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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 May 2013 (10.05.2013)

WIPO | PCT

(10) International Publication Number
WO 2013/064163 A1

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

(21) International Application Number:
PCT/EP2011/005555

(22) International Filing Date:
1 November 2011 (01.11.2011)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant *(for all designated States except US)*: **ACADEMISCH MEDISCH CENTRUM** [NL/NL]; Meibergdreef 9, NL-1105 AZ Amsterdam (NL).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **MEDEMA, Jan Paul** [NL/NL]; Goudensteinpark 12, NL-2151 HC Nieuw Vennep (NL). **VERMEULEN, Louis** [NL/NL]; Wijsmullerstraat 49-11, NL-1058 JH Amsterdam (NL). **DE SOUSA E MELO, Felipe** [CH/NL]; Fagestraat 38-3, NL-1052 GD (NL).

(74) Agent: **DE PAUW, Elmar**; Bureau Kennistransfer AMC, Meibergdreef 9, TKs0-136 (De Rode Luifel), NL - 1105 AZ Amsterdam (NL).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, 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, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

Declarations under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

Published:

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

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

WO 2013/064163 A1

(54) Title: METHYLATION MARKERS FOR COLORECTAL CANCER

(57) Abstract: The present invention provides a method of determining the risk of developing metastasis and/or tumor relapse in a colorectal cancer patient, comprising the steps of providing a biological sample of said colorectal cancer patient, and determining in said biological sample the methylation status and/or the expression level of a gene selected from the group consisting of LGR5, AP-CDD1, AXIN2, DKK1, and ASCL2, and determining an increased risk of developing metastasis and/or tumor relapse, under the condition that hypermethylation and/or under expression is determined in a gene selected from the group consisting of LGR5, AP-CDD1, AXΓN2, DKK1, and ASCL2.

METHYLATION MARKERS FOR COLORECTAL CANCER

TECHNICAL FIELD OF THE INVENTION

The present invention relates in general to the use of the methylation status and or the expression level of particular genes and their potential to predict the recurrence of colorectal cancer or predict drug resistance/susceptibility. More specifically, the invention relates to the methylation status and/or expression level of a gene selected from the group consisting of LGR5, APCDD1, AXIN2 and ASCL2 and DKK1 as a biomarker for prognosis and treatment of colorectal cancer.

BACKGROUND OF THE INVENTION

Colorectal cancer is a major contributor to cancer-related death. Colorectal cancers are surgically removed in the majority of the cases. However, despite the fact that operations are declared to be successful and the patient is considered tumor-free by pathological analysis, recurrences occur in a subset of the patients. This is frequent in stage III patients, but even patients with a relatively low stage of disease at the time of operation (stage I and II) have a (small) risk of developing recurrences later on. Due to the fact that the risk is relatively small (<20%) patients are not receiving adjuvant treatment to limit the side-effects and to prevent overtreatment of patients that will never develop a recurrence. Identification of the patients at risk is therefore an important issue as these patients would clearly benefit from adjuvant therapy. There is therefore an urgent need for a test that helps to identify these patients.

Several biomarkers have been identified that are implicated in the prognosis of colorectal cancer. Specific mutations, microsatellite instability and epigenetic changes are being used to evaluate prognosis (Genetic and Epigenetic Biomarkers of Colorectal Cancer.Choong MK, Tsafnat G.Clin Gastroenterol Hepatol. 2011 Apr 29. More recently, it has been suggested that a higher prevalence of cancer stem cells is predictive for prognosis. Also immature stem cell signatures (based on Sox2, Oct4 and Nanog targets) have been used previously to determine prognosis in different types of cancers (Ben-Porath et al., 2008). It has also been suggested that intestinal stem cell signatures can predict recurrence in colorectal cancer (Merlos-Suarez et al., 2011). However, none of the existing biomarkers have been shown to provide sufficient sensitivity and/or specificity to predict prognosis in all patients or to justify routine use in a clinical setting.

In view of the foregoing, there remains a need in the art for a reliable method for prognosticating metastasis and/or tumor relapse in colorectal cancer patients. It is an object of the present invention to provide a method of prognosticating metastasis and/or tumor relapse in a colorectal cancer patient that does not suffer from the disadvantages of the currently available

methods, and materials for use in the method. This and other objects and advantages, as well as inventive features, will become apparent from the detailed description provided herein.

SUMMARY OF THE INVENTION

The present invention provides a method of determining the risk of developing metastasis and/or tumor relapse in a colorectal cancer patient, comprising determining in a biological sample of said patient the methylation status and/or the expression level of a gene selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2, and determining whether one or more of said genes is hypermethylated or under expressed, wherein hypermethylation or under expression of one or more of said genes indicates an increased risk of developing metastasis and/or tumor relapse. In a preferred embodiment, hypermethylation or under expression of at least two, preferably three, four or five of said genes indicates an increased risk of developing metastasis and/or tumor relapse. More preferably, hypermethylation in at least two, preferably three, four or five genes of the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2 indicates an increased risk. In a highly preferred embodiment, an increased risk of developing metastasis and/or tumor relapse is determined, if hypermethylation and/or under expression is determined in the genes selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2; LGR5, APCDD1, AXIN2 and DKK1; LGR5, APCDD1, AXIN2, and ASCL2; and AXIN2 and ASCL2. Preferably, said colorectal cancer patient is suffering from colorectal cancer in stage I, II or III, more preferably stage II and III.

The invention further provides a method of predicting the response in a colorectal cancer patient to treatment with a demethylating agent comprising determining the risk of developing metastasis and/or tumor relapse in a colorectal cancer patient according to the method of the invention, and predicting a positive clinical response to said treatment, if an increased risk is determined. Preferably, said demethylating agent is 5-aza-cytidine, 5-aza-2'-deoxycytidine (decitabine), zebularine [1-(β-D-ribofuranosyl(dihydro-pyrimidin-2-1)] or L-methionine.

The invention further provides a method for identifying and/or selecting a colorectal cancer patient with colorectal cancer or treated for colorectal cancer suitable for treatment with a demethylating agent comprising determining the response in a colorectal cancer patient to treatment to colorectal cancer with a demethylating agent according to the method of the invention, and identifying and/or selecting the colorectal cancer patient for treatment with said demethylating agent if a positive clinical response is determined.

Further provided is a method for selecting a suitable treatment regimen for colorectal cancer in a colorectal cancer patient comprising identifying and/or selecting a colorectal cancer patient suitable for treatment with a demethylating agent according to the method of the invention, and selecting said demethylating agent for treatment if the colorectal cancer patient is identified and/or selected for treatment with said demethylating agent.

