTGTAGCGTTTAGGGTCGTCGTAGGTCGGGGGTA GGGTTTTTAGCGGTTTTTTCGCGGTTAGCGGTTACG TAGGAAAAATTCG
38	DAPK1	1612	DAPK1	NM_004938	GGATAGTCGGATCGAGTTAACGTCGGGGATTTTGTT TTTTTCGCGGAGGGGATTCGGTAATTCGTAGCGGTA GGGTTTGGGGTCGGCGTTTGGGAGGG
39	DBC1_23879	1620	DBC1	NM_014618	AGGATAGGTATGAATTTCGGTTTCGGAAGGCGGTTA TTATTTTTTTTGTTTTTCGGTTTTTTCGTTTTCGTTTTC GTTGTTGTTCGTTCGTTT
40	DDX19B_22963	11269	DDX19B	NM_007242, NM_001014449, NM_001015047	CGGGTTTGAGGGTAATAGAATCGATAGTTTTAAGTG GGTAAAGGGTGGTTAAATAGGAGTGGTTTTCGACGT TATTGTGGCG
41	DKK2_23970	27123	DKK2	NM_014421	GTGCGGGGTAAGAAGGAACGGAAGCGGTGCGATTT ATAGGGTTGGGTTTTTTTGTATTTTGGGTTACGTTTT TTTGGCGAGAAAGCGTTTCGTATTTGATTGTTTTT
42	DKK2_23973	27123	DKK2	NM_014421	GAGAGAGAAAGCGGGAGTTCGCGGCGAGCGTAGC GTAAGTTCGTTTTTTAGGTATCGTTGCGTTGGTAGC GATTCGTTGTTTTTTGTGAGTTAGGGGATAACGTTTC GGGGTAATTGTGA
43	EGFR_23302	1956	EGFR	NM_201283, NM_005228	TAGGAGCGTTGTTTCGGTCGTTTCGGAGGGTCGTAT CGTTGTTTTTCGAAGAGTTCGTTTCGGTTTTTTCGAT TAATATTGGACGGAGTTAGGGGGTCGTG
44	EGR4_24277	1961	EGR4	NM_001965	TTTAGGTGGGAAGCGTATTTATCGGACGGTCGGTTC GGTGAGGCGTAGCGTTTTAGATTGGCGTATTCGCG GTTTTAGCGTTT
45	EPB41L3_19071	23136	EPB41L3	NM_012307	GGGATAGTGGGGTTGACGCGTGGTTTCGGCGTCGC GCGGTTTTTCGAATTTCGAGTTTCGCGTTCGGCGCG GTCGGGGTTTTTAATCGTTTTTTCGTTCGTCGGGATT TTTAT
46	EPB41L3_19072	23136	EPB41L3	NM_012307	GCGTGGGTTTTCGTCGTAGTTTCGCGGAGTTTCGGT GTTTTTTGTAATAGGGGGCGGGGGGAATAGCGGCG AGTAGTTTTGGG
47	FOS_22338	2353	FOS	NM_005252	CGGGTTGTAGTTAATATCGAGGGTGTAGTGCGGGG GGAGGCGGGGGTCGCGGTTGGGGGAGGGGAGGC GGGAACGGCGTAGAATGAGAGAG
48	FOXE1_13314	2304	FOXE1	NM_004473	TTTGTTCGTTTTTCGATTGTTCGTTTTTCGGGGTTCG GGCGTATTTTTTTAGGTAGGAGTAGTTGTGGCGGCG CGGTAGGAGTTTTATAGCGTTA
49	GADD45A_24463	1647	GADD45A	NM_001924	CGTTAATCGGATAAGAGTGCGCGCGGGATTCGTTTT TTTTTTTCGGTATCGTTTTCGTTTTCGTTTTTTCGGTT

					CGTTTTTCGCGTGGTTTT
50	GATA4	2626	GATA4	NM_002052	GTATAGTTTCGTAGTTTGCGTTTAGCGGAGGTGTAG TCGGGGTCGCGTATTTTCGTTTCGTTTTTGTACGTG ATTTTTATAGGTTAGTTAGCGTTTTAGGGTCGAGTTG TTGGGTCGGGGATTCGAGTCGCGAGTT
51	GATA4_13295	2626	GATA4	NM_002052	GGTATTGTTATTTTGCGTTTTCGGAGTCGTTGGTGG GCGATAAGTTTTCGTTTATTTTTTTTTATGTGCGAGTT GTCGTGTAGTTTGTTTCGGG
52	GDAP1L1_19773	78997	GDAP1L1	NM_024034	GATTTCGGGTTGTTATGGCGATTTTTAATAATTTGAT TTTTATTAATTGTAGTTGGTGGTTTATTTTCGCGTTG GAGAGCGATGCGGTTAAGTTAG
53	GDAP1L1_19775	78997	GDAP1L1	NM_024034	GAAAGAAGGAGGTTTCGGCGCGGCGGTTTTTTTTTCG TTTAGTATTATATGGTTTCGTCGAGTTTGTTTTTTTTT TTTTTTTTTTTTCGTTTCGTGAATGTTATTATTTATCG GG
54	GNB4	59345	GNB4	NM_021629	GTTGTGAGTTGCGTTTTTTACGTCGGTTTCGCGTTTT AGGGGTTGTTGAGCGTTTAGCGGATATCGGTAGCG
55	GPNMB_52607	10457	GPNMB	NM_001005340	GGTCGTAGTCGTAGTCGGGAGATTGAGGGTTAGGG CGCGGTCGCGGGGTTTTTTGGGTCGGGGCGCGGTT TACGTTTTAGGTTTTTGCGG
56	GREM1_29777	26585	GREM1	NM_013372	GAATTTGGTACGATTTTACGGAGATTTCGTTTTTTTT AGCGTAGTTTTCGTTATTGAGCGCGGGATTAACGTA GGCGATGTCGGGCGGTCGATAGGGAAAGTTTAGAT
57	Gst-Pi_New3	2950	GSTP1	NM_000852	ATTTAGTATTGGGGCGGAGCGGGGCGGGATTATTTT TATAAGGTTCGGAGGTCGCGAGGTTTTCGTTGGAGT TTCGTCGTCGTAGTTTTCGTTA
58	HHIP_23319	64399	HHIP	NM_022475	AGTAGTAGGAATAGAAACGGCGACGGCGGCGGCG GGGTAGGCGGAGGTAGGGTTAGCGTTGGGTTTTAG ATGATGTTGAGGTTTTTTTTGTCGGCGGTTGTAGTTT T
59	HIN1_3	92304	SCGB3A1	NM_052863	GAAGTTGGTTAGGGTACGGTCGTGAGCGGAGCGGG TAGGGTTTTTTTAGGAGCGCGGGCGAGGTCGGCGT TGGAGGGGCGAGGATCGGGTATAAGAAGTT
60	HOOK2_19741	29911	HOOK2	NM_013312	GGATCGTTGGATTTTGGTTCGAGTATTCGTTTTCGTT ACGTGGTAAGTTTGCGTGGAAAGGATAGGTGAGGTT TCGTTTTTTTGTGGTTGGTTTACGTGGGGCGAGGAT ATA
61	HOXA1_27316	3198	HOXA1	NM_153620	TTTTTAGAGTAAATAGCGGGAGCGTATTGGGGGTAT TTATTATTTACGTTTGTTTTTTGATTTAACGCGTAGG GGTTGTAAGGCGTAT
62	HOXA11_23844	3207	HOXA11	NM_005523	TTTTATTTATTCGGGGAGTTGCGGGTGGGAGGTGG GGACGAGAGTTGAGTTTTTATCGTTTTTTGTATATTC GAGAACGAGGATTTTGT
63	HOXA7_2	3204	HOXA7	NM_006896	TCGTAGGGTTCGTAGTCGTTTAGAATGGAAGGGTAA GAGGTTTAAATATGCGGTTAAAGAATTCGTTCGCGT TCGGCGGGTTTGGCGCGTTTCGCGGAAAAAGATTT GGA
64	HOXD1(2)	3231	HOXD1	NM_024501	GTCGGTTGACGTTTTGAGATAAGTCGGAAAAGGGTC GGGTTCGTCGAAGGTCGCGTAATTTATTTGGTCGTT GAGGAGGAAAGAGTCGTCGTTCGAGAAGACGGT
65	IGSF4_18987	23705	CADM1	NM_014333	TCGGATTTCGTTTTTAGCGTATGTTATTAGTATTTTAT TAGTTGTTCGTTCGGGTTTCGGAGGTAGTTAACGTC GTTAGTTTGAGGTAGGTGTTC
66	ISYNA1_19726	51477	ISYNA1	NM_016368	TAGGTTGGTTTGGTTTCGGTCGTTTAGAGTTTTCGTT GATTTTTTGTTTATTTCGGGTTTTTAGTTCGTCGCGA TGGAGGTCGTCGTTTA
67	JAM3	83700	JAM3	NM_032801	GGGATTATAAGTCGCGTCGCGTTGTCGTTGGTTTTT TAGTAATTTTCGATATGGCGTTGAGGCGGTTATCGC GATTTCGGTTTTGCGTTCG
68	JPH3_12611	57338	JPH3	NM_020655	TTAGATTTCGTAAACGGTGAAAACGGATTTAGGCGA TCGATATAGTAGAGTCGCGGTCGTCGGCGGTTTTG GGTCGCGAGCGTTTTTCGGAGGAGA
69	KNDC1_19691	85442	KNDC1	NM_033404 NM_152643	TGGATGGAGTTTAGGTTATATCGTCGAGTTGTTTGT GCGTGTTATTTTTGGAAGTTATTTCGTGTGTTAATTA TGTAGGGCGGTTTCGTAGTATTTT
70	KRAS_24235	3845	KRAS	NM_033360	AGGAGGGATTGTCGGATTTACGCGGCGGTTCGTTTT TTGTTTAGTCGTAAGGTTGTTTTCGTAGTCGTTAATT TTGATTCGGAGC
71	LAMA1_63431	284217	LAMA1	NM_005559	TTTTTAGATTTATCGAGTGGCGGCGGAGGCGAGATG CGCGGGGGCGTGTTTTTGGTTTTGTTGTTGTGTGTC GTCGCGTAGTGTCGGTAGAGAGGTGAGTTCG

72	LMX1A_9513	4009	LMX1A	NM_177398, NM_177399, NM_001033507	CGGTATCGTTGTTTAGGAGGCGTCGATATTTTCGTA AAGGTTTAGTCGGGGTGAGGGGTATTGGGGGGCGA TCGGGTTAGAGCGTTTCGAGGTAATAGTTATACG
73	LOC285016_22940	285016	hCG_1990170	NM_001002919	AGTTGTTTGGTATTCGCGGTTTTTAAAGGGGAAAGA AAGTTGCGTTCGCGTTAGGCGTAGCGCGTTCGGCG GACGCGGTTTTTCGGGCGAAAGTTAGGAGGGGTCG
74	LOX_23395	4015	LOX	NM_002317	GTTAGATTGATTTCGTTCGAGGAGGACGTGGTTTAT AGAAAATAAAAACGGGGTTTAAATTACGTGAGGGAA GGAGAAATTTTTAATTAAGGAGGCGAGCGGAGTACG GGTATTTTAGTT
75	LTB4R_31250	1241	LTB4R	NM_181657	TAGTAGATTTTTAGCGGTGAAGACGTAGAGTATCGG GTTGACGTTAGAATTGAAGAAGGTTAAGGTCGTAGT TTTCGTTCGCGTCGTTTGGTCGGTTTCGTTTAGTTTC GTTAAGGTTTT
76	MAL	4118	MAL	NM_002371	TTCGGGTTTTTTTGTTTTTAATTCGCGCGCGGGGGC GTTTAGGTTATTGGGTTTCGCGGAGTTAGCGAGAGG TTTGCGCGGAGTTTGAGCGGCGTTCGTTTCGTTTTA AGGTCGACGTTAGTACGTCGTTATGGTTTTC
77	MTAP_24628	4507	MTAP	NM_002451	GTAAGTGAGTTTCGAGTGTCGCGTTTTAGTTTTTTT CGCGGCGGTAAGGGACGTACGGGTCGGGCGTATG GTTTTCGGAG
78	MYO18B_24620	84700	MYO18B	NM_032608	GAAAGGTCGGATTTGTTTTTCGAGGGTCGAGTTAGT TTTTGTAGATGGTTGTAGTTTTAGTTATGAGTGTTAT TTTTTTTTTGTTTTTATAGGGCGAGGCGTGATGAGAT GGT
79	NDRG2_56603	57447	NDRG2	NM_201540,NM_201539,NM_201535,NM_201537	AGATTTTGTGGTTTCGTCGTTAATTTTTTTTAGTTCG GTTTAGAATAGGAGATTAGTTTAGGTTCGTTGAATCG TAATGTTCGGGGGAT
80	NID2_9091	22795	NID2	NM_007361	GCGGTTTTTAAGGAGTTTTATTTTCGGGATTAAATGG TTCGTAAGGTTTGGGGTAGCGGCGTTGTAGGAGAT GAGTTTAGCGTAAAGGGAATTTCGTAG
81	NOL4_19645	8715	NOL4	NM_003787	GAGAGATTCGGGATTCGTGTGTTTTTCGGGGTTTAA AGGCGTTGGGCGGGCGGTTGTTTTCGGGAGAGGC GTAGTTTTTATTTTTGGATTAC
82	NPTX1_2	4884	NPTX1	NM_002522	AGTACGTTGTTTCGGAGTTTTTCGGCGTCGTCGGCG GTTACGGACGCGGCGTATATGTCGGCGTTTACGGG TATCGAGGTGTAGATGAAG
83	NPTX2_57779	4885	NPTX2	NM_002523	GCGTCGTTTTGTATGGGTATCGCGGGTAGCGGGTA GTCGGCGTGTATCGTTTTTGGGGGTAGTGTCGTGTA TACGAAGCGGTTATCGGG
84	OGFOD2_23131	79676	OGFOD2	NM_024623	CGAGTAGTAGTTGCGTCGGGATTACGGTTCGGTGA GTGGTCGTTGTCGTTTTTACGGAGTAGTGGGTAGAG AGGGGTAGTGGAGGAGGGAAGTTCGTTTTTAGGGT CGTTTGT
85	PAK3_1	5063	PAK3	NM_002578	TGTATGATTTTAGTTCGCGGATAAGTGGGTGTGTTA GGGTCGTTTTTAGAGGGTCGGGGTTTTTTCGTTTGG TTAAATTTTAGATTCGTTTATTGGGGTTTGGGTCGTT TGAGGTAAAATTCGT
86	PAK3_3	5063	PAK3	NM_002578	GCGGGATTTATTTGTTACGGATTTAGTTATTTCGTTA AGATTTTTTTTTTATTTTCGAGCGTTTTAGTTGGCGG GGTTGGGGAGTCGTAGTTTCGCGGTCGTAGTCGTA GTTTCGGGTT
87	PAX1_27210	5075	PAX1	NM_006192	ATTGCGTCGGGTTTAGTTTCGGTTATTTCGGTTATTT CGGCGTTAGGTAGTTGGTCGGTTCGTTCGTTATGGG TAAGGGGC
88	PAX1_27211	5075	PAX1	NM_006192	GTTTAGGGAAAGCGGACGAGAGGGAAGGGAGGTA GGCGGATTCGATTTATTTTATTAGTTTTTTCGAGTTTT GTTTGTCGTTC
89	PDCD4_11827	27250	PDCD4	NM_145341, NM_014456	GTTCGTAGTTCGGGGCGTTGGGGAGGGCGCGGTTG GATTTGCGGGGTTATAAGAAGGTAGTCGGATTTTCG TATCGTAGGTTCGGATTTGATAGGATCGC
90	PHACTR3_11692	116154	PHACTR3	NM_080672 NM_183244 NM_183246	TTATTTTGCGAGCGGTTTCGCGATACGAGGTAGTCG TTTTCGTTTTTCGACGCGGTTATGGGTTCGGTCGGC GCGGGGGTAAGTTAGAGCGAGTCGCGTGGAATTAG AGTATTC
91	POMC	5443	POMC	NM_000939	GATTTGGGCGTTTTTGGTTTTTCGCGGTTTCGAGTTT TCGATAAATTTTTTGCGTCGATTGCGGTATGAGAAGT C
92	PRKCE_24134	5581	PRKCE	NM_005400	GTGGGTTTTAAGTTTACGGTTTCGTAGATTTTGATTT

					TAAGAAGGTTATTGAATATTATTATGGTCGGGGCGG GGAGTGGGGGTCGGGGTTATTTCGTTTGTTTCGAG GGTAGG
93	RAD23B_1	5887	RAD23B	NM_002874	GGCGGAGTTTGTATAGAGGCGGAGTCGCGGTAGTC GGAGAGAACGTTTTAGTAATAGTCGTTAGGAGGAAG TTTTAGGAGTTTTTGTCGTTTACGGAACGCGTTTGC GTAATTCGGGTT
94	RALY_19607	22913	RALY	NM_007367	TTTTTGGGTTTCGTTGTTTCGAGTTGGCGTCGTTCG CGCGTTTCGTCGTATTGATAGCGGCGCGAGTTTCGT AATCGCGAGTTTTGTTTTCGGTCGGTATTATTGAGG CG
95	RARA_24121	5914	RARA	NM_00102480 9	TTCGTTTCGTTTAGGTATCGTTTTTGGTTTAATTTATT TTCGGCGCGTTCGGTTGTAGCGGGAGAAACGTAGG GAATCGAGAGG
96	RARA_24129	5914	RARA	NM_000964, NM_00103360 3	TTTAGGATTATAGTGAGCGACGGGAGAGGAGGGAT GGGGAAAGTTAGAATTGGCGAGAAGGAAATGGTTA GATTAGAAGTAGAGGTCGGGGTTAAAGGCGGTTA
97	RASSF1A	11186	RASSF1	NM_007182. NM_170712 NM_170714	GCGTTGAAGTCGGGGTTCGTTTTGTGGTTTCGTTCG GTTCGCGTTTGTTAGCGTTTAAAGTTAGCGAAGTAC GGG
98	RBP4_24106	5950	RBP4	NM_006744	GGTCGTTTCGTTGTTTTATAGCGTCGGGGGGAGGG GGTCGCGTTTTCGTAATCGCGCGGGGTGAAAGATC GAAGGGGAGGCGTCGGGGGTATTTGTATAACGC
99	RECK_18940	8434	RECK	NM_021111	TTACGGTTAGTAGAAGGAGTAGCGTATTTCGTAGAG AGGTTCGGACGGTCGTTATGTTCGGGTCGGGCGGT TTTAGAGTCGCGGGATGTTCGGATTTAGTTTGGTCG TAG
100	RPRM_2	56475	RPRM	NM_019845	TCGAGGAAGAAGATGTCGAAGATTACGGTGAGTGA GAGTACGTATATGATCGCGATTTGTATTACGCGTATT ATGTATAGGTTACGTTCGTTCGGGTTTTT
101	SALL4_12833	57167	SALL4	NM_020436	GAGGCGTAAGTAGGCGAAATTTTAGTATATTAATTC GGAGGAGGATTAGGGCGAGTAGTAGTCGTAGTAGT AGATTTCGGAGTTTGTAGATGCG
102	SEMA3F_23485	6405	SEMA3F	NM_004186	GATTAGAGCGAGCGAACGAATCGCGGCGGTTCGGA GAGTTTCGAGCGTAGCGTAGGATTTGGGTACGTCG CGAGGAATCGTGTAGTTTAGCGCGGTCGTTCGGTTC GGGTTTAGTAGTTA
103	SLC5A8_24598	160728	SLC5A8	NM_145913	GGTTTGTTGGTCGTTTTTAGCGAAGGCGTAGTAGAT GTCGATGGCGGTCGAGATGATTAGTATGTTCGCGAA TATTACGTAGTTTTATATTACGAAGGTGTCGATGTTT CG
104	SLC5A8_24601	160728	SLC5A8	NM_145913	GTATTTAGGGTAGCGGGTCGATTTTTCGAGGTTTTA TATTTGGGTTTGAGGGGCGCGGTTCGTAGCGGCGG GTGTAGGGGCGATTGTTAGTTTTTATTTCG
105	SLIT1_23651	6585	SLIT1	NM_003061	GCGTTATGGTGTTTTTATAGCGTTTCGTTCGCGAGTT AGACGGTAGTAGTCGTTGATTATTTTCGTTCGGGGT CGTTTTTAGGTGTAGTTTCGGGGTAGAGTTATCGAA GA
106	SLIT1_23653	6585	SLIT1	NM_003061	TTGTAGGCGGTTTGTAGTCGTTGAGTGGTCGTCGG GAGAGGGGGGTTGCGGCGGGGGAGGGCGGGGAG GAGTTTGGTTTTGGATGTGTGTTTTTCGATTGATGGA TGATTGTC
107	SLIT2_23672	9353	SLIT2	NM_004787	GAGGATCGGTTTAGGTTGCGGCGGAGTCGAGGGCG AGGGAGAGGTCGCGTGAGTGAGTAGAGTTTAGAGT CGTGCGTTTTTAGAATTG
108	SLIT2_23676	9353	SLIT2	NM_004787	AGGGGAAGACGAAGAGCGTATATTTATAGTTTTTCG GTGTTGCGGGGGATATTTTTGGGTACGTTGCGTAGC GTTAGTTCGTG
109	SLIT2_23681	9353	SLIT2	NM_004787	TAGCGGAGAGGAGATTACGCGTTTTTTGTTTTTTAAG GATGAATTTGGCGGTAAAAGAGTTGGGGTTTTTAAC GGTTGTTAAGATGTAGGGGTC
110	SLIT3_23619	6586	SLIT3	NM_003062	AGGGGTATTTATAGGCGTTTAGCGTTGCGGGGGAT GTTTCGAGGAATCGCGCGGAGGTTTAGTTCGTGGTA GTTTACGTTGGTAGCGGAGTAGGTA
111	SMPD1_24061	6609	SMPD1	NM_000543	GAAGGGTAATCGGGTGTTTTCGGCGTCGTTCGGGG TTTTGAGGGTTGGTTAGGGTTTAGGTCGGGGGGGA CGGGATAGACGAATTAG
112	SOCS1_23595	8651	SOCS1	NM_003745	GATAGGGTTTTGTTTTCGGCGGGTGTGGAGATAGTT GGGGCGGAGGAGGGTGTGTTAGGGCGCGTTTTAAG AGGGTTTGGCGGTAGAAAGTGGAATTCGAGGTAGC GGGGTAAAAT