The invention further provides a method of treating a patient suffering from colorectal cancer with a demethylating agent comprising determining in a biological sample from said patient the methylation status and/or the expression level of one or more genes selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2, determining whether one or more of said genes is hypermethylated or under expressed, and administering to said patient a therapeutically effect amount of a demethylating agent if one or more of said genes is hypermethylated or under expressed.

The invention further provides a kit comprising one or more primers for determining the methylation status of at least one gene selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2, wherein said one or more primers is an oligonucleotide having 18 to 40 nucleotides and having at least 80%, preferably at least 95% sequence identity to a nucleic acid sequence selected from Seq ID No. 31, Seq ID No. 32, Seq ID No. 33, Seq ID No. 34, Seq ID No. 35, Seq ID No. 36, Seq ID No. 37, Seq ID No. 38, Seq ID No. 39, Seq ID No. 40, Seq ID No. 49, Seq ID No. 50, Seq ID No. 55, Seq ID No. 56, Seq ID No. 57 and Seq ID No. 58. Preferably, said kit comprises two or more primers for determining the methylation status of at least two genes selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2, wherein said two or more primers are oligonucleotides having 18 to 40 nucleotides and having at least 80%, preferably at least 95% sequence identity to a nucleic acid sequence selected from Seq ID No. 31, Seq ID No. 32, Seq ID No. 33, Seq ID No. 34, Seq ID No. 35, Seq ID No. 36, Seq ID No. 37, Seq ID No. 38, Seq ID No. 39, Seq ID No. 40, Seq ID No. 41, Seq ID No. 42, Seq ID No. 43, Seq ID No. 44, Seq ID No. 45, Seq ID No. 46, Seq ID No. 47, Seq ID No. 48, Seq ID No. 49, Seq ID No. 50, Seq ID No. 51, Seq ID No. 52, Seq ID No. 53, Seq ID No. 54, Seq ID No. 55 and Seq ID No. 56.

The invention further provides a kit for assessing methylation in a test sample, comprising in a package a reagent that modifies methylated cytosine residues but not non-methylated cytosine residues, or that modifies non-methylated cytosine residues but not methylated cytosine residues; and one or more oligonucleotide primers and/or pair of oligonucleotide primers that specifically hybridizes under amplification conditions to a gene selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2, preferably one or more primer according to the invention.

The invention further provides the use of the kit of the invention for determining the methylation status of one or more genes selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2.

The invention further provides the use of the kit of the invention for determining whether a sample is from an individual at risk of developing metastasis and/or tumor relapse. Further provided is the use of the kit of claim the invention for determining a risk of developing metastasis and/or tumor relapse in a colorectal cancer patient.

Further provided is a method for classifying a biological sample of a colorectal cancer patient, comprising determining in said sample the methylation status and/or the expression level of a gene selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 shows that the cancer stem cell profile predicts poor prognosis, and Wnt targets and intestinal stem cell markers predict a favorable prognosis. Figure A shows a schematic representation of cancer stem cell spheroid transduction with TOP-GFP lentiviral vector. Microarray analysis was performed on 10% lowest and highest TOP-GFP sorted fractions to generate a colon cancer stem cell signature. The heat map depicts the most differentially regulated genes, including differentiation markers (FABP1, MUC2) and Wnt canonical targets (LGR5, LEF1 and APCDD1). Figure B shows the results of a gene set enrichment analysis reveals a strong relationship between cancer stem cell signature and tumor relapse in the AMC-AJCCII-90 patient set. Also previously described LGR5- and intestinal stem cell-EphB2 signatures associate with recurrence in the present set. ES, enrichment score; NES, normalized enrichment score; FDR, false discovery rate. Figure C shows a Kaplan-Meier graph (relapse free survival) based on overall adherence to the TOP-GFPhigh/cancer stem cell profile as identified by gene ranking analysis (see example section for details). Figures D and F show the results of K-means clustering analysis of colorectal cancer samples from the AMC-AJCCII-90 patient set according to the colon-cancer stem cell signature (D) and the dnTCF4 signature representing a defined set of Wnt target genes (F). Several canonical Wnt target genes are denoted for each signature. Figures E and G show the results of a Kaplan-Meier analysis on relapse free survival which is drawn for each corresponding signature. Note that the poor prognosis cluster (black) is associated with low expression of indicated Wnt targets (See also Figures 5C and D for similar conclusions based on the LGR5- and intestinal stem cell-EphB2 signatures). P-values are calculated with the log-rank test. (D-G) Black represents Wnt-target-low cluster (WntLow), Grey Wnt-target-high cluster (WntHigh).

Figure 2 shows that Wnt target genes are regulated by methylation in colorectal cancer. Figure A shows the relative expression of established Wnt target genes (LGR5, APCDD1, ASCL2, DKK1 and AXIN2) in normal, adenoma, and colorectal cancer tissue. Patients in the colorectal cancer group that developed a recurrence are indicated with open symbols. Horizontal line indicates mean value. Figure B shows the results of methylation-specific PCR for Wnt target genes LGR5, APCDD1, ASCL2 and DKK1 in a panel of colorectal cancer lines. U, unmethylated; M, methylated; Figure C shows the relative expression of indicated Wnt target genes following (48hrs) demethylating treatment with 5-Aza of colorectal cancer cell lines. Confidence intervals represent standard deviation.

Figure 3 shows the functional relevance of Wnt target gene methylation. Figures A and B show the results of a clonogenic analysis of indicated colorectal cancer cell lines (A) and primary human colon-cancer stem cell cultures (B) in the absence or presence of demethylating treatment with 5-Aza for 48hrs. Figure C shows subcutaneous xenografts of human primary cancer stem cell cultures treated with 5-Aza or PBS intraperitoneally. Figure D shows expression of Wnt target genes in 5-Aza-treated primary human cancer stem cell culture-derived xenografts from the experiment depicted in (C). Figure E shows the TOP/FOP analysis to determine Wnt signaling levels following 5-Aza treatment. Figure F shows the effect of transient overexpression of APCDD1 or AXIN2 on TOP/FOP activity in colorectal cancer cell lines and in primary human cancer stem cell cultures (Co100, right bars). Error bars represent standard deviation.

Figure 4 shows that methylation of the Wnt target gene biomarkers identifies poor prognosis patients. Figure A shows the percentage of CpG island methylation in the promoter region of the indicated Wnt target genes for a subset of patients from the AMC-AJCCII-90 set. Figure B shows the unsupervised cluster analysis using the methylation levels (ranking) of the Wnt target gene biomarkers which reveals two clusters. Patient number, High or Low Wnt-target cluster and recurrence are indicated. Grey scale depicts rank order of methylation level within the patient set for each gene, green; lowly methylated, red; highly methylated. Grey squares represent data not available. Figure C shows a Kaplan-Meier curve depicting the two different patient groups as identified in (B). P-value is calculated with the log-rank test.

Figure 5 shows gene set enrichment analysis, K-means clustering and Wnt targets validation by qPCR. Figure A displays a heat map of unsupervised cluster analysis based on the 187 gene-colon cancer stem cell signature on 6 TOP-GFP high and low sorted spheroid cultures reveals clear separation of cancer stem cells from differentiated cells. Figure B shows the results of gene set enrichment analyses performed with the different (cancer) stem cell associated gene signatures on an additional, publically available patient dataset. ES, enrichment score; NES, normalized enrichment score; FDR, false discovery rate. Figures C and D show that K-means clustering of the AMC-AJCCII-90 with the LGR5- (C) or ISC-EphB2 (D) signatures also shows two clusters. Kaplan-Meier and prognostic value of these signatures is depicted (right panels). P-value is calculated with the log-rank test. (E) qPCR validation of Wnt target gene expression levels in the Wnt-target-High (WntHigh) and Wnt-target-Low (WntLow) clusters by qPCR. Each dot represents a patient. Horizontal lines represents the mean value.

Figure 6 shows that Wnt target genes are prognostic but their expression levels do not relate to the activity of the Wnt cascade or to additional mutations. Figure A shows the results of

multivariate analysis using Cox proportional hazard model to assess dependency between the cancer stem cell signature and several reported mutations in colorectal cancer in relation to prognosis. Figure B shows representative β-catenin stainings are shown for two patients in both the Wnt-target-High (WntHigh) and Wnt-target-Low (WntLow) clusters. Figure C shows automated scoring of nuclear β-catenin fractions in patients from the AMC-AJCCII-90 set. Figure D shows that Kaplan-Meier curves depict disease free survival for groups based on expression levels of the individual Wnt target genes indicated in the AMC-AJCCII-90 patient set. Groups comprise 45 highest (red) and 45 lowest (blue) expressing patients for each gene. P-value is calculated with the log-rank test. (E) Graphs depict relationship between fraction of CD133+ cells as determined by FACS analysis in freshly isolated colorectal cancer specimens and expression levels of the genes indicated by qPCR. Figure F shows graphs which depict a relation between fraction of cells demonstrating nuclear β-catenin and expression levels of the genes indicated from the AMC-AJCCII-90 patient set. Both in (E) and (F) no clear relation can be observed between colon stem cell content and colon stem cell-associated genes in these samples. Each dot represents an independent colorectal cancer sample.

Figure 7 shows that poor prognosis colorectal cancer patients associate with an immature phenotype. Figure A demonstrates the results of gene set enrichment analysis for the association of embryonic stem cells associated gene signatures with either the Wnt-target-Low or Wnt-target-High cluster of patients. Results of both the AMC-AJCCII-90 patient set as well as a publically available, and larger, patient set are included. Values indicate an association of the particular signature with the Wnt-target-Low cluster. Size indicates number of genes in the profile. NOS indicates Nanog/Oct4/Sox2. ES, enrichment score; NES, normalized enrichment score; FDR, false discovery rate. FDRs < 0.25 are indicated in bold. Clearly the Wnt-target-low cluster of tumors in both patients sets adhere to a more immature phenotype. (B) qPCR validation of Embryonic stem cell core genes expression levels in the Wnt-target-High (WntHigh) and Wnt-target-Low (WntLow) clusters by qPCR. Each dot represents a patient. Horizontal lines represents the mean value. Figure C shows that the Wnt-target-Low (WntLow) cluster of patients is enriched in patients displaying a poorly differentiated morphology. P-value calculated by Chi-square test. Figure D shows Wnt target genes that are downregulated during the progression from adenoma to carcinoma were manually deleted from the colon stem cell-signature (n=16 genes). This curated signature was subsequently used in a gene set enrichment analysis showing increased association with poor prognosis colorectal cancers. Figure E shows the microarray expression levels of the different Wnt target genes in normal, adenoma and colorectal cancer samples derived from a publically available dataset (Galamb et al., 2008) depicting identical up and downregulation during tumor progression. Figure F shows that heat map analysis of this dataset using the dnTCF4 gene signature reveals high and low Wnt target gene expression groups. Of note some of the colorectal cancers group with low/normal samples and some with the high/adenoma samples.

DETAILED DESCRIPTION OF THE INVENTION

Definitions

The term "colorectal cancer patient" as used herein refers to a patient (also referred to as subject) suffering from colon or rectal cancer, a subject suspected of having colorectal cancer, identified as having an increased risk for colorectal cancer, treated for colorectal cancer, in remission for colorectal cancer, being monitored for recurrence of colorectal cancer or a subject who has been treated for colorectal cancer. The patient could already be diagnosed as having colorectal cancer, and provided methods may be useful in determining the stage of disease progression. In a preferred embodiment of any method according to the invention, said patient is suffering from colorectal cancer in stage 0, I, II, III or IV and even more preferably stage II, III. In a highly preferred embodiment, said patient has been treated for colorectal cancer, preferably not longer than 10, 9, 8, 7, 6, 5 years ago. Most preferably, said patient has been treated for colorectal cancer stage I or II. The colorectal patient may be an animal or human being. In preferred embodiments of the invention, the patient is a human being.

The term "suffering from" when used to describe a patient and in reference to a disease, is used herein to describe patients who have been diagnosed as having the disease, and/or are experiencing symptoms related to the disease. Thus, a patient who is diagnosed with colorectal cancer but not experiencing symptoms related to the cancer is "suffering from" colorectal cancer.

The term "colorectal cancer" as used herein, also called colon cancer or bowel cancer, is defined to include cancerous growths in the colon, rectum and appendix.

The term "susceptible" is used herein to mean having an increased risk for and/or a propensity for something, i.e. a disease such as cancer. The term takes into account that an individual "susceptible" for a disease may never be diagnosed with the disease.

As used herein, the terms "stage I cancer," "stage II cancer," "stage III cancer," and "stage IV" refer to the TNM staging classification for cancer. Stage I cancer typically identifies that the primary tumor is limited to the organ of origin. Stage II intends that the tumor has spread through the muscle wall of the colon. Stage III intends that the tumor has spread to lymph nodes. Stage IV intends that the primary tumor has spread to other organs.