113	SOX1_27153	6656	SOX1	NM_005986	TTGTAGTTTTCGAGTTGGAGGTCGTTGAGGATCGAG CGTAGGAGGAAGGAGATAGCGCGTAGCGGCGGTC GGCGAGGAGATAGTATATTTCGGGTCGGGTTTAGC GTATCGTTTT
114	SOX1_27159	6656	SOX1	NM_005986	GTTAGGAGTTCGTCGGTTAGCGAGTATTTGTTTTTTT TGAGTAGCGTTTTGGTTTTGCGGCGCGGTCGGTATT TGTAATTCGGGTG
115	SOX17_66072	64321	SOX17	NM_22454	GAGATGTTTCGAGGGTTGCGCGGGTTTTTCGGTTCG AAGTCGTCGTTCGTGTTTTGGTTTGTCGCGGTTTGG TTTATAGCGTATTTAGGGTTTTTAGTCGGTTTAGTGA TATTGCGG
116	SPARC_Wis	6678	SPARC	NM_003118	TTTCGCGGTTTTTTAGATTGTTCGGAGAGCGCGTTTT GTTTGTCGTTTGTTTGTTTGTTATTGAGGTATG
117	SPN_24052	6693	SPN	NM_003123, NM_001030288	ATCGTAGGTTGGGTTTGGTCGTTGGTAGGGAAGTG GGTAGAGGGGAGGTTCGGTTAGGTTTTTCGGTAATT TTCGCGTGTTTTGTTTTT
118	SST_23808	6750	SST	NM_001048	TGGTTGCGTTGTTTATCGTTTTGGTTTTGGGTTGTGT TATCGGCGTTTTTTCGGATTTTAGATTTCGTTAGTTT TTGTAGAAGTTTTTGGTTGTTGTCGCGGGGAAGTAG GTAA
119	TAC1_56187	6863	TAC1	NM_003182	GGGTATTTATTGCGACGGATAGTTTCGCGGGGTGTT GAGTTTTTTTGGTTTTTTCGAGCGTACGTTGGTCGTT TCGTATTTTCGGTAGTTGTCGTCGG
120	TERT_23702	7015	TERT	NM_003219, NM_198255	GGTTTCGATAGCGTAGTTGTTTCGGGCGGATTCGG GGGTTTGGGTCGCGTTTTTTCGTTCGCGCGTCGTTC GCGTTTTTAGGGTGTAG
121	TFPI-2	7980	TFPI2	NM_006528	GTTCGTTGGGTAAGGCGTTCGAGAAAGCGTTTGGC GGGAGGAGGTGCGCGGTTTTTTGTTTTAGGCGGTT CGGGTGTTCGTTTTATG
122	TLL1_24051	7092	TLL1	NM_012464	TAAGGAATTTTGTATTCGGAGGCGGGGAGGGCGTA GGTAAATTCGGTTTTGGCGGCGTTGGCGTTCGTAGT TTGTTAGGT
123	TNFAIP1_23212	7126	TNFAIP1	NM_021137	GTGGTTAGCGGATTTCGAGTCGTTTTTAGTTTGTAGT CGTTTGTTTTTTAGTAGTTTTAAGTTGTGAGTTTATAT TTTGCGTTCGTCGATTTCGTTCGGAGTGTTGTTTAGT T
124	TRMT1_19794	55621	TRMT1	NM_017722	TTTCGTAGGGTTCGGTGTCGTTTTTTATCGTTGTTGT ATTCGGTAGTTTTGGAGATTGTTATTCGAAAAATCGG GTTTTAGAGAGTATTCGG
125	TWIST1_3	7291	TWIST1	NM_000474	GTTAGGGTTCGGGGGCGTTGTTCGTACGTTTCGGC GGGGAAGGAAATCGTTTCGCGTTCGTCGGAGGAAG GCGACGG
126	TWIST1_9329	7291	TWIST1	NM_000474	TTTAGTTCGTTAGTTTCGTCGGTCGACGATAGTTTGA GTAATAGCGAGGAAGAGTTAGATCGGTAGTAGTCGT CGAGCGGTAAGCGCGGGGGACGTAAGCGGCGTAG TAGTA
127	UGT1A1_22912	54658	UGT1A1	NM_007120	TTTTGTGGTTAGTCGCGGTAGGGGAATTTGGAGTTT TTTGGTTATTTTAGTAGAAGTTATCGATAGTTGATTG TTTATTTTACGT
128	WIF1_9096	11197	WIF1	NM_007191	GCGTCGTTAGATATTTTGTTGCGTTGTAGTTTTTTTA GTTAGGGTTGTTTTCGTTTAGACGGTTGGGCGCGTC GTTTTTCGGTTTGGGTGTTA
129	WIT1_24567	51352	WIT1	NM_015855	GTATGGAGCGTTTTGCGATTGTAGGAGTACGTTAGT TTTTTAGCGTTGGTTTAGTGTCGTTTGGGTTTTCGGG TATGTGGATTCGTT
130	WT1_1	7490	WT1	NM_024426, NM_024424	TGTGTTATATCGGTTAGTTGAGAGCGCGTGTTGGGT TGAAGAGGAGGGTGTTTTCGAGAGGGACGTTTTTTC GGATTCGTTTTTATTTTAGTTGCGAGGGCGTTTTTAA GGAGTAGCG
131	XRCC3_9322	7517	XRCC3	NM_005432	CGTTTGTTTTTTATAGGTTCGGGTAATGGAGATTCGC GGTCGTTTTCGTTTTTTGATTTTGTTTTTATTTTTTACG TTCGTTGTCGTTTACGATTTTCGATTTCGTTGT
132	ZGPAT_23961	84619	ZGPAT	NM_032527 NM_181484 NM_181485	TGTATGCGGAGAGGTCGTAGTTATTGTTGTGAGTAG GATATAGTGGCGGTTGATTTGGGAGAAGTTATAGAG GGACGGGGTGGGAGAGGGACGAGGAGTCGGGAAT GGT

Table 3: qMSP Molecular Beacon sequences

Row Nr	Assay Name	Gene ID	Official Gene Symbol	Refseq	Molecular beacon sequence (5' – 3') (modification beacons: 5' FAM, 3' DABCYL) SEQ ID NO's 397-425
1	ALX3_25180	257	ALX3	NM_006492	CGACATGCGCGGTTGATTCGTTTTTCGGTTTGC GGGCATGTCG
2	C13orf18_Gron	80183	C13orf18	NM_025113	CGACATGCCGTCGTAGGTATCGAGACGTCGTTT AGATGGGCATGTCG
3	GATA4	2626	GATA4	NM_002052	CGACATGCGTAGTCGGGGTCGCGTATTTTCGTT TCGGCATGTCG
4	HOXA11_23844	3207	HOXA11	NM_005523	CGACATGCGATAAAAACTCAACTCTCGTCCCCA CCGCATGTCG
5	JAM3	83700	JAM3	NM_032801	CGACACGATATGGCGTTGAGGCGGTTATCGTGT CG
6	JPH3_12611	57338	JPH3	NM_020655	CGTCTGCAACCGCCGACGACCGCGACGCAGAC G
7	LMX1A_9513	4009	LMX1A	NM_177398, NM_177399, NM_001033507	CGACATGCCCGATCGCCCCCCAATACCGCATGT CG
8	NOL4_19645	8715	NOL4	NM_003787	CGACATGCGGCGTTGGGCGGGCGGTTGCATGT CG
9	PAK3_1	5063	PAK3	NM_002578	ACATGCCGTTTTTAGAGGGTCGGGGTTTTTTCG GCATGT
10	TERT_23702	7015	TERT	NM_003219, NM_198255	CGACATGCGACCCAAACCCCCGAATCCGCGCAT GTCG
11	TFPI2	7980	TFPI2	NM_006528	CGACATGCACCGCGCACCTCCTCCCGCCAAGC ATGTCG
12	TWIST1_3	7291	TWIST1	NM_000474	CGACATGCCGGCGGGGAAGGAAATCGTTTCGC ATGTCG
13	CCNA1_Gron	8900	CCNA1	NM_003914	CGACATGCACGACGCCCCCGAACCTAACGCATG TCG
14	CDO1_55929	1036	CDO1	NM_001801	CGACATGCCCGACTTCCCCGAACTCCGCATGTC G
15	CDO1_55928	1036	CDO1	NM_001801	CGACATGCGCGATTTCGGATTTATTGCGTTGTTA GGGCATGTCG
16	GREM1_29777	26585	GREM1	NM_013372	CGACATGCGGGATTAACGTAGGCGATGTCGGG CATGTCG
17	GPNMB_52607	10457	GPNMB	NM_001005340	CGACATGCGGTTTTTTGGGTCGGGGCGCGGCA TGTCG
18	HIN1_3	92304	SCGB3A1	NM_052863	CGACATGCAGGGTTTTTTTAGGAGCGCGGGCGA GG-GCATGTCG
19	HOXD1(2)	3231	HOXD1	NM_024501	CGACATGCGGGTCGGGTTCGTCGAAGGTCGGC ATGTCG
20	LAMA1_63431	284217	LAMA1	NM_005559	CGACATGCCAAAAACACGCCCCCGCGCATGTCG
21	LTB4R_31250	1241	LTB4R	NM_181657	CGACATGCGTAGTTTTCGTTCGCGTCGTTTGGT CGGCATGTCG
22	MAL	4118	MAL	NM_002371	CGACATGCAAACGAACGCCGCTCAAACTCCGCG CGCATGTCG
23	NDRG2_56603	57447	NDRG2	NM_201540 NM_201539 NM_201535 NM_201537	CGACATGCGTTCGGTTTAGAATAGGAGATTAGTT TAGGTTCGTTGCATGTCG
24	NID2_9091	22795	NID2	NM_007361	CGACATGGGTTCGTAAGGTTTGGGGTAGCGGC CATGTCG
25	NPTX2_57779	4885	NPTX2	NM_002523	CGACATGCGCGGGTAGTCGGCGTGTATCGCAT GTCG
26	RASSF1A	11186	RASSF1	NM_007182. NM_170712 NM_170714	CGTCTGCGTGGTTTCGTTCGGTTCGCGTTTGTT AGGCAGACG
27	SALL4_12833	57167	SALL4	NM_020436	CGACATGCGGAGGATTAGGGCGAGTAGTAGTC GTAGCATGTCG
28	SOX17_66072	64321	SOX17	NM_22454	CGACATGCGTTCGTGTTTTGGTTTGTCGCGGTTT GGCATGTCG
29	TAC1_56187	6863	TAC1	NM_003182	CGACATGCGGTTTTTTCGAGCGTACGTTGGTCG CATGTCG

EXAMPLES

Example 1: Discovery of methylation markers in cervical cancer, using relaxation ranking

To identify genes that are downregulated in cervical cancer due to promoter hypermethylation and to enrich for those genes that are most frequently involved in cervical cancer, a multistep approach was used combining:

- Affymetrix expression microarray analysis on a panel of frozen tissue samples from 39 human primary cervical cancers to identify cancer-specific down-regulated genes.
- Affymetrix expression microarray analysis on a panel of 4 different cervical cancer cell lines in which the expression of (hyper)methylated genes was re-activated upon treatment with 5-aza-2'deoxycytidine (DAC) (blocking DNA methylation), and/or trichostatin A (TSA) (inhibiting histone deacetylase - HDAC).

Data from both approaches were combined, and a novel non-parametrical ranking and selection method was applied to identify and rank candidate genes. Using in silico promoter analysis we restricted the analysis to those candidate genes that carry CpG-islands. The new approach resulted in a significant enrichment of hypermethylated genes: we compared the first 3000 high-ranking candidate probes with lists of imprinted genes, X-chromosome located genes and known methylation markers. In addition, we determined the hypermethylation status of the 10 highest ranking candidate genes in both cervical cancers and normal cervices using COBRA (COmbined Bisulfite Restriction Analysis).

MATERIAL AND METHODS

Primary cervical tissue samples:

For the expression microarray analysis, tissues from 39 early stage frozen cervical cancer samples were used from a collection of primary tumors surgically removed between 1993 and 2003 (University Medical Center Groningen, Groningen, The Netherlands). All cervical cancer patients underwent gynecological examination for staging in accordance with the International Federation of Gynecology and Obstetrics (FIGO) criteria (Finan *et al.*, 1996). Tumor samples were collected after surgery and stored at -80 °C. The stage of cervical cancer patients included 33 FIGO stage IB (85%) and 6 FIGO stage IIA (15%). The median age of the cervical cancer patients was 46 years (IQ range 35 – 52 yr.).

For COBRA and BSP (Bisulfite Sequencing PCR), 10 (of the 39) primary cervical cancers and 5 controls (normal cervix) were used. The age-matched normal cervical controls were women without a history of abnormal PAP smears or any form of cancer and planned to undergo a hysterectomy for benign reasons during the same period. Normal cervices were collected after surgery and histologically confirmed.

Informed consent was obtained from all patients participating in this study.

Cervical cancer cell lines:

Four cervical carcinoma cell lines were used: HeLa (cervical adenocarcinoma, HPV18), SiHa (cervical squamous cell carcinoma, HPV16), CSCC-7 (non-keratinizing large cell cervical squamous cell carcinoma, HPV16) and CC-8 (cervical adenosquamous carcinoma, HPV45). HeLa and SiHa were obtained from the American Tissue Type Collection. CSCC-7 and CC-8 (Koopman et al., 1999) were a kind gift of Prof. GJ Fleuren (Leiden University Medical Center, Leiden, The Netherlands). All cell lines were cultured in DMEM/Ham's F12 supplemented with 10% fetal calf serum.

Cell lines were treated for 3 days with low to high dose (200 nM, 1 μM or 5 μM) 5-aza-2'deoxycytidine (DAC), 200 nM DAC with 300 nM trichostatin A (TSA) after 48 hours, or left untreated. Cells were split to low density 24 hours before treatment. Every 24 hours DAC was refreshed. After 72 hours cells were collected for RNA isolation.

RNA and DNA isolation:

From the frozen biopsies, four 10-μm-thick sections were cut and used for standard RNA and DNA isolation. After cutting, a 3-μm-thick section was stained with haematoxylin/eosin for histological examination and only tissues with >80% tumor cells were included. Macrodissection was performed to enrich for epithelial cells in all normal cervices.

For DNA isolation, cells and tissue sections were dissolved in lysis buffer and incubated overnight at 55°C. DNA was extracted using standard salt-chloroform extraction and ethanol precipitation for high molecular DNA and dissolved in 250 μl TE-4 buffer (10 mM Tris; 1 mM EDTA (pH 8.0)). For quality control, genomic DNA was amplified in a multiplex PCR containing a control gene primer set resulting in

products of 100, 200, 300, 400 and 600 bp according to the BIOMED-2 protocol (van Dongen et al., 2003).

RNA was isolated with TRIzol reagent (Invitrogen, Breda, The Netherlands) according to manufacturer's protocol. RNA was treated with DNAse and purified using the RNeasy mini-kit (Qiagen, Westburg, Leusden, The Netherlands). The quality and quantity of the RNA was determined by Agilent Lab-on-Chip analysis.

Expression data:

Gene expression for 39 primary cancers and 20 cell line samples was performed using the Affymetrix HGU 133 Plus 2.0 array with 54,675 probes for analysis of over 47,000 human transcripts. The labeling of the RNA, the quality control, the microarray hybridization and scanning were performed by ServiceXS according to Affymetrix standards. For labeling, ten microgram of total RNA was amplified by *in vitro* transcription using T7 RNA polymerase.

Quality of the microarray data was checked using histograms, boxplots and a RNA degradation plot. One cell line sample was omitted because of poor quality. Using BioConductor (Gentleman *et al.*, 2004), present (P), absent (A) or marginal (M) calls were determined with the MAS5 algorithm. MAS5 uses a non-parametric statistical test (Wilcoxon signed rank test) that assesses whether significantly more perfect matches show more hybridization signal than their corresponding mismatches to produce the detection call for each probe set (Liu *et al.*, 2002). The relaxation ranking approach only relied on P-calls. Some samples were analyzed in duplicate, and the profile of P-calls is highly similar (93-95 % of the probesets have an identical P/M/A call).

Relaxation ranking algorithm:

In order to identify the most promising markers that are methylated in cervical cancer, we assumed that such markers should be silenced in cancer cells and upregulated upon re-activation after DAC/TSA treatment; therefore, the best methylation markers will be genes represented by probes with:

- no expression in primary cervical cancers: P-calls=0 out of 39 cancers
- no expression in (untreated) cervical cancer cell lines: P-calls=0 out of 4 cell lines

- expression in cervical cancer cell lines treated with DAC (or DAC in combination with TSA): P-calls=15 out of 15 treated cell lines

To select for those gene probes that would be the best candidate hypermethylated genes in cervical cancer, we present the relaxation ranking algorithm. Probe sets were ranked, not primarily based on the number of P-calls and thus explicitly setting thresholds, but primarily driven by the number of probe sets that would be picked up, based on selection criteria (the number of P-calls in primary cancers, untreated and treated cell lines). The stricter (e.g. P-calls: 0 - 0 - 15) these selection criteria, the lower the number of probes that meet with these criteria; while if the conditions become more and more relaxed (higher number of P-calls in primary cancers and untreated cell lines, and lower number of P-calls in treated cell lines), the more probes will comply. In the end, using P-calls: 39 - 4 - 0 as criteria, all probe sets were returned. This way, there was no need to define a 'prior' threshold for the number of P-calls.

The following sorting method was applied:

(1) All possible conditions were generated and the number of probes that were picked up under these conditions was calculated:
   a. the number of samples with expression (P) of a certain probe in
      i. primary cervical cancer samples is called x_{sample}
      ii. cervical cancer cell lines is called y_{sample}
      iii. treated cervical cancer cell lines is called z_{sample}
   b. all combinations of x, y and z are made
      i. x (the number of P-calls in primary cancers) varies from 0 to 39
      ii. y (the number of P-calls in untreated cell lines) from 0 to 4
      iii. z (the number of P-calls in treated cell lines) from 0 to 15
      iv. In total, 3200 combinations of x, y and z can be made
   c. a probeset was found under each of these generated conditions x, y and z if:
      i. $x_{sample} \leq x$ (number of P-calls for probe in primary cancers smaller or equal compared to condition) AND
      ii. $y_{sample} \leq y$ (number of P-calls for probe in untreated cell lines smaller or equal compared to condition) AND
      iii. $z_{sample} \geq z$ (number of P-calls for probe in treated cell lines larger or equal compared to condition)

d. under very strict conditions (x=0, y=0, z=15) no probes were found, while under the most relaxed conditions (x=39, y=4, z=0) all probes were returned. For all combinations of x, y and z, the number of probes that complied (w), was stored

(2) The data was sorted with w as primary criterion (ascending), followed by x (ascending), y (ascending) and z (descending)

(3) This sorted dataset was analyzed row per row. In row i, the w_i probes retrieved with criteria x_i y_i z_i were compared with the list of probes, already picked up in rows 1 to i-1. If a probe did not occur in this list, it was added to the list

(4) This process continued until there were m (user-defined) probes in the list

DNA methylation analysis using COBRA and bisulphate sequencing:

To validate the (hyper)methylated status of candidate gene probes, DNA extracted from 10 cervical cancers and 5 normal cervices were analyzed using BSP and COBRA. Bisulfite modification of genomic DNA was performed using the EZ DNA methylation kit (Zymogen, BaseClear, Leiden, The Netherlands). The 5' promoter region of the tested gene was amplified using bisulfite treated DNA. PCR primers for amplification of specific targets sequences are listed in Table 4. COBRA was performed directly on the BSP products as described by Xiong et al. (Xiong and Laird, 1997) using digestions with BstUI, Taq1 and/or HinfI according the manufacture's protocol (New England Biolabs Inc., Beverly, MA). For sequence analysis, the BSP products were purified (Qiagen) and subjected to direct sequencing (BaseClear, Leiden, The Netherlands). Leukocyte DNA collected from anonymous healthy volunteers and in vitro CpG methylated DNA with SssI (CpG) methyltransferase (New England Biolabs Inc.) were used as negative and positive control, respectively.