The term "tumor relapse" as used herein refers to colorectal cancer that comes back again after treatment, either in the colon, or in some other part of the body.

The term "biological sample" as used herein refers to any sample from a colorectal cancer patient for diagnostic, prognostic, or personalized medicinal uses and may be obtained from surgical samples, such as biopsies or fine needle aspirates, from paraffin-embedded tissues, from frozen tumor tissue samples, from fresh tumor tissue samples, from a fresh or frozen body fluid. Most preferably the sample contains cells derived from colon or colorectal tissue or nucleic acids from such cells. However, any other suitable biological samples (e.g. bodily fluids such as blood, stool,

etc...) in which the methylation status or the gene expression level of a gene of interest can be determined are included within the scope of the invention. In certain embodiments of the invention, the biological sample is processed or treated before being used in accordance with the inventive methods. For example, the sample may be processed such that the sample contains mostly nucleic acids such as DNA.

By "methylation status" is meant the level of methylation of cytosine residues (found in CpG pairs) in the gene of interest. When used in reference to a CpG site, the methylation status may be methylated or unmethylated. When used in reference to a CpG island or to any stretch of residues, the methylation status refers to the level of methylation, which is the relative or absolute concentration of methylated C at the particular CpG island or stretch of residues in a biological sample. Methylation of a CpG site or island at a promoter usually prevents expression of the gene. The sites or islands can also surround the 5' region of the coding region of the gene as well as the 3' region of the coding region. Thus, CpG sites or islands can be found in multiple regions of a nucleic acid sequence including 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, as appropriate. The levels of methylation of the gene of interest are determined by any suitable means in order to reflect whether the gene is likely to be downregulated or not.

The term "expression level" as used herein refers to a determined level of gene expression.

The term "hypermethylation" as used herein refers to the average methylation status corresponding to an increased presence of 5-mCyt at one or a plurality of CpG dinucleotides within a DNA sequence of a test DNA sample, relative to the amount of 5-mCyt found at corresponding CpG dinucleotides within a normal control DNA sample.

The term "under expression" as used herein refers to a lesser expression level of a gene in a biological sample, compared to expression in a control or reference sample. In preferred embodiments, under expression is defined as a decrease in expression level of at least twofold compared to the expression level value in a reference sample.

The phrase "corresponding to" when used to describe positions or sites within nucleotide sequences, is used herein as it is understood in the art. As is well known in the art, two or more nucleotide sequences can be aligned using standard bioinformatic tools, including programs such as BLAST, ClustalX, Sequencher, and etc. Even though the two or more sequences may not match exactly and/or do not have the same length, an alignment of the sequences can still be performed and, if desirable, a "consensus" sequence generated. Indeed, programs and algorithms used for alignments typically tolerate definable levels of differences, including insertions, deletions, inversions, polymorphisms, point mutations, etc. Such alignments can aid in the determination of which positions in one nucleotide sequence correspond to which positions in other nucleotide sequences.

The abbreviation "CpG" is used herein to refer to a dinucleotide comprised of a cytosine nucleotide (deoxycytidine) linked via a phosphate group to a guanine nucleotide (deoxyguanosine) through linkages to the 5' position of the deoxycytidine and the 3' position of the deoxyguanosine. The cytosine in this dinucleotide is said to be in the "5' position" of the dinucleotide, and the guanine is said to be in the "3' position" of the dinucleotide. As is understood by one of ordinary skill in the art, the abbreviation "CpG" also refers to modified dinucleotides similar, so long as the 5' nucleotide is still identifiable as deoxycytidine and the 3' nucleotide is still identifiable as deoxyguanosine. For example, a deoxycytidine-deoxyguanosine dinucleotide in which the cytosine ring is methylated at the 5 position is still considered a CpG dinucleotide, and may be referred to as a methylated CpG or abbreviated as 5mCpG. As used herein, the abbreviation CpG can also refer to a CpG site, defined below.

The term "CpG island" refers to a contiguous region of genomic DNA that satisfies the criteria of (1) having a frequency of CpG dinucleotides corresponding to an "Observed/Expected Ratio" >0.6, and (2) having a "GC Content" >0.5. CpG islands are typically, but not always, between about 0.2 to about 1 kb in length.

The term "Observed/Expected Ratio" ("O/E Ratio") refers to the frequency of CpG dinucleotides within a particular DNA sequence, and corresponds to the [number of CpG sites/(number of C bases X number of G bases)] X band length for each fragment.

The term "CpG site" is used herein to refer to a position within a region of DNA corresponding to a position where a CpG dinucleotide is found in a reference sequence. One of ordinary skill in the art will understand the term CpG site to encompass the location in the region of DNA where a CpG dinucleotide is typically found, whether or not the dinucleotide at that position is a CpG dinucleotide in a particular DNA molecule. For examples, the DNA sequence of a gene may typically contain a CpG dinucleotide at a particular position, but may contain other dinucleotides at the corresponding position in mutant versions, polymorphic variants, or other variations of the gene. Some mutations such as single base substitutions may alter the identity of the dinucleotide to be something other than CpG. Other mutations such as insertions or deletions may alter the position of the CpG dinucleotide typically found at a particular site. In cases such as these, the term CpG site is understood by one of ordinary skill in the art to encompass the site corresponding to the position where a CpG dinucleotide is typically found, for example, in a wild type version of the DNA. Similarly, the term CpG site also encompasses the corresponding site in a nucleic acid that has been modified experimentally, for example by labeling, methylation, demethylation, deamination (including conversion of a cytosine to uracil by a chemical such as sodium bisulfite), etc.

The term "Prognosis" is defined to include an assessment or prediction of the probable course, outcome, recovery or survival from a disease. Most physicians give a prognosis based on statistics of how a disease acts in studies on the general population. Prognosis can vary with

colorectal cancer depending on several factors, such as the stage of disease at diagnosis, type of cancer, and even gender.