Table 4: list of primers used for BSP ([1]: +1 is transcription start site (TSS); [2:] Several primer pairs were tested, however, none worked)

Name	Forward primer (5'-3')	Reverse primer (5'-3')	Ta	Start position[1]	End position	RefSeq
DAZL *ADARB1*[2]	TTTGGGGGTGATGTGTGTGTTT	TCTCCCTCAACTCACCATAATA	54	-161	312	NM_001351 NM_015834
SYCP3	AAAATTTAAAAATTGGAAGGTATT AGG	ACCTCACTAATCAAAAACAACCTCT	54	-208	+186	NM_153694
AUTS2	TTTTAAAAGTGATAAAGTTGGTTA TGG T	CCCTTTTCTTTCTCCTCTCTTTCT	56	+300	-184	NM_015570
NNAT	GGTTAGGGATTGGGGAGAA	GCTAAACTTACCTACAACAACAC	54	-271	210	NM_005386
SST	GGGGTATGTGGAATTGTGTG	AAA TCT CCT TAC CTA CTT CCC C	54	-185	+276	NM_001048
HTRA3	GTYGGTTTTGTYGTTATGTAGGY	AAC TTC ACT TCC TCC CTA ACC	57	+190	+622	NM_053044
ZFP42	AGTAGGTGTTTGTTGAAGATAG	ACT CAT AAC ACA CAT AAC CAT C	60	+308	+580	NM_174900
NPTX1	GGTAGTGGGGGTTTGATAG	AAATAATCTCCTTCTACTACAACAC	54	-2	+372	NM_002522
GDA	TATAGAAGGTGGAGGAAGTTGA	CACCTCCATAAAACAAATCCAAA	54	-239	+194	NM_004293
CCNA1	TATAGTTGGAGTTGGAGGGT	AAACAACTAACAAATACACTAAAA	54	-279	+146	NM_153694

RESULTS

To identify novel markers that are methylated in cervical cancer, we applied a multistep approach that combines re-expression of silenced hypermethylated genes in cervical cancer cell lines (using DAC and DAC/TSA), downregulated expression in 39 cervical cancers expression, and selection of candidate markers using a relaxing ranking algorithm. The best profile of a candidate marker would be: no expression in any of the 39 cervical primary cancers and 4 untreated cancer cell lines, but re-activation of expression after demethylation and/or blocking of histone deacetylation in all 15 cell lines treated with various combinations of DAC/TSA (P-calls: 0 – 0 – 15). However, none of the probe sets showed this ideal profile. To generate a list of candidate genes, a relaxation ranking algorithm was applied.

The only variable used in the relaxation ranking is the number of probes we would like to retrieve. As shown in Figure 1, the number of probes retrieved (w) with parameters x, y and z (the number of P-calls in respectively primary tumor samples, untreated and treated cell lines) follows a complex profile which consists not only of additive elements, but also interactions between the parameters. In general, the number of P-calls in primary cancer samples (x) has the largest influence on w. The sorting methodology has the advantage that no cut-off values have to be chosen for x, y and z, and therefore there is no need to implicitly link a relative weight factor to the parameters.

To calculate the most optimal number of potentially hypermethylated candidate markers for further analysis, we estimated this number based on known (i.e. described in literature) methylation markers in cervical cancer. Forty-five known methylation markers were found using text-mining using GeneCards (Rebhan *et al.*, 1997) for aliases/symbols to query PubMed through NCBI E-Utils. The position of the markers after ranking ("observed") was determined as shown in the step plot in Figure 2. If the markers would be randomly distributed in the ranking, the profile would be similar to the curve, marked 'expected'. This 'expected' curve is not a straight line, but is calculated based on whether a probe could be assigned with a gene symbol and taking probes into account that are associated with a gene that is already associated with an earlier selected probe. The number of observed methylation markers has in general the same slope as expected. However, until about 3000 probes, the slope of the number observed markers versus the number of selected probes (in dashed lines) cannot be explained if the markers would be randomly distributed as its steepness is much higher. When selecting more than 3000 probes, the slope suddenly decreases to a level that is close to random distribution. This enrichment can also statistically be proven. Therefore, we selected the first 3000 probes, referred to as TOP3000, in the ranking for further analysis. In this TOP3000 list, 2135 probes are associated with a gene symbol, of which 1904 are unique.

Validation of the 10 highest-ranking candidate genes using COBRA:

In order to validate whether the highest ranking genes represent markers that are functionally hypermethylated in cervical cancer, we performed COBRA on bisulfite-treated DNA of 10 cervical cancers and 5 normal cervices. For this analysis we focused on those first 10 genes from the highest ranking probe-list (Table 5) that:

- represent a known gene (i.e. gene symbol)
- contain a CpG-island surrounding the TSS
- are located on any chromosome except chromosome X
- are expressed in less than 15 carcinomas

BSP was used to amplify the CpG-islands of these candidate genes using bisulfite-treated DNA and COBRA to determine the methylation status. *CCNA1* (at position 49) was included as a positive control for the highest listed, reported cervical cancer specific methylation gene promoter. BSP/COBRA of *CCNA1* revealed that 6 of 10 carcinomas are methylated at the restriction enzyme sites (T1, T3, T5, T7, T9 and T10 in Figure 3). Sequence analysis of the BSP-products (on average 7-9 independent

clones for each carcinoma) of these 10 carcinomas revealed that in 6 carcinomas the promoter is hypermethylated in good agreement with the COBRA results (Figure 3C).

Table 5: Methylation status using COBRA of the 10 highest ranking gene promoters. Gene selected for further validation after applying additional criteria. Included is CCNA1 on position 47 (original position 241) as the highest ranking cervical-cancer-associated hypermethylated gene. Methylation status was determined by BSP/COBRA (see Figure 3 and Figure 4).

Rank	Gene symbol	Chromosomal location	Methylation in cancer	Methylation in normal
1	DAZL	3p24.3	9/9	5/5
2	ADARB1	21q22.3	Nd	Nd
3	SYCP3	12q	9/9	5/5
4	AUTS2	7q11.22	0/9	0/5
5	NNAT	20q11.2	9/9	5/5
6	SST	3q28	7/9	0/5
7	HTRA3	4p16.1	1/9	0/5
8	ZFP42	4q35.2	9/9	5/5
9	NPTX1	17q25.1	5/10	0/5
10	GDA	9q21.13	0/9	0/5
47	CCNA1		6/10	0/5

Table 5 summarizes the methylation status of the 10 highest ranking genes in 10 cervical cancer and 5 normal cervices using COBRA. One gene (ADARB1 at rank 2) could not be analyzed for methylation as no specific BSP products could be amplified using several combinations of primer pairs. Interestingly, using the BSP products of the other 9 listed genes, 7 (78%) showed methylation in carcinomas (Table 5). Four genes are hypermethylated in all 9 tested cancers, while for SST (7 of 9 carcinomas), HTRA3 (1 of 9 carcinomas) and NPTX1 (5 of 10 carcinomas) not all tested carcinomas are hypermethylated. Figure 4 shows representative methylation analysis of 3 genes using COBRA. Three (NNAT, SST and NPTX1) of the 7 hypermethylated gene promoters have been reported to be methylated in tumors previously. Taken

these data together, these findings showed that the relaxation ranking algorithm resulted in a very significant enrichment for genes with a positive methylation status.

A cervical-cancer-specific hypermethylated marker is only of relevance for the diagnosis of (pre-) malignant disease in case normal cervical epithelium is not methylated. COBRA analysis of 5 normal cervices for all 9 genes revealed that 4 genes (DAZL, SYCP3, ZFP42 and NNAT) are hypermethylated in all 5 samples (Table 5). On the other hand, of the 7 genes hypermethylated in cervical cancer specimens, 3 genes (*SST, HTRA3 and NPTX1*) did not show DNA methylation in any of the normal cervices of 5 independent individuals. We observed the same methylation profile for CCNA1 that was reported previously as a cervical cancer specific gene (Kitkumthorn *et al.*, 2006) with hypermethylation in only 6 of 10 tumors but none of the 5 normals (Table 5; Figure 3).

Example 2: BROAD analysis: Genome-wide Promoter Alignment.

The "Database of Transcription Start Sites" (DBTSS) (Suzuki *et al.*, 2004) mapped each transcript sequence on the human draft genome sequence to identify its transcriptional start site, providing more detailed information on distribution patterns of transcriptional start sites and adjacent regulatory regions. The promoters of the above identified TOP3000 genes were separately mapped on the genome-wide alignment of all promoter associated CpG islands. All the promoter sequences were subsequently aligned by clustalW algorithm (Li 2003; Thompson *et al.*, 1994). Treeillustrator (Trooskens *et al.*, 2005) was used to visualize the large guide tree in addition to indicating the location of the known markers. Some regions on the "circle" are denser in known markers than others, indicating that there might be a sequence mechanism located in the small region around the TSS which makes certain genes more methylation-prone. The genes were selected as candidates to be methylated if they were located in a cluster, i.e. less than 9 nodes (distance to the closest neighboring marker) away from a marker already described in the literature. These genes were assigned a score, calculated as follows: if the gene is a known literature marker, score +10, if a known marker is one node away, score +9, if there are markers two nodes away: addition to score = number of markers * 8, etc. The genes were ranked according to this score.

A final gene selection was made based on the ranking, the opportunity to design primers, genes to be known as tumor suppressor genes and expert knowledge on their

function, history and mutation status in other cancer types. Also known genes from literature and previous research were included for confirmation.

A final selection of markers resulting from the above set out approaches, were tested on tissue using the Base5 methylation profiling platform (Straub *et al.* 2007). Differential methylation of the particular genes was assessed using Base5 methylation profiling platform as follows: DNA was extracted from cervical samples, bisulfite converted, and selected regions of the particular genes were amplified using primers whose sequence represented converted or non-converted DNA sequences. Amplification was monitored in real-time set up using SYBR® green. Data analyses designed to cope with inherent variance (*i.e.,* noise) in measured Ct and Tm values were applied to withhold 112 different assays for detecting differential methylation of ALX3, ALX4, AR, ARID4A, ATM, AURKA, B4GALT1, BMP2, BMP6, BNIP3, C13orf18, C16orf48, C9orf19, CALCA, CAMK4, CCNA1, CCND2, CDH1, CDH4, CDK6, CDKN1B, CDKN2B, CLSTN2, CLU, COL1A1, CPT1C, CTDSPL, CYCLIND2, DAPK1, DBC1, DDX19B, DKK2, EGFR, EGR4, EPB41L3, FOS, FOXE1, GADD45A, GATA4, GDAP1L1, GNB4, Gst-Pi, HHIP, HOOK2, HOXA1, HOXA11, HOXA7, IGSF4, ISYNA1, JAM3, JPH3, KNDC1, KRAS, LMX1A, LOC285016, LOX, MTAP, MYO18B, NOL4, NPTX1, OGFOD2, PAK3, PAX1, PDCD4, PHACTR3, POMC, PRKCE, RAD23B, RALY, RARA, RBP4, RECK, RPRM, SEMA3F, SLC5A8, SLIT1, SLIT2, SLIT3, SMPD1, SOCS1, SOX1, SPARC, SPN, SST, TERT, TFPI-2, TLL1, TNFAIP1, TRMT1, TWIST1, UGT1A1, WIF1, WIT1, WT1, XRCC3, and ZGPAT in cervical cancer tissue samples.

MATERIAL AND METHODS

Samples:

A total of 201 frozen tissue samples (87 cervical cancer samples, the majority derived from squamous cell carcinomas; and 114 normal tissues) were collected by UMC Groningen. If the tissue contained more than 20% stromal cells, the samples were macro-dissected to enrich for tumor cells.

DNA isolation and bisulphite modification:

DNA was isolated using proteinase K digestion and phenol/chloroform extraction. DNA concentration was measured using NanoDrop Spectrophotometer. From each sample, up to 2 µg of genomic DNA was converted using a bisulphite based protocol (EZ DNA Methylation Kit™, ZYMO Research, Orange, CA.).

Detection of hypermethylation:

Methylation specific PCR (MSP) primers were designed for each of the genes assessed for (hyper)methylation. An example on primer design spanning a large region of the promoter is provided in Figure 5A and 5B for ALX4.

For some genes more primer pairs were designed giving a total of 424 different assays. These assays were applied on 8 sub-arrays of 2 OpenArray™ plates by BioTrove Inc. The beta-actin assay was applied on each sub-array as an internal control. Quality control was performed using an *in vitro* methylated DNA sample and a negative control sample. The selectivity and the reproducibility were checked. After DNA conversion and purification, beta-actin copy number was determined by qMSP. The equivalent of 1500 beta-actin copies per sample was applied per sub-array of an OpenArray™ plate on a real-time qPCR system (BioTrove Inc.) using the DNA double strand-specific dye SYBRgreen for signal detection.

The cycling conditions were: 90°C-10 seconds, (43°C 18 seconds, 49°C 60 seconds, 77°C 22 seconds, 72°C 70seconds, 95°C 28 seconds) for 40 cycles, 70°C for 200 seconds, 45°C for 5 seconds. A melting curve was generated in a temperature range between 45°C and 94°C.

Analysis of methylation:

For each combination of assays and samples two parameters were collected using an algorithm which is part of the standard data analysis package offered by the supplier. The parameters were the Ct value (threshold cycle number) of the assessed amplicon and the melting temperature of the assessed amplicon. The following data analysis workflow was applied to the results created by the software which came with the system OpenArray™ system: Data was collected for each combination of assays and samples in the two sets of samples used. Results were filtered using the following approach. Read outs from not loaded reaction spaces were removed from analysis. Technical Control assays were removed from the data set. Assays known to not work for other than biological reasons were removed from the analysis. Per sub-array, signals were only interpreted if there was a positive beta-Actin call. Ct values > 0 for each gene were normalized using the Ct values collected for the gene beta-Actin. This resulted in two files containing the results for each set of sample. 201 samples were tested of which 6 gave invalid results. In total 79,170 reactions were performed of which 74,110 were valid. For the data analysis, 2 boundaries were defined: an upper

bound on beta-Actin-normalized-Ct (banCt) and a lower bound on Melting Temperature (Tm). Samples below the banCt boundary and above the Tm boundary are considered to be "methylated", others (including all samples with no signal, i.e. Ct>40) are classified as "unmethylated". In both dimensions the set of candidate boundaries consists of all values in between 2 measurements, plus infinity (the equivalent of no boundary). The set of candidate models for "methylated" then consists of all combinations of candidate Tm lower bound and a banCt upperbound. A score is computed for each of these candidate models, as follows. Count: cancers inside boundaries = true positives (TP), cancers outside boundaries = false negatives (FN), normals inside boundaries = false positives (FP), normals outside boundaries = true negatives (TN). A binomial test was applied to find out how unusual it is to have at least TP successes in (TP+FP) trials where the probability of success is (TP+FN). The lower this probability value is the better. Then quality control data were taken into account to determine the most robust boundaries. Using the standard deviations (StDevQC) observed in the QC, a series of increasingly "noisy" datasets were generated. The measurements are replaced by a value randomly selected from a normal distribution with average equal to the observed measurement and standard deviation equal to StDevQC multiplied by a value that gradually (10 noise levels) increases from 0 to 2. Each time the score of the candidate model is computed by applying the 2 steps above (i.e., count and binomial test). All these scores (11 in total: 1 for "no noise" and 10 for noise levels 0.2, 0.4, ..., 2) are added up to obtain the ultimate accumulated score. The candidate model with the best (i.e. lowest) accumulated score is retained. This same score of the best candidate model for each marker is also used for ranking the markers.

RESULTS

A high throughput, real-time methylation specific detection platform was applied on two groups of samples isolated from cervical cancer tissue and from corresponding normal cervical tissue. In this study it was shown that a number of genes are differentially methylated in cervical cancer. We identified 112 different assays for detecting 96 different genes being differentially methylated in human cervical cancer tissue and normal cervical tissue control samples. The genes identified are ALX3, ALX4, AR, ARID4A, ATM, AURKA, B4GALT1, BMP2, BMP6, BNIP3, C13orf18, C16orf48, C9orf19, CALCA, CAMK4, CCNA1, CCND2, CDH1, CDH4, CDK6, CDKN1B, CDKN2B, CLSTN2, CLU, COL1A1, CPT1C, CTDSPL, CYCLIND2, DAPK1, DBC1, DDX19B, DKK2, EGFR, EGR4, EPB41L3, FOS, FOXE1,

GADD45A, GATA4, GDAP1L1, GNB4, Gst-Pi, HHIP, HOOK2, HOXA1, HOXA11, HOXA7, IGSF4, ISYNA1, JAM3, JPH3, KNDC1, KRAS, LMX1A, LOC285016, LOX, MTAP, MYO18B, NOL4, NPTX1, OGFOD2, PAK3, PAX1, PDCD4, PHACTR3, POMC, PRKCE, RAD23B, RALY, RARA, RBP4, RECK, RPRM, SEMA3F, SLC5A8, SLIT1, SLIT2, SLIT3, SMPD1, SOCS1, SOX1, SPARC, SPN, SST, TERT, TFPI-2, TLL1, TNFAIP1, TRMT1, TWIST1, UGT1A1, WIF1, WIT1, WT1, XRCC3, and ZGPAT

The resulting assays have the assay details provided in Table 1, Table 2, and Figure 5B.

Example 3: Further Assay selection: Base 5 – Lightcycler platform

Of the different assays listed in Table 1 previously identified using the Base5 methylation platform, the top 63 ranked assays plus β-actin (ACTB) were transferred to the Lightcycler platform in order to further fine-tune the selection of the best cervical cancer methylation markers. This platform allows the assessment of markers in a system which is closer to, and provides information valuable for the subsequent development of, a final, scaled up MSP assay. The 64 assays (Table 6) were applied on a 384 well plate by Sigma. Six repeats of the assay set fitted on a 384 well plate. The samples were randomized per plate.

The sample set selected for the Lightcycler analysis was also previously used in the Base 5 analysis in order to make a compared analysis: a total of 27 cervical tumor samples and 20 controls (frozen tissue) were collected by UMC Groningen.

Table 6: The 64 selected assays which were applied on the Lightcycler platform

N°	Assays	Base 5 ranking
1	LMX1A_9513	1
2	SLIT2_23681	2
3	ISYNA1_19726	3
4	EPB41L3_19071	4
5	WT1_1	5
6	DKK2_23973	6
7	ALX3_25180	7
8	JAM3	8
9	JPH3_12611	9
10	SLIT2_23672	10
11	SOX1_27153	11
12	SOX1_27159	12
13	RALY_19607	13

14	RPRM_2	14
15	CDH4_24735	15
16	CPT1C_23912	16
17	SLIT2_23676	17
18	PAX1_27211	18
19	DKK2_23970	19
20	TERT_23702	20
21	NOL4_19645	21
22	HOXA11_23844	22
23	CALCA_2	23
24	C13orf18_19885	24
25	PAX1_27210	25
26	WIT1_24567	26
27	GATA4_13295	27
28	SLIT1_23651	28
29	LOC285016_22940	29
30	POMC	30
31	Gst-Pi_New3	32
32	DAPK1	34
33	GDAP1L1_19773	35
34	TFPI-2	36
35	TWIST1_9329	37
36	SST_23808	38
37	EGR4_24277	39
38	C16orf48_22922	45
39	DBC1_23879	46
40	GDAP1L1_19775	47
41	OGFOD2_23131	48
42	ALX4_25062	49
43	TLL1_24051	51
44	CTDSPL_23795	52
45	CYCLIND2_1	58
46	COL1A1_23253	65
47	CDK6_9703	71
48	CDH1_17968	76
49	SOCS1_23595	78
50	FOXE1_13314	91
51	BMP2_17901	94
52	AURKA_24802	110
53	SEMA3F_23485	120
54	PAK3_3	121
55	HOXA7_2	125
56	CTDSPL_23804	127
57	NPTX1_2	136
58	SLIT1_23653	164
59	SMPD1_24061	174

60	GADD45A_24463	250
61	KRAS_24235	281
62	RECK_18940	321
63	UGT1A1_22912	341
64	Beta_Actin	Internal control

Tissue slides were deparaffinized using 100% xylene followed by 100% ethanol. Pellet was resuspended in a buffer containing SDS-proteinase K, and DNA was extracted with phenol-chloroform followed by ethanol precipitation. DNA concentration was measured using NanoDrop Spectrophotometer. From each sample, up to 3 µg of genomic DNA was converted using a bisulphite based protocol (EZ DNA Methylation Kit™, ZYMO Research). After DNA conversion and purification, equivalent of 20 ng of gDNA was used per reaction. All the samples were tested on Lightcycler using Sybergreen as detector and the amplicon size was determined by capillary electrophoresis.

Quality control was performed using *in vitro* methylated DNA sample, unmethylated DNA sample (Chemicon International, CA, USA; Cat.# S7821 and Cat.# S7822) and no template control sample (H_2O). From the Lightcycler platform, the Ct values (cycle number at which the amplification curves cross the threshold value, set automatically by the software) and melting curves (Tm) were generated. From the capillary electrophoresis platform, size of the amplicon and intensity of the signal detected were generated. For each assay, Tm and amplicon size parameters were determined in *in vitro* methylated DNA sample, unmethylated DNA sample and no template control sample. The measured Tm and amplicon size values were compared to the calculated values. If the Tm or amplicon size values were out of the range of the calculated ones, the assay was considered as non specific and disqualified. All the 64 assays were specific.

A sample is considered methylated if Ct is under 40 and if Tm and amplicon size are within the boundaries of Tm +/- 2 degrees and amplicon size +/- 10 bp. The intensity of the band detected by capillary electrophoresis had to be higher than 20. Those evaluation criteria have been developed based on concordance with existing Molecular Beacon based qMSP assays, to ensure that the conclusions drawn from these data would be predictive of MSP assays developed subsequently.

DNA methylation calls were compared between cervical cancer and control patients. An assay ranking with the set of samples was generated and the results are summarized in the methylation table of Figure 6. A one-tailed Fisher's exact test was used as a scoring function to rank the candidate markers. The calculation of Fisher's exact test

was based on a formula as described by Haseeb Ahmad Khan in "A visual basic software for computing Fisher's exact probability" (*Journal of Statistical Software*, vol. 08, issue i21, 2003).

A comparison between the results coming from the Base 5 (Biotrove) and the Lightcycler platforms has been performed. Most of the interesting assays discovered on the Base 5 platform were confirmed on the Lightcycler platform.

Example 4: QMSP

Seventeen assays (ALX3, C13ORF18, DBC1, EPB41L3, GATA4, HOXA11, JAM3, JPH3, LMX1A, NOL4, PAK3, SLIT2_23672, SLIT2_23676, SOX1, TERT, TFPI2 and TWIST1_3) were further selected based on their performance on the Biotrove and Lightcycler platforms and on complementarity analysis to maximize discriminatory power. For these assays, qMSPs using Molecular Beacon as detection system were designed (3 designs, if possible, were evaluated per assay) and tested on control samples. For this selection, assays were judged on several criteria, including background fluorescence, dynamic of the curve, and level of fluorescence generated. PCR material was used for generating standard curves for quantification of the results.

Five assays did not meet the desired specifications (EPB41L3, SOX1, SLIT2_23672, DBC1, and SLIT2_23676) and may be redesigned in a later phase or can be used on another detection platform. The remaining 12 assays were further tested on converted DNA of cervix cancer cell lines.

All these results were taken into account to decide which assays should be further verified on cervical tissue samples collected by Ulg (normal PE tissue samples #13, cancer PE tissue samples #17) and/or UMCG (normal frozen tissue samples #20, cancer frozen tissue samples #27).

Seventeen (CCNA1, CDO1_55928, CDO1_55929, GREM1, GPNMB, HIN1, HOXD1, LAMA1, LTB4R, MAL, NDRG2, NID2, NPTX2, RASSF1A, SALL4, SOX17, and TAC1) additional good performing assays were also selected for further verification on the cervix tissue samples. These candidates were taken from other in-house cancer projects, and were not tested on the Biotrove/Lightcycler platform as described above.

DNA was isolated from the cervix tissue samples using a phenol-chloroform procedure, quantified using the picogreen method and 1.5 µg of DNA was bisulphite treated using the ZYMO kit.

qMSPs were carried out in a total volume of 12 µl in 384 well plates in an ABI PRISM 7900HT instrument (Applied Biosystems). The final reaction mixture consisted of in-house qMSP buffer (including 80.4 nmol of MgCl2), 60 nmol of each dNTP, 0.5 U of Jump Start Taq polymerase (SIGMA), 72 ng of forward primer, 216 ng of reverse primer, 1.92 pmol of Molecular Beacon detection probe, 6.0 pmol of ROX (passive reference dye) and 72 ng of bisulphite converted genomic DNA. Thermal cycling was initiated with an incubation step of 5 minutes at 95°C, followed by 45 cycles (95°C for 30 seconds, 57°C for 30 seconds, 72°C for 30 seconds). A finalizing step was performed at 72°C for 5 minutes to conclude cycling. These conditions were similar for all the test genes as well as for ACTB. Cell lines [*in vitro* methylated DNA sample and unmethylated DNA sample (Chemicon International, CA, USA; Cat.# S7821 and Cat.# S7822)] were included in each run as positive and negative controls, and entered the procedure at the DNA extraction step. Primers and molecular beacon sequences used for the different qMSPs are summarized in Table 1 and Table 3. Corresponding amplicons are summarized in Table 2.

Ct values were determined using the SDS software (version 2.2.2.) supplied by Applied Biosystems with automatic baseline settings and threshold. The slopes and R^2 values for the different standard curves were determined after exporting data into MS Excel.

As an example, Figure 7 shows the amplification plot obtained for the standard curve for *TAC1_56187* (960000 copies to 9.6 copies of the gene) and Figure 8 shows the amplification plot obtained for the standard curve and for all samples for *TAC1_56187*. The Ct values plotted against the Log Copies of *TAC1_56187* (Figure 9) give a R^2 of 0.9995 and the efficiency of the reaction is 99.35%.

In addition to the test genes, the independent reference gene β-actin (ACTB) was also measured. The ratios between the test genes and ACTB were calculated to generate the test result. The samples were classified as methylated, unmethylated, or invalid based on the decision tree shown in Figure 10.

A provisional cut-off was defined for each gene, chosen based on the greater of either the highest value seen among the controls or a value 3 times the standard deviation of the values from control samples.

The one-tailed Fisher's exact test as described above was used as a scoring function to rank the candidate markers (*Journal of Statistical Software*, vol. 08, issue i21, 2003).

Table 7 summarizes the results obtained for *TAC1_56187*. Table 8 summarizes the results obtained for all the tested markers on tissue samples. The individual

performances of the assays are shown in Figure 11 and the assays are ranked according their p-value (Fisher's exact test). The best performing markers were further tested on clinical samples (scrapings).

Table 7: Summary of the test results for *TAC1_56187* on cervix tissue samples. In column "methylation status", the black boxes indicate the methylated results; white boxes indicate the unmethylated results.

Sample	*TAC1_56187*_Ratio	Methylation Status
normal	5.0	0
normal	0.0	0
normal	0.0	0
normal	0.0	0
normal	4.3	0
normal	0.0	0
normal	1.4	0
normal	13.2	0
normal	3.5	0
normal	0.8	0
normal	0.0	0
normal	0.0	0
normal	3.0	0
normal	6.1	0
normal	2.5	0
normal	0.0	0
normal	2.2	0
normal	5.2	0
normal	0.8	0
normal	20.0	0
normal	0.0	0
normal	0.0	0
normal	0.0	0
normal	0.0	0
normal	0.0	0
normal	13.7	0
normal	0.0	0
normal	0.0	0
normal	0.0	0
normal	0.0	0
normal	0.0	0
normal	0.0	0
normal	0.0	0
normal	10.8	0
normal	9.8	0
normal	21.9	0
normal	15.8	0
normal	0.0	0
normal	0.0	0

N Normals	38

Sample	*TAC1_56187*_Ratio	Methylation Status
cancer	19.1	0
cancer	0.0	0
cancer	330.7	1
cancer	92.6	1
cancer	7.3	0
cancer	784.9	1
cancer	23.5	0
cancer	245.2	1
cancer	1235.0	1
cancer	0.0	0
cancer	1256.1	1
cancer	570.0	1
cancer	3.3	0
cancer	30.0	1
cancer	191.0	1
cancer	11.6	0
cancer	110.0	1
cancer	777.1	1
cancer	2944.3	1
cancer	66.5	1
cancer	843.7	1
cancer	208.7	1
cancer	257.0	1
cancer	750.0	1
cancer	542.5	1
cancer	378.5	1
cancer	720.5	1
cancer	323.6	1
cancer	937.3	1
cancer	657.8	1
cancer	482.7	1
cancer	772.8	1
cancer	781.5	1
cancer	67.5	1
cancer	742.9	1
cancer	797.1	1
cancer	304.9	1
cancer	765.0	1
cancer	331.0	1
cancer	1111.0	1
cancer	383.5	1
cancer	703.7	1
cancer	237.6	1
cancer	718.2	1
cancer	541.1	1
cancer	348.3	1

N Cancers	46
Sensitivity	85
Specificity	100
Cut off	25
STDEV (Normals) *3	18.0
Cncr Meth +	39
Cncr Meth -	7
Cntrl Meth +	0
Cntrl Meth -	38
p-value (Fisher test)	3.94E-17

Table 8: Summary of the performance results of all the tested markers on tissue samples.

Ranking Lightcycler	NA	12	NA	3	1	NA	NA	NA	21	8	7	10	NA	NA
Ranking Base5	NA	1	NA	21	8	NA	NA	NA	27	20	9	22	NA	NA
Ranking qMSP tissue	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Assays	**NID2_9091**	**LMX1A_9513**	**TAC1_56187**	**NOL4_19645**	**JAM3**	**CDO1_55929**	**CDO1_55928**	**SOX17_66072**	**GATA4**	**TERT_23702**	**JPH3**	**HOXA11_23844**	**LAMA1_63431**	**CCNA1_Gron**
Sensitivity	**95**	**85**	**85**	**91**	**83**	**88**	**81**	**80**	**71**	**65**	**72**	**65**	**59**	**60**
Specificity	**97**	**100**	**100**	**95**	**100**	**100**	**97**	**97**	**97**	**100**	**97**	**100**	**100**	**97**
Cut off	2	10	25	2	5	0	20	35	2	1	1	20	5	15
RatioMax (Normals)	2	10	22	2	5	0	28	51	2	1	1	20	3	15
STDEV (Normals) *3	1.5	6.5	18.0	1.9	3.5	0.0	17.7	32.0	1.3	0.5	0.6	14.2	2.2	11.6
Cncr Meth +	41	39	39	42	38	38	38	37	42	39	33	28	27	36
Cncr Meth -	2	7	7	4	8	5	9	9	17	21	13	15	19	24
Cntrl Meth +	1	0	0	2	0	0	1	1	1	0	1	0	0	1
Cntrl Meth -	31	38	38	36	38	32	37	37	38	38	37	32	38	37
p-value (Fisher test)	1.42 E-17	3.94 E-17	3.94 E-17	7.55 E-17	2.27 E-16	2.79 E-16	2.08 E-14	3.66 E-14	8.90 E-13	2.40 E-12	1.06 E-11	4.82 E-10	5.53 E-10	1.03 E-09

Ranking Lightcycler	15	NA	NA	16	NA	NA	NA	25	NA	NA	6	NA	NA	NA
Ranking Biotrove	24	NA	NA	36	NA	NA	NA	7	NA	NA	121	NA	NA	NA
Ranking qMSP tissue	15	16	17	18	19	20	21	22	23	24	25	26	27	28
Assays	**C13ORF18_Gron**	**MAL**	**GREM1_29777**	**TFPI2**	**GPNMB_52607**	**NDRG2_56603**	**NPTX2_57779**	**ALX3_25180**	**HIN1_3**	**SALL4_12833**	**PAK3_1**	**HOXD1 (2)**	**LTB4R_31250**	**RASSF1a**
Sensitivity	**65**	**53**	**51**	**60**	**50**	**59**	**55**	**55**	**50**	**33**	**30**	**30**	**60**	**45**
Specificity	**97**	**100**	**97**	**100**	**100**	**95**	**100**	**100**	**100**	**97**	**100**	**95**	**68**	**74**
Cut off	2	1	2	5	5	5	30	5	1	20	10	50	20	0
RatioMax (Normals)	3	0	3	5	4	6	29	3	0	35	6	59	28	0
STDEV (Normals)	1.8	0.3	1.8	3.1	2.7	4.5	24.3	2.0	0.2	18.5	4.2	45.2	17.9	0.2

*3														
Cncr Meth +	28	23	22	12	17	20	11	11	10	14	6	6	12	9
Cncr Meth -	15	20	21	8	17	14	9	9	10	29	14	14	8	11
Cntrl Meth +	1	0	1	0	0	1	0	0	0	1	0	1	6	5
Cntrl Meth -	31	32	31	19	19	18	19	19	19	32	19	18	13	14
p-value (Fisher test)	9.66 E-09	8.07 E-08	2.91 E-06	3.22 E-05	7.22 E-05	8.61 E-05	1.00 E-04	1.00 E-04	2.91 E-04	9.64 E-04	1.19 E-02	5.29 E-02	7.19 E-02	1.89 E-01

Example 5: Best performing markers tested on clinical cervical scraping samples

Cervical scraping samples were collected under the Cervical Cancer Clinical Collaborative Research Agreement study of ONCO with the Gynecology Department of the UMCG hospital. The scraping samples were taken from patients who were referred to the hospital with an abnormal PAP smear or because they were suspected for cervical carcinoma. Gynecological examination under general anesthesia was performed in all cervical cancer patients for staging in accordance with the International Federation of Gynecology and Obstetrics (FIGO) criteria. Control scraping samples were taken from women who visited the hospital for a non-malignant condition, e.g. fibroids, prolaps uteri or hypermenorrhea, and who were scheduled to undergo a hysterectomy. While the patient was under general anesthesia, the cervix was scraped with an Ayres spatula and brush. The scraped cells were suspended in 5-ml PBS. Cytospins for cytomorphological assessment were made (1/5 volume). Cytospins were Papanicolaou stained and routinely classified according to a modified Papanicolaou system (Hanselaar AG. Kwaliteit van cytopathologisch onderzoek in het herziene bevolkingsonderzoek naar baarmoederhalskanker. Nederlands Tijdschrift voor Obstetrie en Gynaecologie 1996; 109:207-210) without knowledge of the clinical data. The remaining 4-ml of the scraped cells was centrifuged, washed, aliquoted, snap-frozen in liquid nitrogen and stored at -80°C. DNA was extracted using standard salt-chloroform extraction and ethanol precipitation. DNA of the pellet was used for qMSP of a panel of good performing markers for cervical cancer and also for HPV typing.

DNA was extracted from the scraped cells using standard salt-chloroform extraction and ethanol precipitation for high molecular DNA, dissolved in 250 µL TE-4 buffer (10 mM Tris; 1 mM EDTA, pH 8.0) and kept at -20°C until tested.

Presence of high risk HPV was analyzed by PCR using HPV16 and HPV18 specific primers on DNA of the scraping samples. On all HPV16- or HPV18- negative cases, general primer-mediated PCR was performed using two HPV consensus primer sets,

CPI/CPIIG and GP5+/6+, with subsequent nucleotide sequence analysis, as described previously [by Wisman et al Int j cancer 2006].

qMSP was performed after bisulphite treatment on denatured genomic DNA. The assays were carried out as described above. The samples were classified as methylated, unmethylated, or invalid as described above. The results obtained for all the tested markers on scraping samples from cervical cancer patients and from control patients were ranked according their p-value (Fisher's exact test) (Table 9). Some markers have a higher sensitivity for squamous cell carcinoma than for adenocarcinoma (NID2, JPH3, CCNA1) and some markers have a higher sensitivity for adenocarcinoma than for squamous cell carcinoma (JAM3, CDO1, HOXA11).

Various combinations of markers were evaluated to see if such a combination could increase the sensitivity while still maintaining a high level of specificity. In all cases, if any marker of a combination panel was positive, the sample was classified as methylated. Examples of the performance of combination of markers are summarized in Table 10. It can be seen that several combinations provided a sensitivity and specificity greater than 90%.

Table 9: Summary of the results obtained for all the tested markers on scraping samples from cervical cancer patients and from control patients (Sens: sensitivity; SCC: squamous cell carcinoma; Ade: adenocarcinoma; cncr: cancer; ctrl: control).

	JAM3	NID2_9091	CDO1_55928	CDO1_55929	LMX1A_9513	TAC1_56187	GREM1_29777	HOXA11_23844	JPH3	GATA-4	C130RF18_Gro n	CCNA1_Gron	TERT_23702	NDRG2_56603	NOL4_19645	LAMA1_63431
Sensitivity	**81.0 %**	**78.5 %**	**82.3 %**	**78.5 %**	**75.9 %**	**72.2 %**	**72.2 %**	**62.0 %**	**64.6 %**	**62.0 %**	**53.2 %**	**51.9 %**	**58.2 %**	**49.4 %**	**43.0 %**	**51.9 %**
Specificity	**98.6 %**	**98.6 %**	**95.7 %**	**97.1 %**	**97.1 %**	**98.6 %**	**97.1 %**	**100.0 %**	**98.6 %**	**97.1 %**	**100.0 %**	**100.0 %**	**97.1 %**	**98.6 %**	**98.6 %**	**94.2 %**
Sens SCC	80.3%	83.3%	81.8%	77.3%	77.3%	72.7%	72.7%	59.1%	69.7%	62.1%	54.5%	57.6%	60.6%	48.5%	43.9%	50.0%
Sens Ade	84.6%	53.8%	84.6%	84.6%	69.2%	69.2%	69.2%	76.9%	38.5%	61.5%	46.2%	23.1%	46.2%	53.8%	38.5%	61.5%
cncr test+	64	62	65	62	60	57	57	49	51	49	42	41	46	39	34	41
cncr test-	15	17	14	17	19	22	22	30	28	30	37	38	33	40	45	38
ctrl test+	1	1	3	2	2	1	2	0	1	2	0	0	2	1	1	4
ctrl test-	68	68	66	67	67	68	67	69	68	67	69	69	67	68	68	65
SCC test+	53	55	54	51	51	48	48	39	46	41	36	38	40	32	29	33
SCC test-	13	11	12	15	15	18	18	27	20	25	30	28	26	34	37	33
Ade test+	11	7	11	11	9	9	9	10	5	8	6	3	6	7	5	8
Ade test-	2	6	2	2	4	4	4	3	8	5	7	10	7	6	8	5
p-val cncr/ctrl	4.75E -26	1.21E -24	4.57E -24	3.11E -23	6.24E -22	1.86E -21	4.17E -20	1.23E -18	4.12E -18	7.73E -16	2.91E -15	8.21E -15	2.03E -14	1.59E -12	1.61E -10	2.17E -10
p-val Ade/SCC	5.32E -01	2.31E -02	5.84E -01	4.33E -01	3.37E -01	4.79E -01	4.79E -01	1.86E -01	3.53E -02	6.43E -01	4.00E -01	2.33E -02	2.54E -01	4.80E -01	4.81E -01	3.25E -01
Cut off	2	5	5	35	15	15	10	1	5	2	0	1	5	150	5	10

Table 10: Examples of the performance of combination of markers on scraping samples from cervical cancer patients and from control patients (Sens: sensitivity; SCC: squamous cell carcinoma; Ade: adenocarcinoma; cncr: cancer; ctrl: control).