The term "predicting the response", as used herein refers to the determination of the likelihood that the patient will respond either favourably or unfavourably to a given therapy. Especially, the term "prediction", as used herein, relates to an individual assessment of any parameter that can be useful in determining the evolution of a patient. As will be understood by those skilled in the art, the prediction of the clinical response to the treatment with a demethylating agent, although preferred to be, need not be correct for 100% of the colorectal cancer patients to be diagnosed or evaluated. The term, however, requires that a statistically significant portion of colorectal cancer patients can be identified as having an increased probability of having a positive response. Whether a subject is statistically significant can be determined without further ado by the person skilled in the art using various well known statistic evaluation tools, e.g., determination of confidence intervals, p-value determination, Student's t-test, Mann- Whitney test, etc. Details are found in Dowdy and Wearden, Statistics for Research, John Wiley & Sons, New York 1983. Preferred confidence intervals are at least 50%, at least 60%, at least 70%, at least 80%, at least 90% at least 95%. The p- values are, preferably, 0.2, 0.1 or 0.05.

The term "clinical response", as used herein, refers to the response of the colorectal cancer patient to a therapy with a demethylating agent. Standard criteria (Miller, et al. Cancer, 1981; 47:207-14) that can be used herewith to evaluate the response to a demethylating agent include response, stabilization and progression. It can be a complete response (or complete remission), which is the disappearance of all detectable malignant disease or a partial response, which is defined as approximately >50%> decrease in the sum of products of the largest perpendicular diameters of one or more lesions (tumor lesions), no new lesions and no progression of any lesion. Patients achieving complete or partial response are considered "responders", and all other patients are considered "non-responders".

The term “positive clinical response” as used herein refers to the outcome being either a complete or partial response.

The term “Wnt target gene biomarker” as used herein refers to a gene selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1 and ASCL2.

“LGR5, APCDD1, AXIN2, DKK1 and ASCL2” is the standard nomenclature as approved by the Human Genome Organisation. The genes encompass not only the particular sequences found in the publicly available database entries, but also encompass transcript variants of these sequences. Variant forms of the encoded proteins may comprise post-translational modification, may result from spliced messages, etc.... Variant sequences have at least 90%, preferably at least 91%, preferably at least 92%, preferably at least 93%, preferably at least 94%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, or preferably at least 99% identity to sequences in the database entries or sequence listing. Computer programs for determining

percent identity are available in the art, including Basic Local Alignment Search Tool (BLAST5) available from the National Center for Biotechnology Information. The genes are available as indicated hereafter.

"APCDD1" (Homo sapiens adenomatosis polyposis coli down-regulated 1, Accession number NM_153000.4) is the gene which encodes an inhibitor of the Wnt signalling pathway. Mutations at this locus have been associated with hereditary hypotrichosis simplex. It has been speculated that increased expression of this gene may be associated with colorectal carcinogenesis.

"LGR5" (Homo sapiens leucine-rich repeat containing G protein-coupled receptor 5, Accession number NM_003667.2) is a gene encoding a 7-transmembrane receptor that binds R-spondins. Overexpression of LGR5 is observed in some types of cancer; and has been reported to be frequently overexpressed in colorectal cancer cell lines. Moreover, LGR5 expression was higher in colorectal cancer cell lines derived from metastatic tumors compared with those from primary tumors. In clinical specimens, a significant overexpression of LGR5 has been described in 35 of 50 colorectal cancers (CRCs), and in seven of seven sporadic colonic adenomas, compared with matched normal mucosa.

"DKK1" (Homo sapiens dickkopf homolog 1, Accession number NM_012242.2) is a gene encoding a protein that is a member of the dickkopf family. It is a secreted protein with two cysteine rich regions and is involved in embryonic development through its inhibition of the WNT signaling pathway. Elevated levels of DKK1 in bone marrow plasma and peripheral blood is associated with the presence of osteolytic bone lesions in patients with multiple myeloma.

"ASCL2" (Homo sapiens achaete-scute complex homolog 2, Accession number NM_005170.2) is a gene which is a member of the basic helix-loop-helix (BHLH) family of transcription factors. It activates transcription by binding to the E box (5'-CANNTG-3'). Dimerization with other BHLH proteins is required for efficient DNA binding. It is involved in the determination of the neuronal precursors in the peripheral nervous system and the central nervous system.

"AXIN2" (Homo sapiens axin 2, NM_004655.3) is the gene encoding the Axin-related protein, AXIN2, which presumably plays an important role in the regulation of the stability of beta-catenin in the Wnt signaling pathway. Apparently, the deregulation of beta-catenin is an important event in the genesis of a number of malignancies. The AXIN2 gene has been mapped to 17q23-q24, a region that shows frequent loss of heterozygosity in breast cancer, neuroblastoma, and other tumors. Mutations in this gene have been associated with colorectal cancer with defective mismatch repair.

Detailed description of certain embodiments of the invention

The inventors have identified genetic markers, which are differentially expressed between samples from patients who had developed a tumor relapse in comparison to patients who remained

disease free. The inventors observed that certain genes belonging to the Wnt target genes can identify poor prognosis patients based on their low expression level both in the analysed patient group as well as in publically available datasets. This finding was entirely unexpected, as it was generally believed that high expression of Wnt target genes is related to poor prognosis. This relation was assumed, as expression of the Wnt target genes is associated with invasion, metastasis, immature phenotype, migration and proliferation. In addition, Wnt target gene expression is associated with cancer stem cell gene profiles and cancer stem cell profiles are associated with poor prognosis. In previous reports the association of a cancer stem cell signature with poor prognosis was attributed to a relative high number of cancer stem cells present in the malignant tissue, which enhances the chance of cancer stem cells shedding from the primary tumor. Indeed, adherence to an intestinal stem cell profile of individual cancer stem cells, but also similarity of breast cancers to a breast-cancer stem cell signature, was translated into an increased risk of metastasis and tumor recurrence (Liu et al., 2007; Merlos-Suarez et al., 2011).

It was therefore surprising that the detection of up regulation of a gene from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2, which are all typical Wnt target genes expressed in cancer stem cells, in a biological sample of a colorectal cancer patient indicates a decreased rather than an increased risk of developing metastasis and/or tumor relapse. The inventors have also studied the relation between hypermethylation of LGR5, APCDD1, DKK1, AXIN2 and ASCL2 and the risk of developing metastasis and/or tumor relapse. Methylation-specific PCR analysis and/or bisulphite sequencing revealed that these genes are methylated in a panel of colorectal cancer cell lines to different extend (Figure 2B). Furthermore, a strong correlation between the hypermethylation and the occurrence of metastasis and/or tumor relapse was found in patients who were treated for colorectal cancer stage II.

In one aspect, the invention is a method of determining the risk of developing metastasis and/or tumor relapse in a colorectal cancer patient, comprising the steps of providing a biological sample of said colorectal cancer patient, and determining in said biological sample the methylation status and/or the expression level of a gene selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2, and determining an increased risk of developing metastasis and/or tumor relapse, under the condition that hypermethylation and/or under expression is determined in a gene selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2.

In a preferred embodiment, an increased risk of developing metastasis and/or tumor relapse is determined, if hypermethylation and/or under expression is determined in at least two, preferably three, four or five genes of the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2. An advantage thereof is that if more of said genes are hypermethylated and/or under expressed, the sensitivity of the method is greater. For example, the inventors found that 100% of relapses occurred in patients that had relatively high levels of methylation in the five Wnt target gene biomarkers, whereas between 65 and 80% of the patients who had only a single Wnt target gene biomarker

which was hypermethylated (and the remaining four unmethylated), developed metastasis and/or tumor relapse.

More preferably, an increased risk of developing metastasis and/or tumor relapse is determined, if hypermethylation is determined in at least two, preferably three, four or five genes of the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2. An advantage thereof is that the detection of hypermethylation improves the sensitivity of the method. In a highly preferred embodiment, an increased risk of developing metastasis and/or tumor relapse is determined, if hypermethylation and/or under expression is determined in the genes selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2; LGR5, APCDD1, AXIN2 and DKK1; LGR5, APCDD1, AXIN2, and ASCL2, and AXIN2 and ASCL2. In a highly preferred embodiment, an increased risk of developing metastasis and/or tumor relapse is determined if hypermethylation is determined in all of LGR5, APCDD1, AXIN2, DKK1, and ASCL2; in all of LGR5, APCDD1, AXIN2 and DKK1; in all of LGR5, APCDD1, AXIN2, and ASCL2, or in both AXIN2 and ASCL2.

Wnt target gene biomarkers

Preferably, said colorectal cancer patient is suffering from colorectal cancer in stage I II or III, more preferably stage II and III.

As shown in the example section, the methods of the invention may involve techniques for measuring the expression level of certain genes. Various techniques for determining the expression level of a gene are known in the art and can be used in conjunction with the present invention.

Determination of Expression Levels

The expression level of a Wnt target gene biomarker can be measured by the biomarker's mRNA level, protein level, activity level, or other quantity reflected in or derivable from the biomarker's gene or protein expression data. The expression products of each of the Wnt target gene biomarkers include both RNA and protein. RNA products of the Wnt target gene biomarkers are transcriptional products of the Wnt target gene biomarkers and include populations of hnRNA, mRNA, and one or more spliced variants of mRNA. Protein products of the Wnt target gene biomarkers may also be measured. The protein products of the Wnt target gene biomarkers include, for example, proteins, protein variants arising from spliced mRNA variants, and post translationally modified proteins.

Any suitable means of measuring the expression of the RNA products of the Wnt target gene biomarkers can be used in accordance with the methods described herein. For example, the methods may utilize a variety of polynucleotides that specifically hybridize to one or more of the RNA products of the Wnt target gene biomarkers including, for example, oligonucleotides, cDNA, DNA, RNA, PCR products, synthetic DNA, synthetic RNA, or other combinations of naturally

occurring of modified nucleotides which specifically hybridize to one or more of the RNA products of the Wnt target gene biomarkers. Such polynucleotides may be used in combination with the methods to measure RNA expression including, for example, array hybridization, RT-PCR, nuclease protection and northern blots.

Array Hybridization

In one embodiment, the expression level of Wnt target gene biomarker may be determined using array hybridization to evaluate the level of RNA expression. Array hybridization utilizes nucleic acid members stably associated with a support that can hybridize with Wnt target gene biomarker expression products. The length of a nucleic acid member attached to the array can range from 8 to 1000 nucleotides in length and are chosen so as to be specific for the RNA products of the Wnt target gene biomarkers. The array may comprise, for example, one or more nucleic acid members that are specific for 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or all of the RNA products of the Wnt target gene biomarkers shown in Table 1, or variants thereof (e.g., splice variants). The nucleic acid members may be RNA or DNA, single or double stranded, and/or may be oligonucleotides or PCR fragments amplified from cDNA. Preferably oligonucleotides are approximately 10-100, 10-50, 20-50, or 20-30 nucleotides in length. Portions of the expressed regions of the Wnt target gene biomarkers can be utilized as probes on the array. More particularly oligonucleotides complementary to the Wnt target gene biomarkers genes and or cDNAs derived from the Wnt target gene biomarker genes are useful. For oligonucleotide based arrays, the selection of oligonucleotides corresponding to the gene of interest, which are useful as probes is well understood in the art. More particularly it is important to choose regions which will permit hybridization to the target nucleic acids. Factors such as the Tm of the oligonucleotide, the percent GC content, the degree of secondary structure and the length of nucleic acid are important factors. See for example U.S. Pat. No. 6,551,784.

Name	*GENE ID*	*Accession number*	*genomic assembly*	*genomic position*	amplified region *start*	*end*
APCDD1	147495	NM_153000.4	Human Feb. 2009 (GRCh37/hg19) Assembly	chr18:10,454,625 -10,488,697	10454683	10454850
LGR5	8549	NM_003667.2	Human Feb. 2009 (GRCh37/hg19)	chr12:71,833,813 -71,978,619	71834155	71834366
DKK1	22943	NM_012242.2	Human Feb. 2009 (GRCh37/hg19)	chr10:54,074,041 -54,077,416	54074538	54074695
ASCL2	430	NM_005170.2	Human Feb. 2009 (GRCh37/hg19)	chr11:2,289,729-2,292,182	2291397	2291677
AXIN2	8313	NM_004655.3	Human Feb. 2009 (GRCh37/hg19)	chr17:63,524,685 -63,557,740	63555273	63555530

Table 1 lists the Wnt target gene biomarkers, showing the Gene ID (NCBI, as published on www.ncbi.nlm.nih.gov on 3 october 2011), the Gene accession number (NCBI, as published on www.ncbi.nlm.nih.gov on 3 october 2011), the genomic assembly, the genomic position and the start and end position of the amplified region.

Arrays may be constructed, custom ordered, or purchased from a commercial vendor. Various methods for constructing arrays are well known in the art.

RT-PCR

In certain embodiments, the level of the expression of the RNA products of the Wnt target gene biomarkers can be measured by amplifying the RNA products of the biomarkers from a sample using reverse transcription (RT) in combination with the polymerase chain reaction (PCR). In certain embodiments, the RT can be quantitative as would be understood to a person skilled in the art.