	NID2_9091 \ HOXA11_23844	JAM3 \ CDO1_55929 \ HOXA11_23844 \ CCNA1_Gron	JAM3 \ HOXA11_23844	JAM3 \ HOXA11_23844 \ GREM1_29777	JAM3 \ NID2_9091 \ HOXA11_23844 \ CDO1_55929	JAM3 \ TAC1_56187 \ HOXA11_23844 \ CDO1_55929	JAM3 \ HOXA11_23844 \ CDO1_55929	JAM3 \ CDO1_55928	JAM3 \ NID2_9091	NID2_9091 \ CDO1_55928
Sensitivity	**89.9%**	**92.4%**	**88.6%**	**91.1%**	**92.4%**	**92.4%**	**92.4%**	**89.9%**	**86.1%**	**88.6%**
Specificity	**98.6%**	**95.7%**	**98.6%**	**95.7%**	**94.2%**	**94.2%**	**94.2%**	**94.2%**	**97.1%**	**94.2%**
Sens SCC	92.4%	92.4%	89.4%	92.4%	92.4%	92.4%	92.4%	89.4%	86.4%	89.4%
Sens Ade	76.9%	92.3%	84.6%	84.6%	92.3%	92.3%	92.3%	92.3%	84.6%	84.6%
cncr test+	71	73	70	72	73	73	73	71	68	70
cncr test-	8	6	9	7	6	6	6	8	11	9
ctrl test+	1	3	1	3	4	4	4	4	2	4
ctrl test-	68	66	68	66	65	65	65	65	67	65
SCC test+	61	61	59	61	61	61	61	59	57	59
SCC test-	5	5	7	5	5	5	5	7	9	7
Ade test+	10	12	11	11	12	12	12	12	11	11
Ade test-	3	1	2	2	1	1	1	1	2	2
p-val cncr/ctrl	8.14E-32	6.60E-31	6.87E-31	6.62E-30	1.17E-29	1.17E-29	1.17E-29	9.85E-28	1.13E-27	7.67E-27

HPV testing will certainly continue to occupy a significant position in the diagnosis of cervical cancer. With this in mind, the best performing methylation markers were tested on scraping samples from patients who were referred to the hospital with an abnormal Pap smear and these samples were also tested for hr HPV and HPV16. The provisional cut off as defined above was reduced in order to obtain the highest possible sensitivity and specificity compared to the performance of hrHPV. The results of these tests are shown in Table 11. For these testing, the classification of pre-cancerous (CIN) conditions were used. Sensitivity was calculated for samples indicating cancer, CIN 2 and CIN 3, while specificity was calculated for those samples from controls, and those indicating CIN 1 or CIN 0 after cytological examination. Overall the specificity of the methylation markers was higher compared to hr-HPV or HPV16 testing but with a lower sensitivity. Combinations of methylation markers (where at least one of the markers scores positive) showed a comparable sensitivity and specificity for cancers and controls, but a much higher specificity for CIN0 and CIN1. The sensitivity for CIN3 and CIN2 is however somewhat lower. In order to increase the sensitivity for CIN3 and CIN2 detection, an analysis was made of combining the

results of methylation markers and HPV16 (Table 12). The sensitivity as well as the specificity increased if HPV16 was combined with the methylation markers.

Table 11: Overall summary of the methylation marker(s) results on scraping samples from patients who were referred to the hospital with an abnormal Pap smear, and from cervical cancer and control patients. (Sens: sensitivity; Spec: specificity; CIN0, CIN1, CIN2, CIN3: cervical intraepithelial neoplasia grade 0, 1, 2, and 3; cncr: cancer; ctrl: control, NA: not applicable).

	hr-HPV	HPV16	JAM3	NID2_9091	LMX1A_9513	CDO1_55928	TAC1_56187	C13ORF18_Gron	JAM3 \ NID2_9091	JAM3 \ LMX1A_9513	JAM3 \ CDO1_55928	JAM3 \ CDO1_55928 \ NID2_9091	NID2 \ LMX1A_9513
Sens Cncr	90%	77%	83%	80%	82%	83%	73%	54%	88%	87%	90%	91%	88%
Sens CIN3	95%	83%	38%	40%	60%	43%	17%	24%	45%	62%	52%	57%	67%
Sens CIN2	74%	45%	21%	29%	29%	24%	7%	5%	33%	36%	31%	38%	43%
Spec Cntrl	96%	99%	99%	93%	94%	91%	93%	100%	93%	93%	90%	84%	88%
Spec CIN0	51%	91%	98%	95%	91%	98%	100%	100%	95%	93%	98%	95%	91%
Spec CIN1	34%	78%	98%	93%	85%	88%	100%	98%	93%	83%	88%	83%	83%
Overall sens	87%	70%	56%	57%	63%	58%	42%	34%	63%	67%	66%	69%	71%
Overall spec	67%	91%	98%	93%	91%	92%	97%	99%	93%	90%	92%	87%	88%
Cut off	NA	NA	1	2	10	3	10	0	NA	NA	NA	NA	NA
cncr test+	74	63	68	66	67	68	60	44	72	71	74	75	72
cncr test-	8	19	14	16	15	14	22	38	10	11	8	7	10
CIN3 test+	40	35	16	17	25	18	7	10	19	26	22	24	28
CIN3 test-	2	7	26	25	17	24	35	32	23	16	20	18	14
CIN2 test+	31	19	9	12	12	10	3	2	14	15	13	16	18
CIN2 test-	11	23	33	30	30	32	39	40	28	27	29	26	24
ctrl test+	3	1	1	5	4	6	5	0	5	5	7	11	8
ctrl test-	66	68	68	64	65	63	64	69	64	64	62	58	61
CIN0 test+	21	4	1	2	4	1	0	0	2	3	1	2	4
CIN0 test-	22	39	42	41	39	42	43	43	41	40	42	41	39
CIN1 test+	27	9	1	3	6	5	0	1	3	7	5	7	7
CIN1 test-	14	32	40	38	35	36	41	40	38	34	36	34	34

Table 12: Overall summary results of methylation marker(s) in combination with HPV16 on scraping samples from patients who were referred to the hospital with an abnormal Pap smear, and from cervical cancer and control patients. (Sens: sensitivity; Spec: specificity; CIN0, CIN1, CIN2, CIN3: cervical intraepithelial neoplasia grade 0, 1, 2, and 3; cncr: cancer; ctrl: control).

	hr-HPV	HPV16	JAM3 \ HPV16	NID2_9091 \ HPV16	LMX1A_9513 \ HPV16	CDO1_55928 \ HPV16	TAC1_56187 \ HPV16	C13ORF18_Grom \ HPV16	JAM3 \ NID2_9091 \ HPV16	JAM3 \ LMX1A_9513 \ HPV16	JAM3 \ CDO1_55928 \ HPV16	JAM3 \ CDO1_55928 \ NID2_9091 \ HPV16	NID2 \ LMX1A_9513 \ HPV16
Sens Cncr	90 %	77 %	95 %	93 %	94 %	99 %	93 %	83 %	98 %	96 %	100 %	100%	96%
Sens CIN3	95 %	83 %	88 %	90 %	88 %	88 %	83 %	86 %	90 %	88 %	88%	90%	90%
Sens CIN2	74 %	45 %	60 %	62 %	60 %	60 %	50 %	45 %	67 %	64 %	64%	67%	67%
Spec Cntrl	96 %	99 %	97 %	93 %	93 %	90 %	91 %	99 %	93 %	91 %	88%	84%	88%
Spec CIN0	51 %	91 %	88 %	86 %	84 %	88 %	91 %	91 %	86 %	84 %	88%	86%	81%
Spec CIN1	34 %	78 %	78 %	73 %	66 %	73 %	78 %	76 %	73 %	66 %	73%	68%	66%
CIN0 test-	22	39	38	37	36	38	39	39	37	36	38	37	35
CIN0 test+	21	4	5	6	7	5	4	4	6	7	5	6	8
CIN1 test-	14	32	32	30	27	30	32	31	30	27	30	28	27
CIN1 test+	27	9	9	11	14	11	9	10	11	14	11	13	14
CIN2 test-	11	23	17	16	17	17	21	23	14	15	15	14	14
CIN2 test+	31	19	25	26	25	25	21	19	28	27	27	28	28
CIN3 test-	2	7	5	4	5	5	7	6	4	5	5	4	4
CIN3 test+	40	35	37	38	37	37	35	36	38	37	37	38	38
cncr test-	8	19	4	6	5	1	6	14	2	3	0	0	3
cncr test+	74	63	78	76	77	81	76	68	80	79	82	82	79
ctrl test-	66	68	67	64	64	62	63	68	64	63	61	58	61
ctrl test+	3	1	2	5	5	7	6	1	5	6	8	11	8
Overall sens	87 %	70 %	84 %	84 %	84 %	86 %	80 %	74 %	88 %	86 %	88%	89%	87%
Overall spec	67 %	91 %	90 %	86 %	83 %	85 %	88 %	90 %	86 %	82 %	84%	80%	80%

As cytology is currently been used and hr-HPV testing has been suggested as primary screening tool in population-based cervical screening, we simulated the effect on the performances of the methylation tests if only cytology (Table 13) or hr-HPV (Table 14) positive patients were selected. The triage simulations were based on the performance results obtained in Table 11 and Table 12. The performance of cytology and hr-HPV testing were based on data from literature.

The performances of the triage tests showed much higher specificity resulting in fewer referrals for colposcopy than did cytology or hr-HPV testing alone but were less sensitive. Testing for hr-HPV types has a higher sensitivity for detecting CIN2+ than cytology. The NPV is close to 100% thus allowing for less frequent screening and longer screening intervals without jeopardizing patients' safety. But, the enthusiasm for using HPV testing in primary screening has been tempered by its somewhat poorer PPV (19%) in comparison with cytological analysis (27%). Using methylation as triage test, the PPVs were much higher.

Taking the limitations of cytology and the decreased disease prevalence due to the introduction of HPV vaccination programs into account, it is proposed to use a highly sensitive and objective screening test such as HPV DNA testing to identify the rare cases of cancer precursors and to combine it, when positive, with another test which has a high degree of specificity, such as methylation testing. Moreover, methylation is measuring changes in the host cells, as precursor of cervix cancer, while HPV is detecting the causative agent. This is an ideal methodology for a screening and a triage assay because they should measure different but complementary biological signals.

Table 13: The simulation of the performance of Cytology test as a first-line screening test on 70000 women and the methylation marker test(s) in- or excluding HPV16 as triage test. (Sens: sensitivity; Spec: specificity; CIN0/1, CIN2+: cervical intraepithelial neoplasia grade 0 and 1, and grade 2 and 3 and cancers; PPV: positive predictive value; NPV: negative predictive value).

	Cytology	Cytology, Triage hr-HPV	Cytology, Triage HPV16	Cytology, Triage JAM3 \ NID2	Cytology, Triage JAM3 \ NID2 \ HPV16	Cytology, Triage NID2 \ LMX1A	Cytology, Triage NID2 \ LMX1A \ HPV16
CIN2+ Test +	540	449	334	213	417	290	417
CIN2+ Test -	230	321	436	557	353	480	353
CIN0/1 Test +	1460	825	219	85	288	187	374
CIN0/1 Test -	67770	68405	69011	69145	68942	69043	68856
Sens	**70.1%**	**58.3%**	**43.4%**	**27.6%**	**54.2%**	**37.7%**	**54.2%**
Spec	**97.9%**	**98.8%**	**99.7%**	**99.9%**	**99.6%**	**99.7%**	**99.5%**
NPV	**99.7%**	**99.9%**	**99.7%**	**99.5%**	**99.8%**	**99.6%**	**99.8%**
PPV	**27.0%**	**35.2%**	**60.4%**	**71.4%**	**59.1%**	**60.8%**	**52.7%**
Colposcopy referrals	**2000**	**1274**	**553**	**298**	**706**	**477**	**791**

Table 14: The simulation of the performance of hr-HPV test as a first-line screening test on 70000 women and the methylation marker test(s) in- or excluding HPV16 as triage test. (Sens: sensitivity; Spec: specificity; CIN0/1, CIN2+: cervical intraepithelial neoplasia grade 0 and 1, and grade 2 and 3 and cancers; PPV: positive predictive value; NPV: negative predictive value).

	hr-HPV	hr-HPV, Triage Cytology	hr-HPV, Triage HPV16	hr-HPV, Triage JAM3 \ NID2	hr-HPV, Triage JAM3 \ NID2 \ HPV16	hr-HPV, Triage NID2 \ LMX1A	hr-HPV, Triage NID2 \ LMX1A \ HPV16
CIN2+ Test +	665	532	433	276	541	376	541

CIN2+ Test -	67	285	333	491	226	390	226
CIN0/1 Test +	2835	553	420	164	553	358	717
CIN0/1 Test -	66433	68631	68814	69070	68681	68875	68517
Sens	**90.8%**	**65.1%**	**56.5%**	**36.0%**	**70.6%**	**49.1%**	**70.6%**
Spec	**95.9%**	**99.2%**	**99.4%**	**99.8%**	**99.2%**	**99.5%**	**99.0%**
NPV	**99.9%**	**99.6%**	**99.6%**	**99.4%**	**99.8%**	**99.5%**	**99.8%**
PPV	**19.0%**	**49.0%**	**50.8%**	**62.7%**	**49.5%**	**51.2%**	**43.0%**
Colposcopy referrals	**3500**	**1085**	**853**	**440**	**1094**	**735**	**1258**

REFERENCES

Badal et al, The human papillomavirus-18 genome is efficiently targeted by cellular DNA methylation, Virology 324 (2004) 483– 492.

Barringer KJ, Orgel L, Wahl G, Gingeras TR. Gene. 1990 Apr 30; 89(1):117-22.

Brink et al, HPV detection methods, Disease Markers 23 (2007) 273-281.

Cottrell, SE et al. A real-time PCR assay for DNA-methylation using methylation-specific blockers. Nuc Acids Res 2004, 32, 1.

Cross, SH et al. Nature Genetics 1994, 6, 236-244.

Deng, D. et al. Simultaneous detection of CpG methylation and single nucleotide polymorphism by denaturing high performance liquid chromatography. 2002 Nuc. Acid Res, 30, 3.

Ehrich, M. et al. 2005. Introduction to DNA methylation analysis using the MassARRAY system. SEQUENOM™ product preview note.

Finan, M.A. et al. (1996) Radical hysterectomy for stage IB1 vs IB2 carcinoma of the cervix: does the new staging system predict morbidity and survival? *Gynecol Oncol.*, 2, 139-147.

Galm et al. Enzymatic Regional Methylation Assay: A Novel Method to Quantify Regional CpG Methylation Density. Genome Research, Vol. 12, Issue 1, 153-157, January 2002.

Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, et al. (2004) Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 5: R80.

Gitan RS et al. Methylation-specific oligonucleotide microarray: a new potential for high-throughput methylation analysis. 2006 Genome Research 12:158-164.

Gonzalgo and Jones, Rapid quantitation of methylation differences at specific sites using methylation–sensitive single nuclear primer extension. 1997 Nuc Acid Res, 25, 12.

Kalantari et al, Conserved Methylation Patterns of Human Papillomavirus Type 16 DNA in Asymptomatic Infection and Cervical Neoplasia, Journal of virology, Dec. 2004, p. 12762–12772.

Kang et al, Inverse correlation between RASSF1A hypermethylation, KRAS and BRAF mutations in cervical adenocarcinoma, Gynecol Oncol 2007 Jun; 105(3):662-6. Epub 2007 Mar 13.

Keating et al, Ki-67, Cyclin E and p16 (INK4A) are complementary surrogate biomarkers for human papillomavirus-related cervical neoplasia. Am J Surg Pathol 25, 884-891 (2001).

Kitkumthorn, N. et al. (2006) Cyclin A1 promoter hypermethylation in human papillomavirus-associated cervical cancer. *Bmc Cancer*, 6.

Koopman,L.A. et al. (1999) Recurrent integration of human papillomaviruses 16, 45, and 67 near translocation breakpoints in new cervical cancer cell lines. *Cancer Research*, 59, 5615-5624.

Kwoh DY, Davis GR, Whitfield KM, Chappelle HL, DiMichele LJ, Gingeras TR. Transcription-based amplification system and detection of amplified human immunodeficiency virus type 1 with a bead-based sandwich hybridization format. Proc Natl Acad Sci U S A. 1989 Feb; 86(4):1173-7.

Li KB. ClustalW-MPI: ClustalW analysis using distributed and parallel computing.

Bioinformatics 2003;19(12):1585-6.

Liu, G. et al. (2003) NetAffx: Affymetrix probesets and annotations. *Nucleic Acids Res.*, 31, 82-86.

Malinowski D. (2007) Multiple biomarkers in molecular oncology. Molecular diagnostics application in cervical cancer detection. *Expert Rev. Mol. Diagn. 7(2)*, 117-131.

Rao et al., Chromosomal amplifications, 3q gain and deletions of 2q33-q37 are the frequent genetic changes in cervical carcinoma. BMC Cancer 4(1), 5 (2004).

Rebhan, M. et al. (1997) GeneCards: Integrating information about genes, proteins and diseases. *Trends in Genetics*, 13, 163.

Shiraisi, M et al. Biol Chem. 1999, 380(9):1127-1131.

Straub, J. et al., A64-AACRMD (2007): Base5, a versatile, highly integrated high-throughput methylation profiling platform for Methylation-Specific PCR based marker identification applied to CRC.

Suzuki Y, Yamashita R, Sugano S, Nakai K. DBTSS, DataBase of Transcriptional Start Sites: progress report 2004. Nucleic Acids Res 2004; 32(Database issue):D78-81.

Thomassin H. et al. MethylQuant: a sensitive method for quantifying methylation of specific cytosines within the genome. 2004 . Nuc Acid Res 32, 21.

Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994; 22(22):4673-80.

Tost, J. et al. Analysis and accurate quantification of CpG methylation by MALDI mass spectrometry. Nuc. Acid Res, 2003, 31(9): e50.

Trinh B. et al. DNA methylation analysis by MethyLight technology. Methods. 2001 Dec; 25(4).

Trooskens G, De Beule D, Decoutere F, Van Criekinge W. Phylogenetic trees: visualizing, customizing and detecting incongruence. Bioinformatics 2005; 21(19):3801-2.

van Dongen,J.J. et al. (2003) Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. *Leukemia*, 17, 2257-2317.

Virmani et al. (2001) Aberrant methylation during cervical carcinogenesis. *Clin Cancer Research*, Mar; 7(3):584-9.

Xiong, Z.G. and Laird, P.W. (1997) COBRA: A sensitive and quantitative DNA methylation assay. *Nucleic Acids Research*, 25, 2532-2534.

Zeschnigk M. *et al.* A novel real-time PCR assay for quantitative analysis of methylated alleles (QAMA): analysis of the retinoblastoma locus. 2004. Nuc Acid Res 32, 16.

CLAIMS

1. A method for identifying, in a test sample, cervical cancer cells, or cervical cells that are predisposed to cancer, said method comprising: providing a test sample comprising cervical cells or nucleic acids from cervical cells; detecting in said test sample methylation of gene erythrocyte membrane protein band 4.1 like 3 (EPB41L3), wherein methylation in said test sample indicates the presence of cervical cancer cells, or cervical cells that are predisposed to cancer, or the presence of nucleic acids from cervical cancer cells or cervical cells that are predisposed to cancer.

2. Method for cervical cancer detection or screening comprising the steps of:

a) providing a test sample comprising cervical cells or nucleic acids from cervical cells;

b) assaying the test sample of step a) for high-risk human papillomavirus (hr- HPV);

c) if b) is positive for the presence of hr-HPV, assaying the methylation status of gene erythrocyte membrane protein band 4.1 like 3 (EPB41L3);

d) if the gene of c) is methylated, refer the woman for colposcopy in order to detect cervical cancer; and

e) if the gene of c) is unmethylated, refer the woman to a more regular screening for the presence of hr-HPV.

3. Method for cervical cancer detection or screening comprising the steps of:

a) providing a test sample comprising cervical cells or nucleic acids from cervical cells;

b) assaying the test sample of step a) for high-risk human papilloma virus (hr-HPV);

c) if b) is positive for the presence of hr-HPV, assaying the methylation status of gene erythrocyte membrane protein band 4.1 like 3 (EPB41L3) and typing the hr-HPV as HPV16 and/or HPV18;

d) if the gene of c) is methylated, and/or the sample is typed HPV16 and/or HPV18 positive, refer the woman for colposcopy in order to detect cervical cancer; and

e) if the gene of c) is unmethylated, refer the woman to a more regular screening for the presence of hr-HPV.

4. The method according to any one of claims 1-3, wherein the test sample comprises squamous cells, nucleic acids from squamous cells, adenocarcinoma cells, nucleic acids from adenocarcinoma cells, adenosquamous cell carcinoma cells, nucleic acids from adenosquamous carcinoma cells, or any combination thereof.

5. The method according to claim 1, wherein the test sample is from a specimen selected from the group consisting of a tissue specimen, a biopsy specimen, a surgical specimen, a cytological specimen, cervical scrapings, cervical smear, cervical washing, vaginal excretions, and blood.

6. The method of claim 5, wherein the test sample is from a biopsy specimen.

7. The method of claim 5, wherein the specimen is a surgical specimen.

8. The method according to claim 1, wherein methylation of at least two genes is detected including the EPB41L3 gene and a gene selected from junctional adhesion molecule C (JAM3) and telomerase reverse transcriptase (TERT).