Total RNA, or mRNA from a sample may be used as a template and a primer specific to the transcribed portion of a Wnt target gene biomarkers is used to initiate reverse transcription. Methods of reverse transcribing RNA into cDNA are well known and are described, for example, in Sambrook et al., 1989, supra. Primer design can be accomplished utilizing commercially available software (e.g., Primer Designer 1.0, Scientific Software etc.) or methods that are standard and well known in the art. Primer Software programs can be used to aid in the design and selection of primers include, for example, The Primer Quest software which is available through the following web site link: biotools.idtdna.com/primerquest/. Additionally, the following website links are useful when searching and updating sequence information from the Human Genome Database for use in biomarker primer design:

1) the NCBI LocusLink Homepage: world wide web at ncbi.nlm.nih.gov/LocusLink/, and 2) Ensemble Human Genome Browser: world wide web at ensembl.org/Homo_sapiens, preferably using pertinent biomarker information such as Gene or Sequence Description, Accession or Sequence ID, Gene Symbol, RefSeq #, and/or UniGene #.

General guidelines for designing primers that may be used in accordance with the methods described herein include the following: the product or amplicon length may be ~100-150 bases, the optimum Tm may be ~60° C., or about 58-62° C., and the GC content may be ~50%, or about 45-55%. Additionally, it may be desirable to avoid certain sequences such as one or more of the following: (i) strings of three or more bases at the 3'-end of each primer that are complementary to another part of the same primer or to another primer in order to reduce primer-dimer formation, (ii) sequences within a primer that are complementary to another primer sequence, (iii) runs of 3 or more G's or C's at the 3'-end, (iv) single base repeats greater than 3 bases, (v) unbalanced distributions of G/C- and A/T rich domains, and/or (vi) a T at the 3'-end.

The product of the reverse transcription is subsequently used as a template for PCR. PCR provides a method for rapidly amplifying a particular nucleic acid sequence by using multiple cycles of DNA replication catalyzed by a thermostable, DNA-dependent DNA polymerase to amplify the target sequence of interest. PCR requires the presence of a nucleic acid to be amplified, two single-stranded oligonucleotide primers flanking the sequence to be amplified, a DNA polymerase, deoxyribonucleoside triphosphates, a buffer and salts. The method of PCR is well known in the art. PCR, is performed as described in Mullis and Faloona, 1987, Methods Enzymol., 155: 335.

QRT-PCR, which is quantitative in nature, can also be performed to provide a quantitative measure of Wnt target gene biomarker gene expression levels. In QRT-PCR reverse transcription and PCR can be performed in two steps, or reverse transcription combined with PCR can be performed concurrently. One of these techniques, for which there are commercially available kits such as Taqman (Perkin Elmer, Foster City, Calif.), is performed with a transcript-specific antisense probe. This probe is specific for the PCR product (e.g. a nucleic acid fragment derived from a gene) and is prepared with a quencher and fluorescent reporter probe complexed to the 5' end of the oligonucleotide. Different fluorescent markers are attached to different reporters, allowing for measurement of two products in one reaction. When Taq DNA polymerase is activated, it cleaves off the quencher of the probe bound to the template by virtue of its 5'-to-3' exonuclease activity. In the absence of the quenchers, the reporters now fluoresce. The color change in the reporters is proportional to the amount of each specific product and is measured by a fluorometer; therefore, the amount of each color is measured and the PCR product is quantified. The PCR reactions are performed in 96 well plates so that samples derived from many individuals are processed and measured simultaneously. The Taqman system has the additional advantage of not requiring gel electrophoresis and allows for quantification when used with a standard curve.

A second technique useful for detecting PCR products quantitatively is to use an intercalating dye such as the commercially available QuantiTect SYBR Green PCR (Qiagen, Valencia Calif.). RT-PCR is performed using SYBR green as a fluorescent label which is incorporated into the PCR product during the PCR stage and produces a fluorescence proportional to the amount of PCR product. Additionally, other systems to quantitatively measure mRNA expression products are known including Molecular Beacons™.

Additional techniques to quantitatively measure RNA expression include, but are not limited to, polymerase chain reaction, ligase chain reaction, Qbeta replicase (see, e.g., International Application No. PCT/US87/00880), isothermal amplification method (see, e.g., Walker et al. (1992) PNAS 89:382-396), strand displacement amplification (SDA), repair chain reaction, Asymmetric Quantitative PCR (see, e.g., U.S. Publication No. US200330134307A1) and the multiplex microsphere bead assay described in Fuja et al., 2004, Journal of Biotechnology 108:193-205.

The level of gene expression can be measured by amplifying RNA from a sample using transcription based amplification systems (TAS), including nucleic acid sequence amplification (NASBA) and 3SR. See, e.g., Kwoh et al (1989) PNAS USA 86:1173; International Publication No. WO 88/10315; and U.S. Pat. No. 6,329,179. In NASBA, the nucleic acids may be prepared for amplification using conventional phenol/chloroform extraction, heat denaturation, treatment with lysis buffer and minispin columns for isolation of DNA and RNA or guanidinium chloride extraction of RNA. These amplification techniques involve annealing a primer that has target specific sequences. Following polymerization, DNA/RNA hybrids are digested with RNase H while double stranded DNA molecules are heat denatured again. In either case the single stranded DNA is made fully double stranded by addition of second target specific primer, followed by polymerization. The double-stranded DNA molecules are then multiply transcribed by a polymerase such as T7 or SP6. In an isothermal cyclic reaction, the RNA's are reverse transcribed into double stranded DNA, and transcribed once with a polymerase such as T7 or SP6. The resulting products, whether truncated or complete, indicate target specific sequences.

Alternatively, visualization may be achieved indirectly. Following separation of amplification products, a labeled, nucleic acid probe is brought into contact with the amplified nucleic acid sequence of interest. The probe may be conjugated to a chromophore, radiolabeled, or conjugated to a binding partner, such as an antibody or biotin, where the other member of the binding pair carries a detectable moiety.

Additionally, detection may be carried our using Southern blotting and hybridization with a labeled probe. The techniques involved in Southern blotting are well known to those of skill in the art and may be found in many standard books on molecular protocols.

Nuclease Protection Assays

In certain embodiments, Nuclease protection assays (including both ribonuclease protection assays and S1 nuclease assays) can be used to detect and quantitate RNA products of the Wnt target gene biomarkers. In nuclease protection assays, an antisense probe (e.g., radiolabeled or nonisotopic labeled) hybridizes in solution to an RNA sample. Following hybridization, single-stranded, unhybridized probe and RNA are degraded by nucleases. An acrylamide gel is used to separate the remaining protected fragments. Typically, solution hybridization can accommodate up to ~100 μg of sample RNA whereas blot hybridizations may only be able to accommodate ~20-30 μg of RNA sample.