9. The method according to any one of claims 1-3, wherein methylation is detected by detecting methylation of a CpG dinucleotide motif in the EPB41L3 gene, or a promoter region thereof.

10. The method of claim 9, wherein methylation is detected by: contacting at least a portion of the EPB41L3 gene or promoter region thereof of the test sample with a chemical reagent that selectively modifies a non-methylated cytosine residue relative to a methylated cytosine residue, or selectively modifies a methylated cytosine residue relative to a non-methylated cytosine residue; and detecting a product generated due to said contacting.

11. The method of claim 10, wherein the step of detecting a product employs amplification with at least one primer that hybridizes to a sequence comprising a modified non-methylated CpG dinucleotide motif but not to a sequence comprising an unmodified methylated CpG dinucleotide motif thereby forming amplification products.

12. The method of claim 10, wherein the step of detecting a product comprises amplification with at least one primer that hybridizes to a sequence comprising an unmodified methylated CpG dinucleotide motif but not to a sequence comprising a modified non-methylated CpG dinucleotide motif thereby forming amplification products.

13. The method of claim 10, wherein the product is detected by a method selected from the group consisting of electrophoresis, hybridization, amplification, sequencing, ligase chain reaction, chromatography, mass spectrometry, and combinations thereof.

14. The method of claim 10, wherein the chemical reagent comprises bisulfite ions.

15. The method of claim 14 further comprising treating the bisulfite ion-contacted portion of the EPB41L3 gene with alkali.

16. The method according to any one of claims 1-3, wherein the step of detecting or assaying employs amplification of at least a portion of the EPB41L3 gene using an oligonucleotide primer that specifically hybridizes under amplification conditions to a region of the EPB41L3 gene, wherein the region is within 10 kb of said gene's transcription start site, and the oligonucleotide primer is selected from SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:177, and SEQ ID NO:178.

17. The method according to any one of claim 1-3, wherein the step of detecting or assaying employs amplification of at least a portion of the EPB41L3 gene using at least one pair of oligonucleotide primers that specifically hybridizes under amplification conditions to a region of the EPB41L3 gene, wherein the region is within 10 kb of said gene's transcription start site, and the oligonucleotide primer pair is selected from SEQ ID NO:45 and SEQ ID NO:177, and SEQ ID NO:46 and SEQ ID NO:178.

18. A kit for assessing cervical cancer, or predisposition to cervical cancer in a test sample containing cervical cells or nucleic acids from cervical cells, said kit comprising in a package:

- a reagent that (a) modifies methylated cytosine residues but not non-methylated cytosine residues, or that (b) modifies non-methylated cytosine residues but not methylated cytosine residues; and

- at least one pair of oligonucleotide primers that specifically hybridizes under amplification conditions to a region of gene EPB41L3, wherein the region is within 10 kb of said gene's transcription start site, and the oligonucleotide primer pair is selected from SEQ ID NO:45 and SEQ ID NO:177, and SEQ ID NO:46 and SEQ ID NO:178.

19. The kit of claim 18 further comprising at least one oligonucleotide probe which hybridizes to an amplicon comprising SEQ ID NO:309 or SEQ ID NO:310.

20. The kit according to claim 18 or 19 further comprising a DNA polymerase for amplifying DNA.

Figure 1A

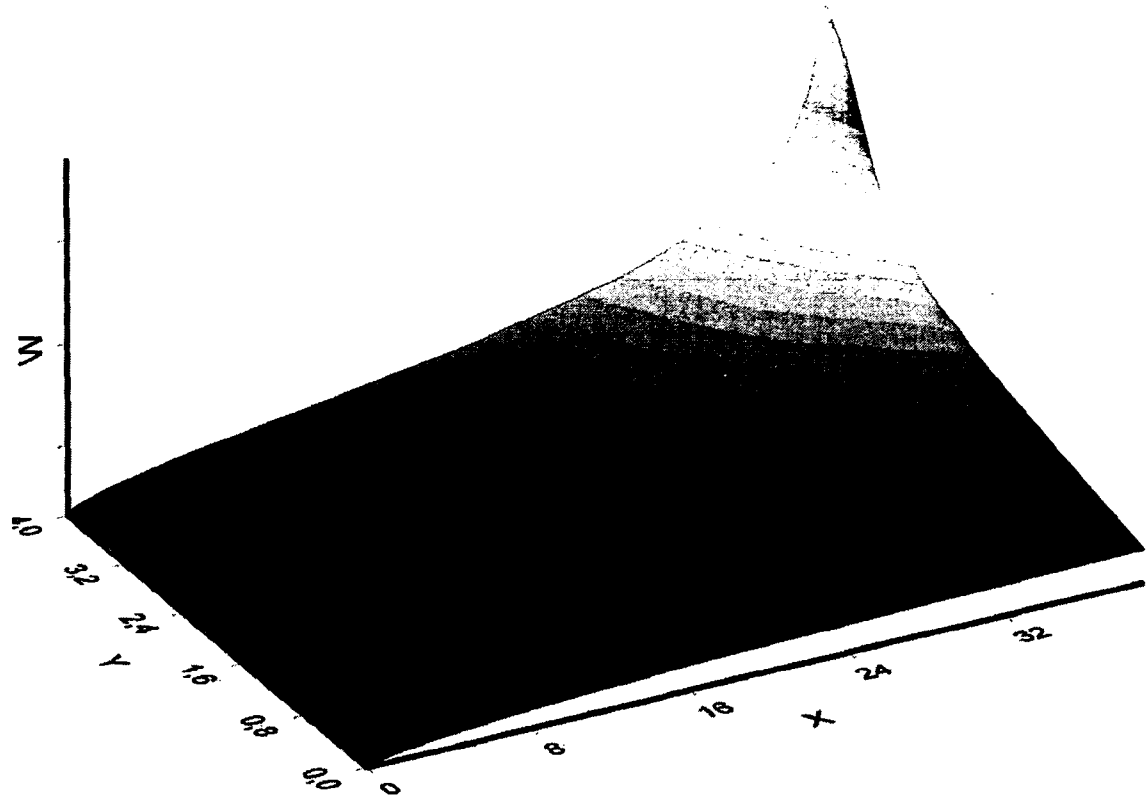


Figure 1B

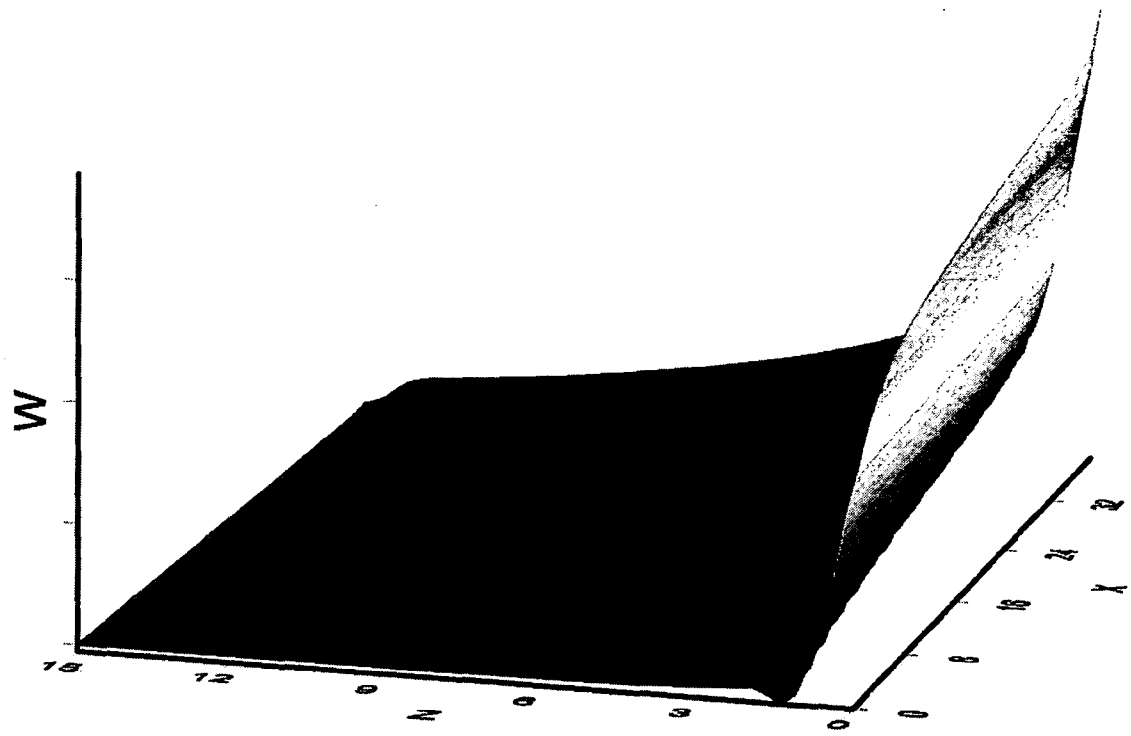


Figure 1C

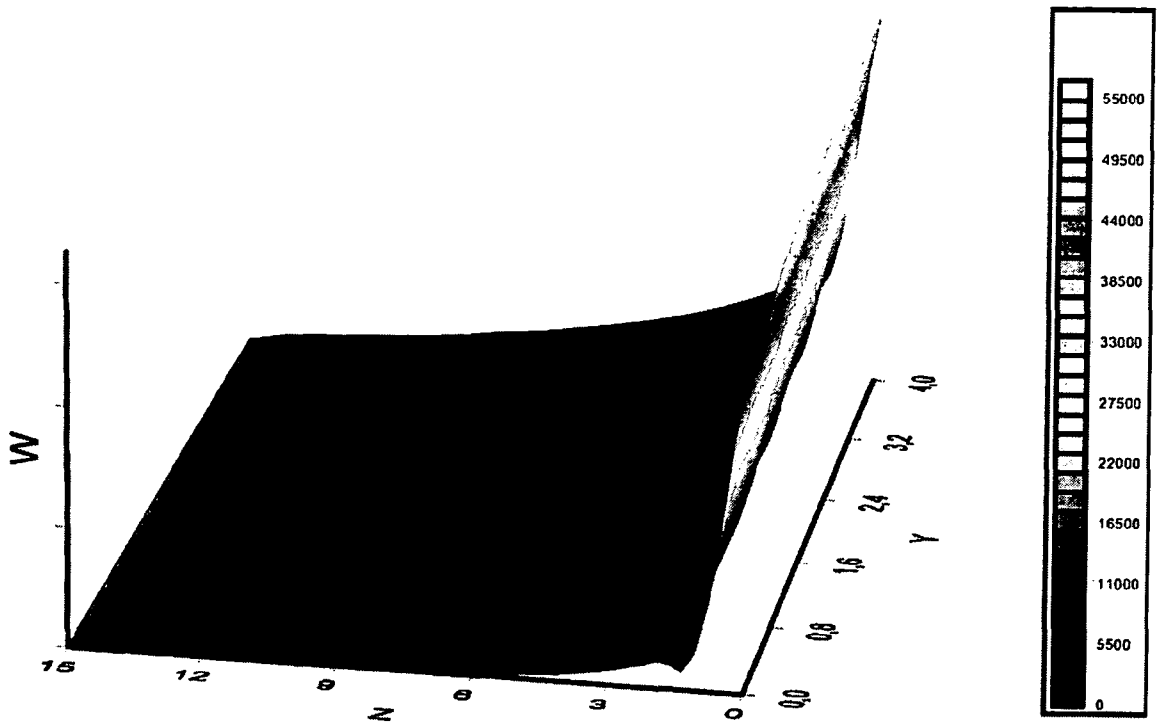

W
15
12
9
6
3
0
Z
0.0
0.8
1.6
2.4
3.2
4.0
Y
55000
49500
44000
38500
33000
27500
22000
16500
11000
5500
0

Figure 2

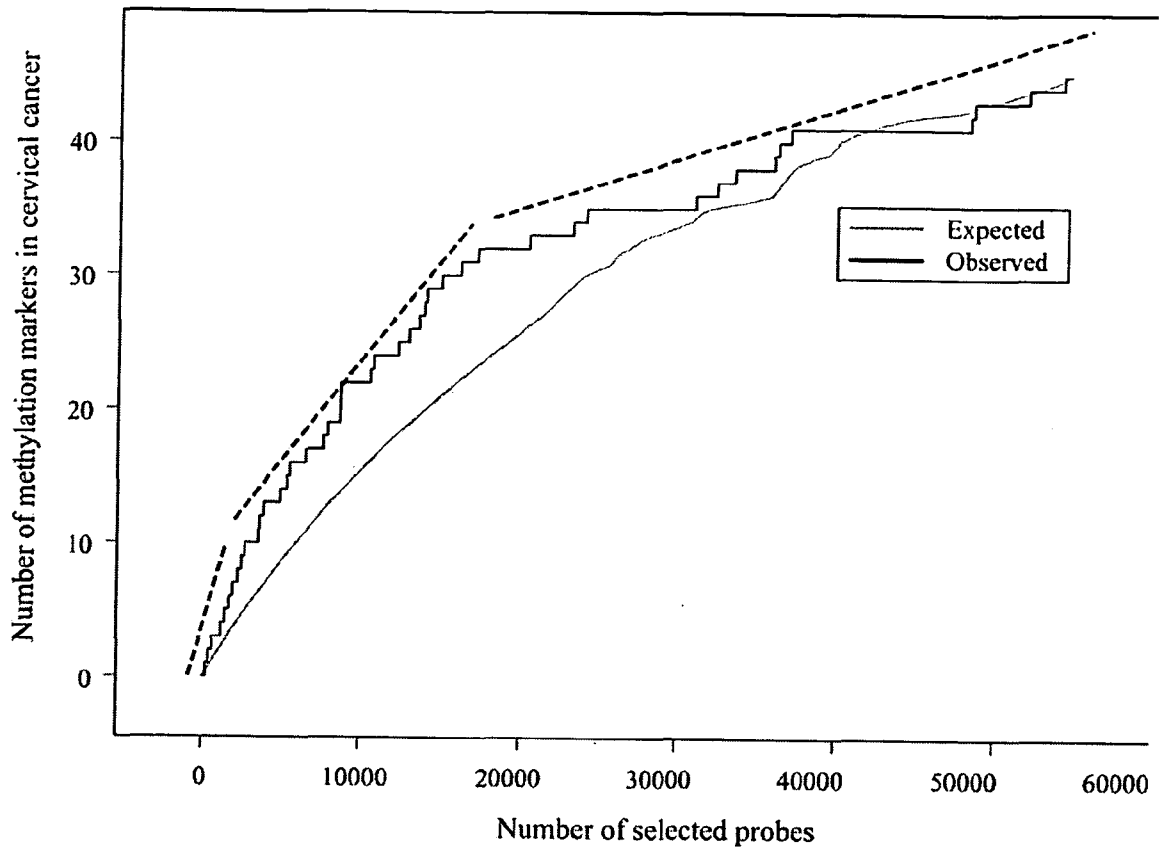

Number of methylation markers in cervical cancer
0
10
20
30
40
0
10000
20000
30000
40000
50000
60000
Number of selected probes
Expected
Observed

Figure 3

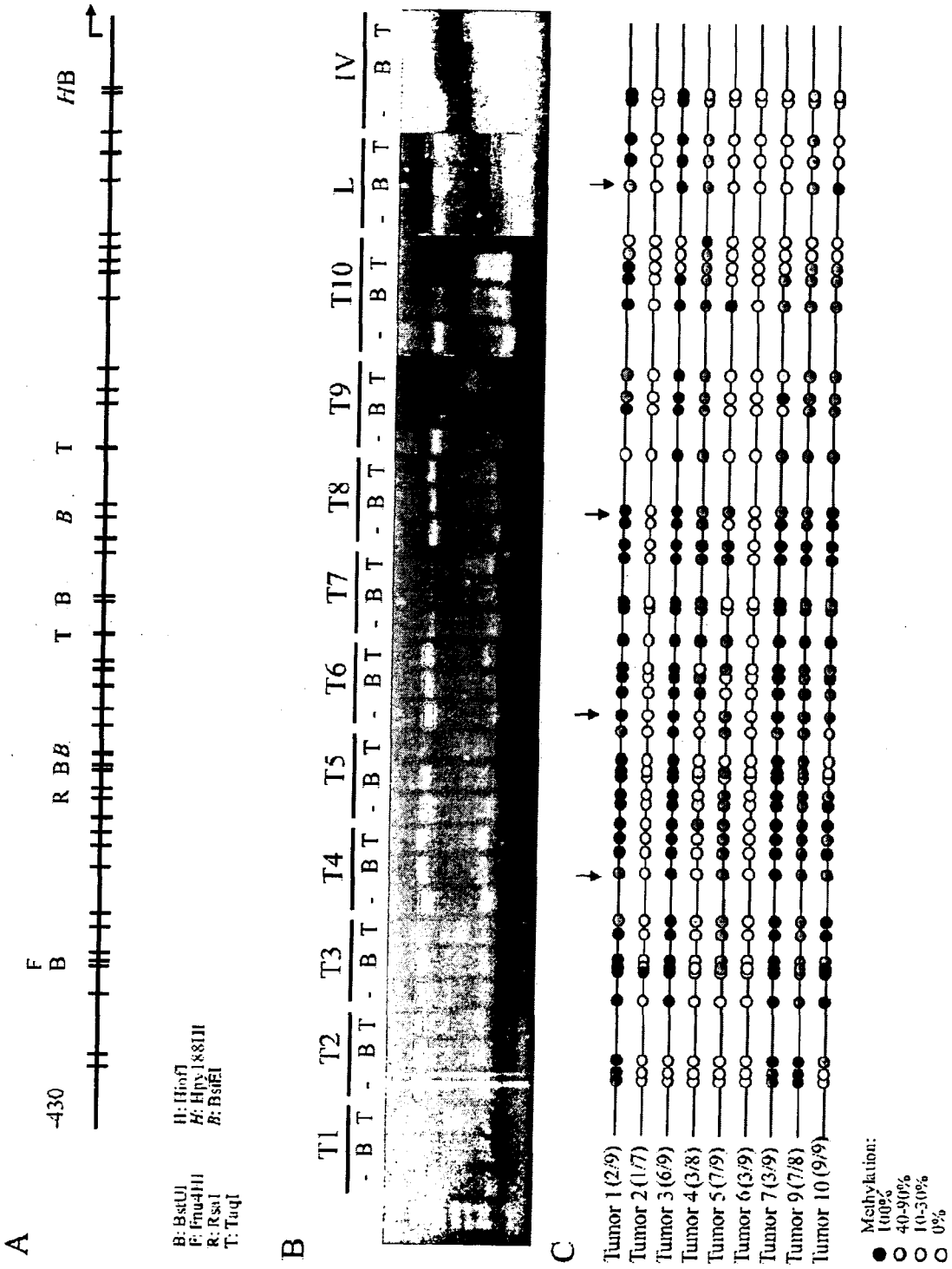

A
-430
F
B
R B B
T B
B
T
H B
L
H: HinfI
H: Hpy188III
B: BstUI
F: Fnu4HI
R: RsaI
T: TaqI
B
T1
T2
T3
T4
T5
T6
T7
T8
T9
T10
L
IV
C
Tumor 1 (2/9)
Tumor 2 (1/7)
Tumor 3 (6/9)
Tumor 4 (3/8)
Tumor 5 (7/9)
Tumor 6 (3/9)
Tumor 7 (3/9)
Tumor 9 (7/8)
Tumor 10 (9/9)
Methylation:
100%
40-90%
10-30%
0%

Figure 4

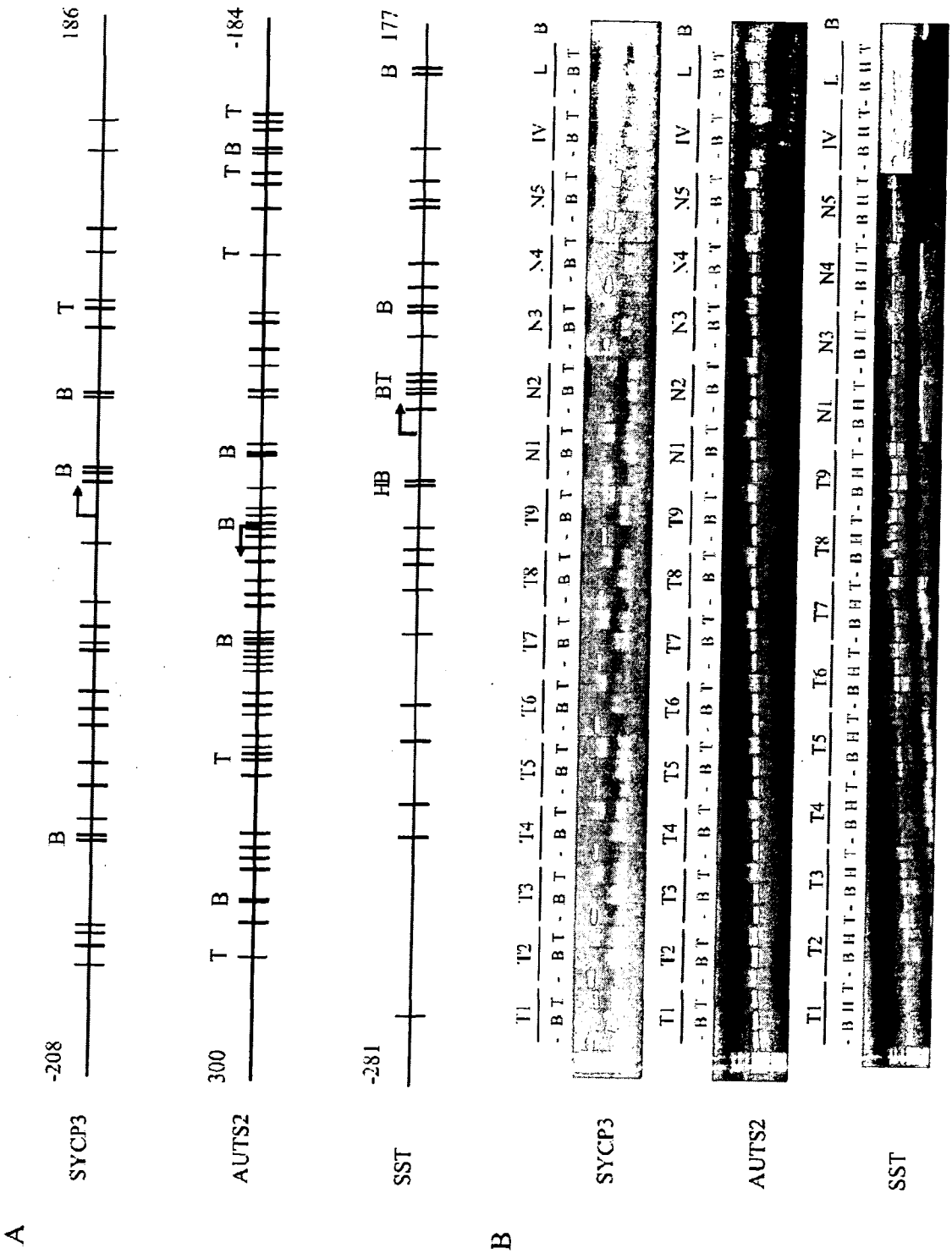

A
SYCP3
AUTS2
SST
-208
186
300
-184
-281
177
B
T1 T2 T3 T4 T5 T6 T7 T8 T9 N1 N2 N3 N4 N5 IV L
SYCP3
AUTS2
SST

Figure 5A

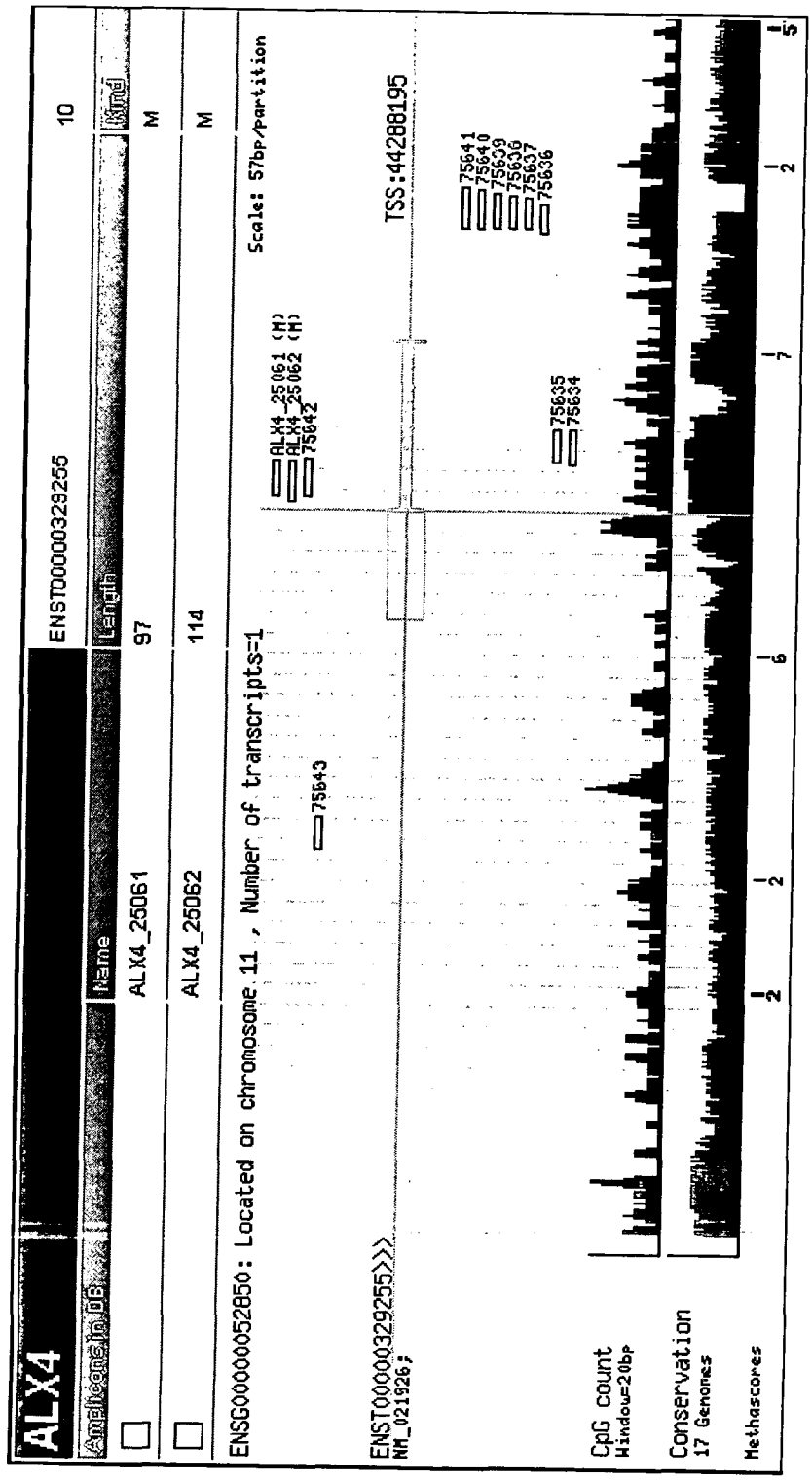

ALX4
ENST00000329255
10
Ameliome in DB
Name
Length
ALX4_25061
97
M
ALX4_25062
114
M
ENSG00000052850: Located on chromosome 11 , Number of transcripts=1
Scale: 57bp/partition
TSS:44288195
ENST00000329255>>
NM_021926;
CpG count
Window=20bp
Conservation
17 Genomes
Methascores

Figure 5B

Nr	Assay Name	Gene ID	Official Gene Symbol	Refseq	Sense Primer sequence (5' - 3') SEQ ID NO's 426-431	Antisense Primer sequence (5' - 3') SEQ ID NO's 432-437
1	ALX4_75634	60529	ALX4	NM_021926	CGTTAGGTTTTGTTGTTTAGCGT	CGACACAACTTTCCTATCGACC
2	ALX4_75635	60529	ALX4	NM_021926	CGTTAGGTTTTGTTGTTTAGCGT	AACTTTCCTATCGACCGCC
3	ALX4_75636	60529	ALX4	NM_021926	GTTTCGTAGTTTAGCGTTAGAGCGT	AAAACGAATACTTCTTACCGACC
4	ALX4_75637	60529	ALX4	NM_021926	GTTTCGTAGTTTCGTAGTTTCGTAG	TTCTTACCGACCCAAAACGTA
5	ALX4_75642	60529	ALX4	NM_021926	GGTCGTTGTTATGGACGTTT	AAAATTATACCGAACTTATCGCCT
6	ALX4_75643	60529	ALX4	NM_021926	TTCGTTATATTTTAGTTTAGCGTTT	ACACAAACCTTCGTCGTCC

Nr	Assay Name	Gene ID	Official Gene Symbol	Refseq	Amplicon Sequence (converted) (5'-3') SEQ ID NO's 438-443
1	ALX4_75634	60529	ALX4	NM_021926	CGTTAGGTTTTGTTGTTTAGCGTCGTAACGGGTTCGGTTTTTGGCGTTT TCGAATTTTTGTGTTTTGGCGGCGGTCGATAGGAAAGTTGTGTCG
2	ALX4_75635	60529	ALX4	NM_021926	CGTTAGGTTTTGTTGTTTAGCGTCGTAACGGGTTCGGTTTTTGGCGTTT TCGAATTTTTGTGTTTTGGCGGCGGTCGATAGGAAAGTT
3	ALX4_75636	60529	ALX4	NM_021926	GTTTCGTAGTTTAGCGTTAGAGCGTTGCGCGGAGATTTTTTGTCGTCGT ACGTTTTGGGTCGGTAAGAAGTATTCGTTTT
4	ALX4_75637	60529	ALX4	NM_021926	GTTTCGTAGTTTCGTAGTTTCGTAGTTTAGCGTTAGAGCGTTGCGCGGA GATTTTTTGTCGTCGTACGTTTTGGGTCGGTAAGAA
5	ALX4_75642	60529	ALX4	NM_021926	GGTCGTTGTTATGGACGTTTATTATAGTTCGGTGTCGTAGAGTCGGGAG GGTTCGTCGTTTTTTAGGGTATTTTTCGGAGGCGATAAGTTCGGTATAA TTTT
6	ALX4_75643	60529	ALX4	NM_021926	TTCGTTATATTTTAGTTTAGCGTTTTTATAGTAGAGGGAAATAGTTAATAA GACGTGTAAGTGATTATGTATTGGACGACGAAGGTTTGTGT

Figure 5B (cont.)

Nr	Assay Name	Gene ID	Official Gene Symbol	Refseq	Amplicon Sequence (not-converted) (5' - 3') SEQ ID NO's 444-449
1	ALX4_75634	60529	ALX4	NM_021926	CGCCAGGTCCTGCTGCCCAGCGCCGTAACGGGCCCGGCTCTTGGCGT CCCCGAATCCCTGTGCTTTGGCGGCGGCCGACAGGAAAGTTGTGCCG
2	ALX4_75635	60529	ALX4	NM_021926	CGCCAGGTCCTGCTGCCCAGCGCCGTAACGGGCCCGGCTCTTGGCGT CCCCGAATCCCTGTGCTTTGGCGGCGGCCGACAGGAAAGTT
3	ALX4_75636	60529	ALX4	NM_021926	GCCCCGCAGCCCAGCGCCAGAGCGCTGCGCGGAGACTCCTTGCCGC CGCACGCCCTGGGCCGGTAAGAAGCATCCGCTTC
4	ALX4_75637	60529	ALX4	NM_021926	GCCCCGCAGCCCCGCAGCCCCGCAGCCCAGCGCCAGAGCGCTGCGC GGAGACTCCTTGCCGCCGCACGCCCTGGGCCGGTAAGAA
5	ALX4_75642	60529	ALX4	NM_021926	GGCCGCTGCCATGGACGCCTACTACAGCCCGGTGTCGCAGAGTCGGG AGGGCTCGTCGCCTTTTAGGGCATTTCCCGGAGGCGACAAGTTCGGCA CAACTTT
6	ALX4_75643	60529	ALX4	NM_021926	TCCGCTACATCCCAGCCCAGCGCCCTTACAGCAGAGGGAAATAGTTAA CAAGACGTGCAAGTGACCATGCACTGGACGACGAAGGCTTGTGT

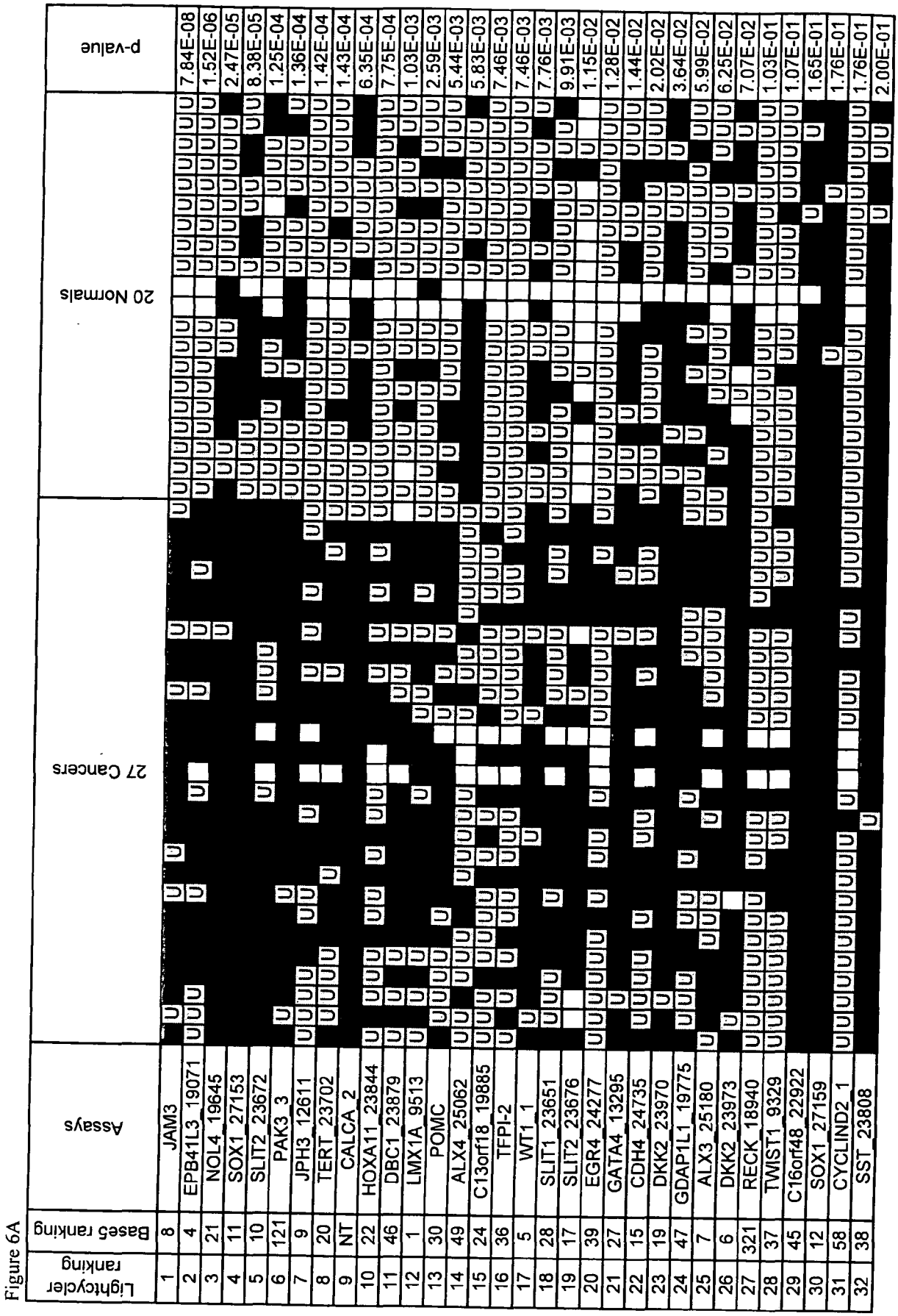

Lightcycler ranking
Based ranking
Assays
27 Cancers
20 Normals
p-value
1 8 JAM3 7.84E-08
2 4 EPB41L3_19071 1.52E-06
3 21 NOL4_19645 2.47E-05
4 11 SOX1_27153 8.38E-05
5 10 SLIT2_23672 1.25E-04
6 121 PAK3_3 1.36E-04
7 9 JPH3_12611 1.42E-04
8 20 TERT_23702 1.43E-04
9 NT CALCA_2 6.35E-04
10 22 HOXA11_23844 7.75E-04
11 46 DBC1_23879 1.03E-03
12 1 LMX1A_9513 2.59E-03
13 30 POMC 5.44E-03
14 49 ALX4_25062 5.83E-03
15 24 C13orf18_19885 7.46E-03
16 36 TFPI-2 7.46E-03
17 5 WT1_1 7.76E-03
18 28 SLIT1_23651 9.91E-03
19 17 SLIT2_23676 1.15E-02
20 39 EGR4_24277 1.28E-02
21 27 GATA4_13295 1.44E-02
22 15 CDH4_24735 2.02E-02
23 19 DKK2_23970 3.64E-02
24 47 GDAP1L1_19775 5.99E-02
25 7 ALX3_25180 6.25E-02
26 6 DKK2_23973 7.07E-02
27 321 RECK_18940 1.03E-01
28 37 TWIST1_9329 1.07E-01
29 45 C16orf48_22922 1.65E-01
30 12 SOX1_27159 1.76E-01
31 58 CYCLIND2_1 1.76E-01
32 38 SST_23808 2.00E-01

Figure 6A

Figure 6B

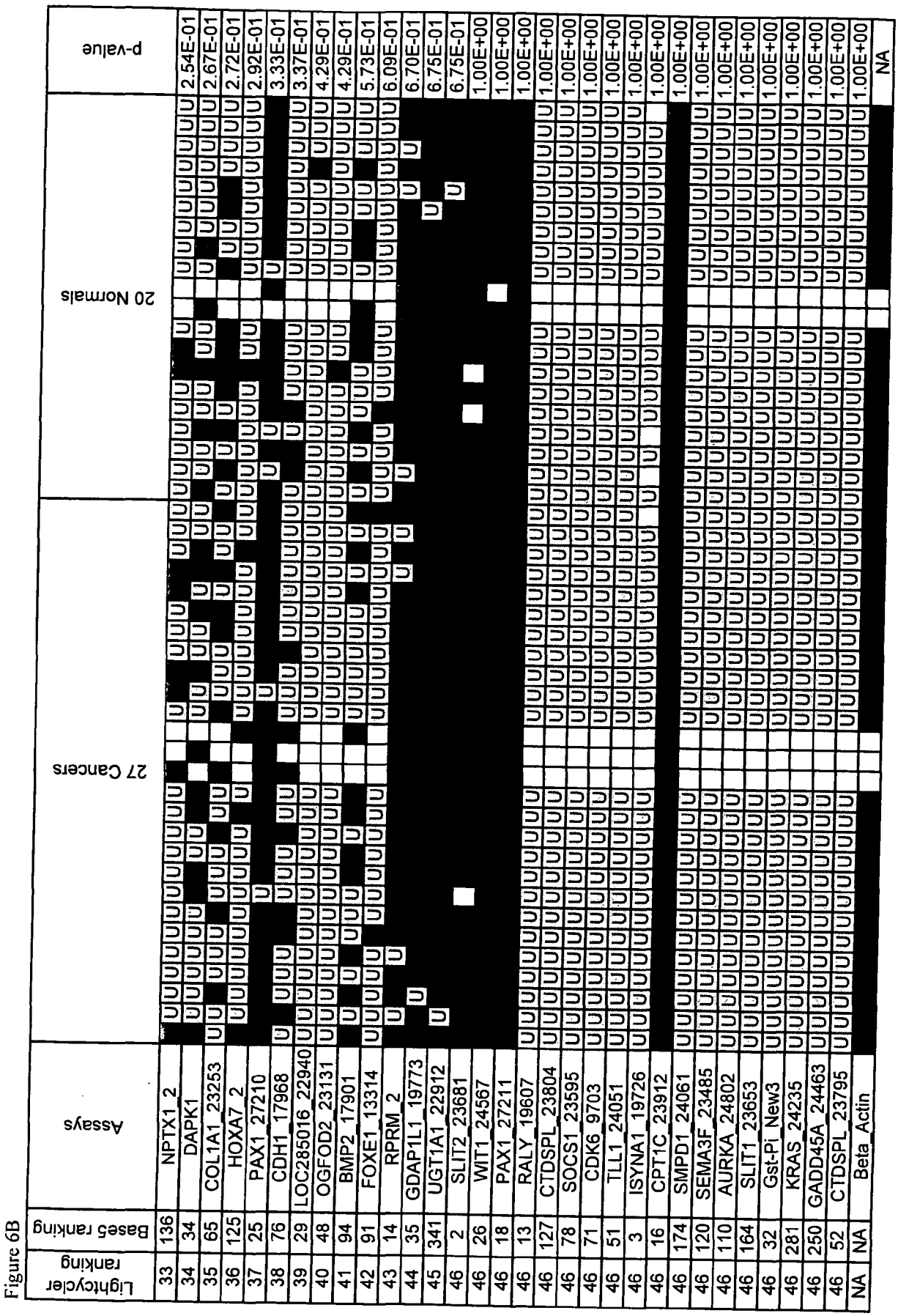

Lightcycler ranking
Based ranking
Assays
27 Cancers
20 Normals
p-value
33 136 NPTX1_2 2.54E-01
34 34 DAPK1 2.67E-01
35 65 COL1A1_23253 2.72E-01
36 125 HOXA7_2 2.92E-01
37 25 PAX1_27210 3.33E-01
38 76 CDH1_17968 3.37E-01
39 29 LOC285016_22940 4.29E-01
40 48 OGFOD2_23131 4.29E-01
41 94 BMP2_17901 5.73E-01
42 91 FOXE1_13314 6.09E-01
43 14 RPRM_2 6.70E-01
44 35 GDAP1L1_19773 6.75E-01
45 341 UGT1A1_22912 6.75E-01
46 2 SLIT2_23681 1.00E+00
46 26 WIT1_24567 1.00E+00
46 18 PAX1_27211 1.00E+00
46 13 RALY_19607 1.00E+00
46 127 CTDSPL_23804 1.00E+00
46 78 SOCS1_23595 1.00E+00
46 71 CDK6_9703 1.00E+00
46 51 TLL1_24051 1.00E+00
46 3 ISYNA1_19726 1.00E+00
46 16 CPT1C_23912 1.00E+00
46 174 SMPD1_24061 1.00E+00
46 120 SEMA3F_23485 1.00E+00
46 110 AURKA_24802 1.00E+00
46 164 SLIT1_23653 1.00E+00
46 32 Gst-Pi_New3 1.00E+00
46 281 KRAS_24235 1.00E+00
46 250 GADD45A_24463 1.00E+00
46 52 CTDSPL_23795 1.00E+00
NA NA Beta_Actin 1.00E+00 NA

Figure 7

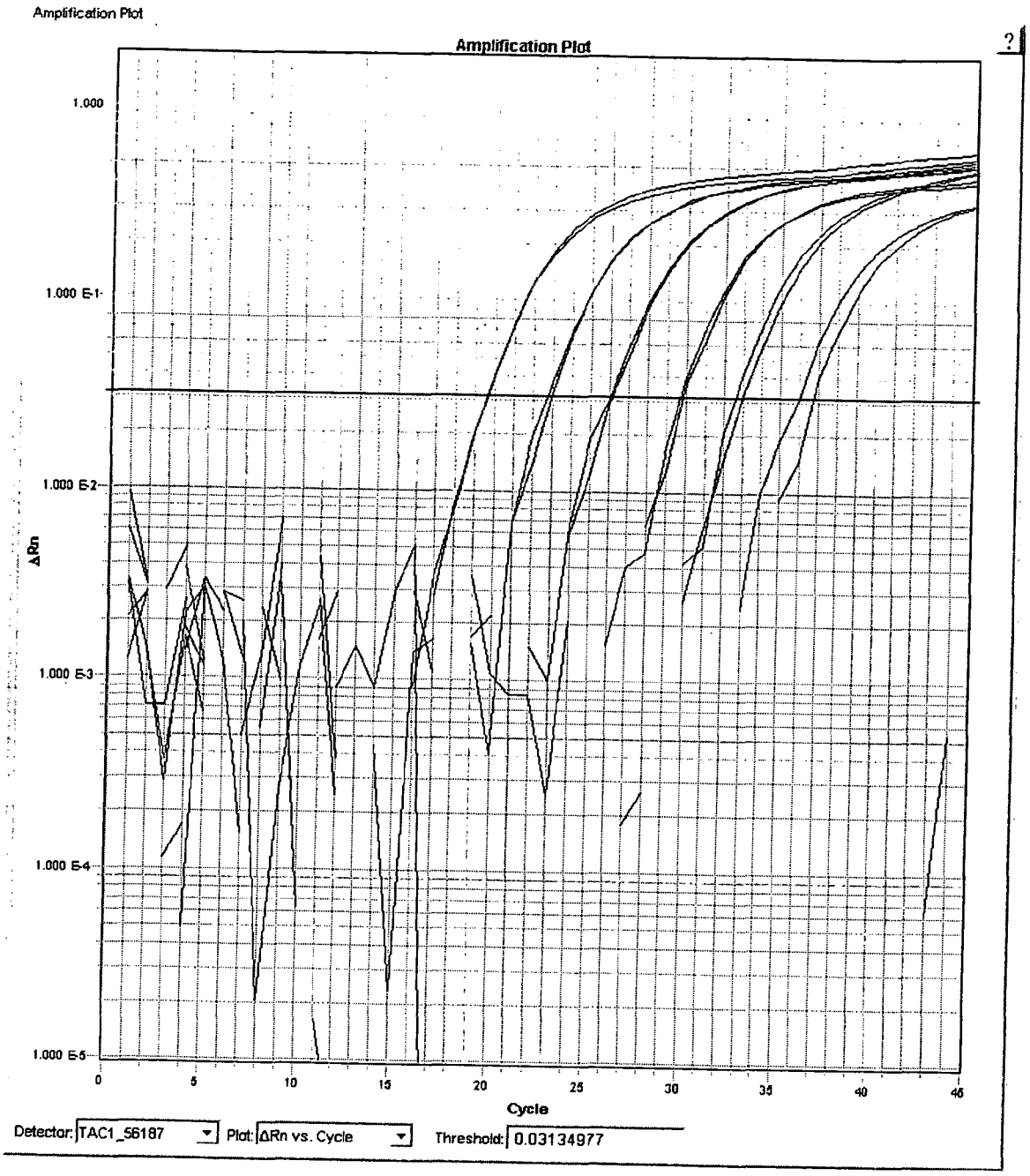

Amplification Plot
Amplification Plot
1.000
1.000 E-1
1.000 E-2
1.000 E-3
1.000 E-4
1.000 E-5
ΔRn
0
5
10
15
20
25
30
35
40
45
Cycle
Detector: TAC1_56187
Plot: ΔRn vs. Cycle
Threshold: 0.03134977

Figure 8

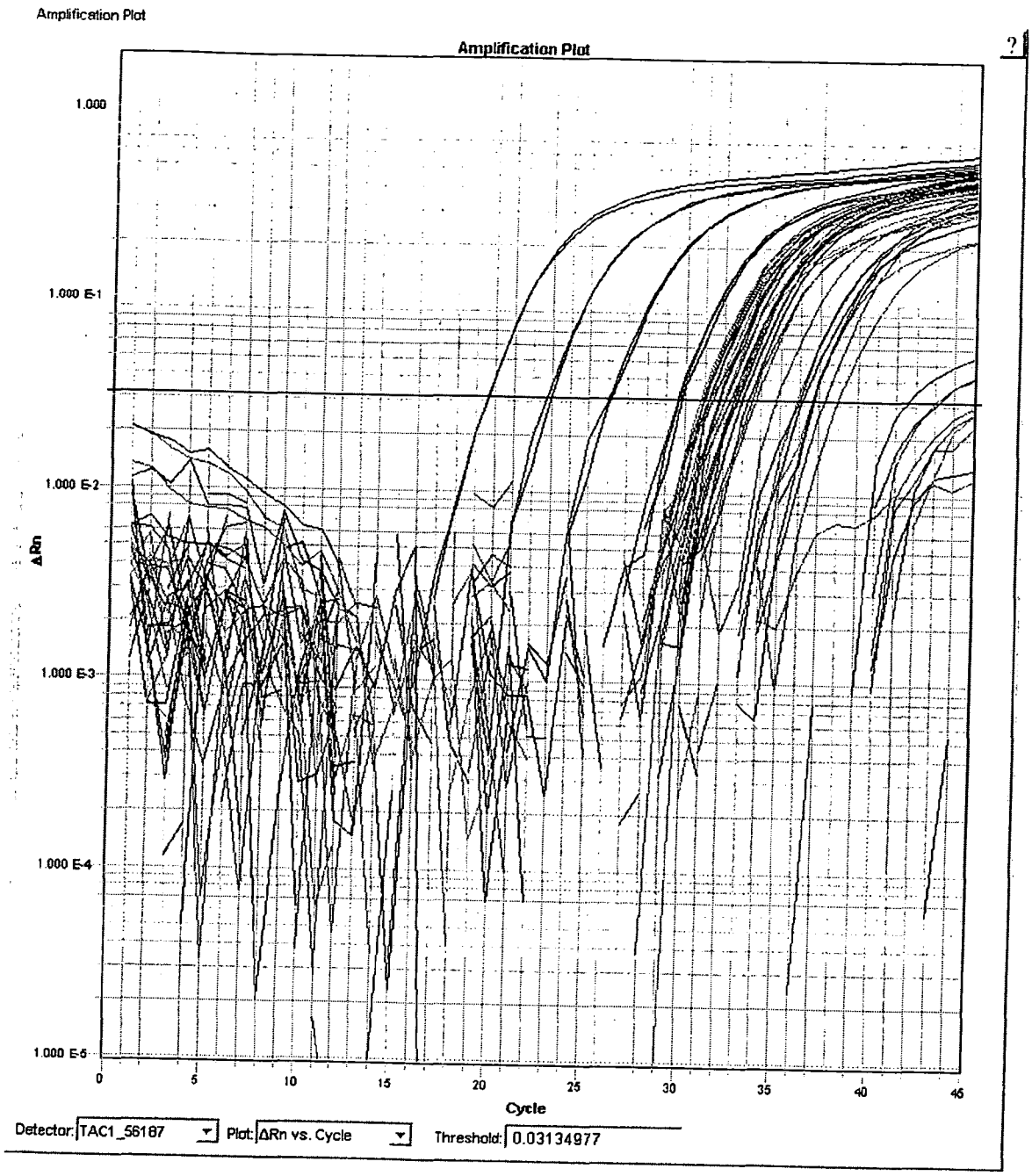

Amplification Plot
Amplification Plot
1.000
1.000 E-1
1.000 E-2
1.000 E-3
1.000 E-4
1.000 E-5
ΔRn
0
5
10
15
20
25
30
35
40
45
Cycle
Detector: TAC1_56187
Plot: ΔRn vs. Cycle
Threshold: 0.03134977

Figure 9

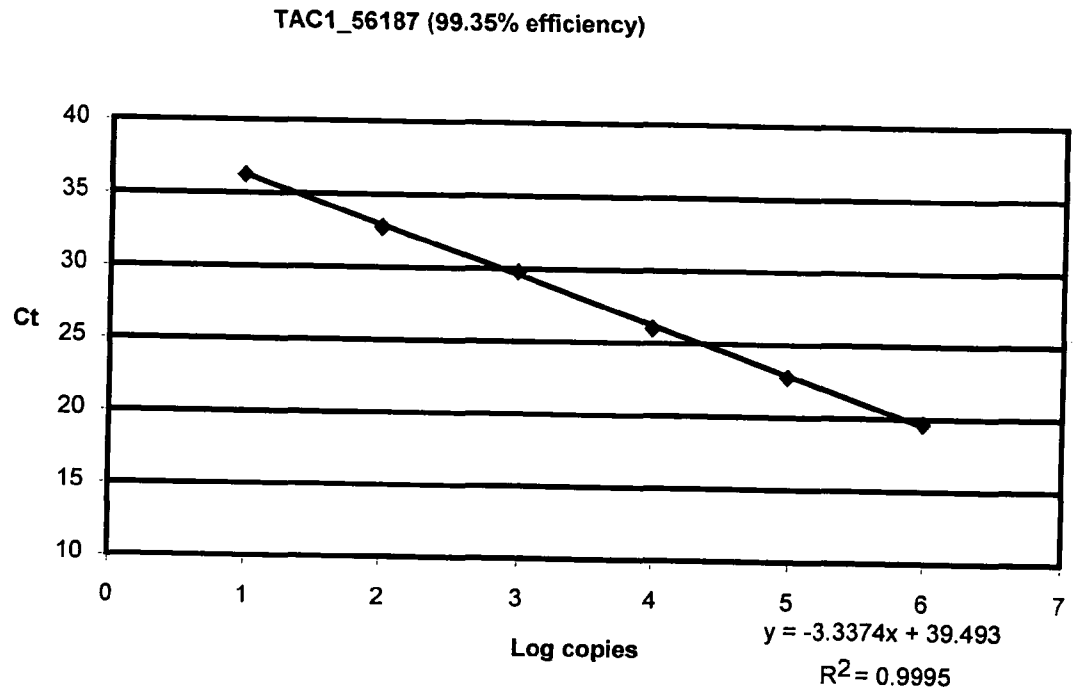

TAC1_56187 (99.35% efficiency)
Ct
40
35
30
25
20
15
10
0
1
2
3
4
5
6
7
Log copies
y = -3.3374x + 39.493
R2 = 0.9995

Figure 10

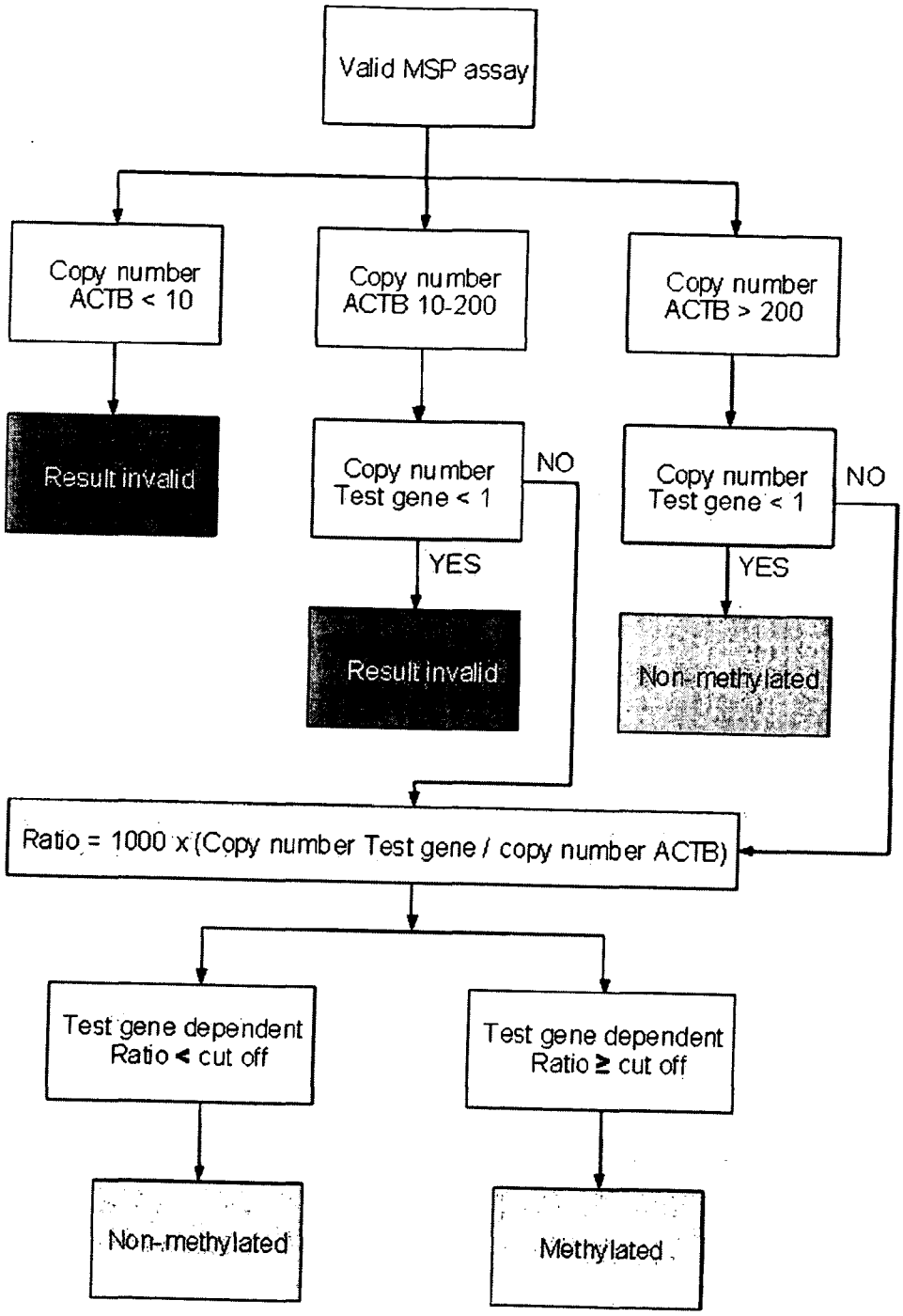
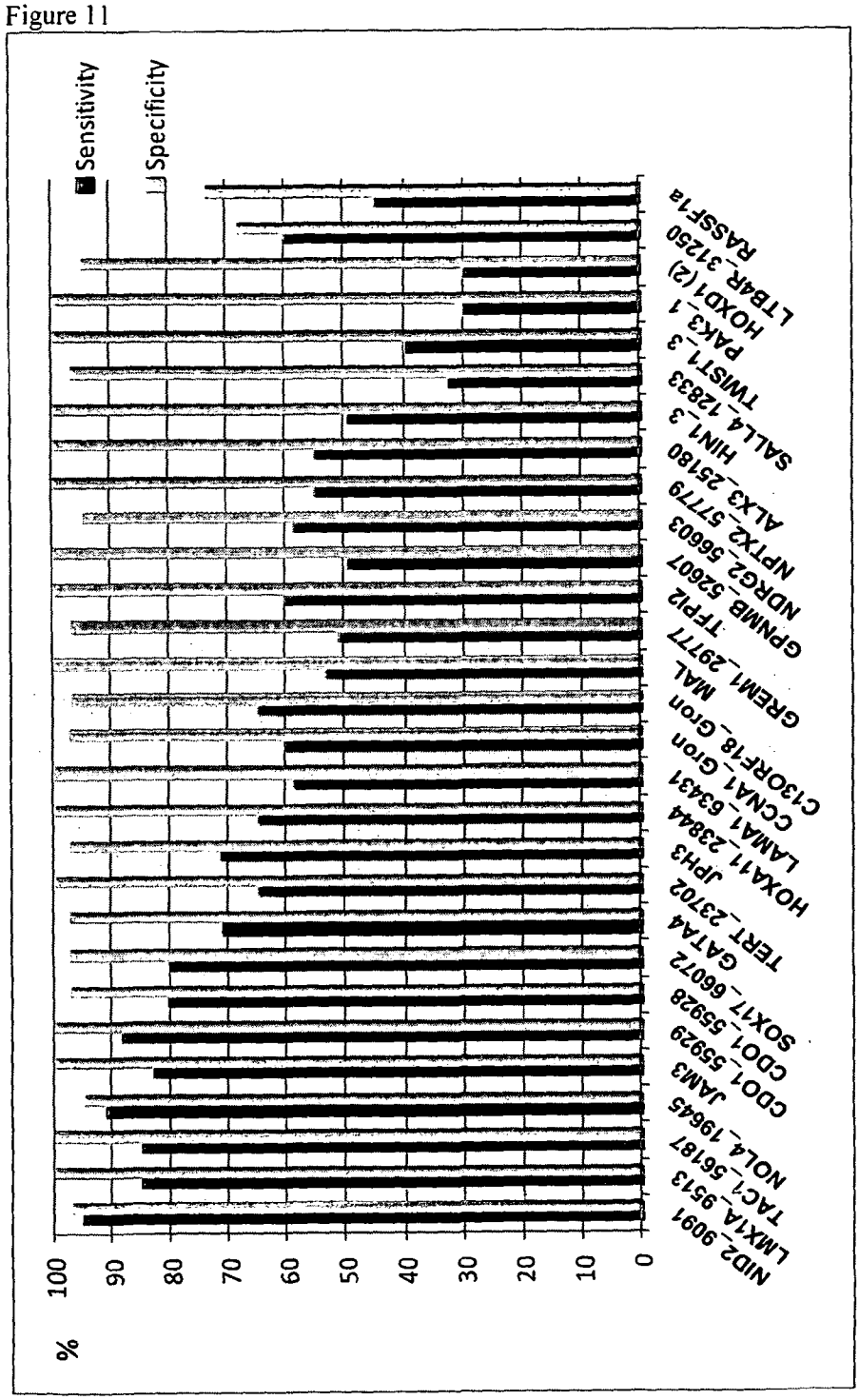

Valid MSP assay
Copy number ACTB < 10
Copy number ACTB 10-200
Copy number ACTB > 200
Result invalid
Copy number Test gene < 1
NO
YES
Result invalid
Copy number Test gene < 1
NO
YES
Non-methylated
Ratio = 1000 x (Copy number Test gene / copy number ACTB)
Test gene dependent Ratio < cut off
Test gene dependent Ratio ≥ cut off
Non-methylated
Methylated

Figure 11



SUBSTITUTE SHEET (RULE 26)