The ribonuclease protection assay, which is the most common type of nuclease protection assay, requires the use of RNA probes. Oligonucleotides and other single-stranded DNA probes can only be used in assays containing S1 nuclease. The single-stranded, antisense probe must typically be completely homologous to target RNA to prevent cleavage of the probe:target hybrid by nuclease.

Northern Blots

A standard Northern blot assay can also be used to ascertain an RNA transcript size, identify alternatively spliced RNA transcripts, and the relative amounts of RNA products of the Wnt target gene biomarkers, in accordance with conventional Northern hybridization techniques known to those persons of ordinary skill in the art. In Northern blots, RNA samples are first separated by size via electrophoresis in an agarose gel under denaturing conditions. The RNA is then transferred to a membrane, crosslinked and hybridized with a labeled probe. Nonisotopic or high specific activity radiolabeled probes can be used including random-primed, nick-translated, or PCR-generated DNA probes, in vitro transcribed RNA probes, and oligonucleotides. Additionally, sequences with only partial homology (e.g., cDNA from a different species or genomic DNA fragments that might contain an exon) may be used as probes. The labeled probe, e.g., a radiolabeled cDNA, either containing the full-length, single stranded DNA or a fragment of that DNA sequence may be any length up to at least 20, at least 30, at least 50, or at least 100 consecutive nucleotides in length. The probe can be labeled by any of the many different methods known to those skilled in this art. The labels most commonly employed for these studies are radioactive elements, enzymes, chemicals that fluoresce when exposed to ultraviolet light, and others. A number of fluorescent materials are known and can be utilized as labels. These include, but are not limited to, fluorescein, rhodamine, auramine, Texas Red, AMCA blue and Lucifer Yellow. A particular detecting material is anti-rabbit antibody prepared in goats and conjugated with fluorescein through an isothiocyanate. Non-limiting examples of isotopes include 3H, 14C, 32P, 35S, 36Cl, 51Cr, 57Co, 58Co, 59Fe, 90Y, 125I, 131I, and 186Re. Enzyme labels are likewise useful, and can be detected by any of the presently utilized colorimetric, spectrophotometric, fluorospectrophotometric, amperometric or gasometric techniques. The enzyme may be conjugated to the selected probe by reaction with bridging

molecules such as carbodiimides, diisocyanates, glutaraldehyde and the like. Any enzymes known to one of skill in the art can be utilized, including, for example, peroxidase, beta-D-galactosidase, urease, glucose oxidase plus peroxidase and alkaline phosphatase. U.S. Pat. Nos. 3,654,090, 3,850,752, and 4,016,043 are referred to by way of example for their disclosure of alternate labeling material and methods.

Protein Products

The expression level of a Wnt target gene biomarker may also be measured by the biomarker's protein level using any art-known method. Traditional methodologies for protein quantification include 2-D gel electrophoresis, mass spectrometry and antibody binding. Preferred methods for assaying biomarker protein levels in a biological sample include antibody-based techniques, such as immunoblotting (western blotting), immunohistological assay, enzyme linked immunosorbent assay (ELISA), radioimmunoassay (RIA), or protein chips. For example, a biomarker-specific monoclonal antibodies can be used both as an immunoadsorbent and as an enzyme-labeled probe to detect and quantify the biomarker. The amount of biomarker present in the sample can be calculated by reference to the amount present in a standard preparation using a linear regression computer algorithm. In another embodiment, Wnt target gene biomarkers may be immunoprecipitated from a biological sample (e.g., directly from urine or serum or from a lysate of cells, etc.) using an antibody specific for said biomarker. The isolated proteins may then be run on an SDS-PAGE gel and blotted (e.g., to nitrocellulose or other suitable material) using standard procedures. The blot may then be probed with an anti-biomarker specific antibody to determine the expression level of the Wnt target gene biomarkers.

Gel electrophoresis, immunoprecipitation and mass spectrometry may be carried out using standard techniques, for example, such as those described in Molecular Cloning A Laboratory Manual, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989), Harlow and Lane, Antibodies: A Laboratory Manual (1988 Cold Spring Harbor Laboratory), G. Suizdak, Mass Spectrometry for Biotechnology (Academic Press 1996).

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (mAb) is meant to include intact molecules as well as antibody portions (such as, for example, Fab, Fab', F(ab')2, Fv, single chain Fv, or Fd) which are capable of specifically binding to a Wnt target gene biomarker.

Expression Level Reference Samples

In some embodiments, expression levels of Wnt target gene biomarkers in a biological sample of interest (e.g., a colorectal cancer biopsy) are compared to a Wnt target gene biomarker expression level in an expression level reference sample. The expression level reference sample may be a biological sample derived from one or more colorectal cancer patients determined to be suffering from metastasis and/or tumor relapse. In other words, the expression level reference

sample serves as a standard with which to compare expression level values for each Wnt target gene biomarker in a test sample. In other embodiments, the expression level reference sample is a sample comprising a colon cancer cell line, preferably multiple cell lines. An increase of Wnt target gene biomarker expression levels compared to the expression level values in a reference sample indicates that the patient has a good prognosis. Accordingly, a decrease of Wnt target gene biomarker expression levels compared to the expression level values in a reference sample indicates that the patient has a bad prognosis. In some embodiments, Wnt target gene biomarker threshold expression level values are optionally set based on one or more statistical criteria for deviation from Wnt target gene biomarker expression level values in an expression level reference sample, e.g., two or more SDs away from the value for a reference sample Wnt target gene biomarker expression level.

In some embodiments, the expression level reference sample is a "negative" reference sample, i.e., a sample of a colorectal cancer patient who did not develop a tumor relapse or a sample from a healthy individual. In another embodiment, said reference sample comprises cells which do not have a mutation in the human APC gene (NC_000005.9), preferably from cell line RKO, which can be purchased from LGC Standards GmbH, Germany under ATCC number CRL-2577. Thus, where expression levels of a Wnt target gene biomarker is significantly lower than the threshold expression level value based on the negative reference sample, the patient's cancer is indicated as a patient having an increased risk of developing metastasis and/or tumor relapse.

In some embodiments, the expression level reference sample is a "positive" reference sample, i.e., a sample from colon cancer carcinoma, a sample comprising cells from a colon cancer cell line or a sample from a patient suffering from colorectal cancer in stage III or IV. In some embodiments, said control sample comprises cell from HCT-15, HCT-116, SW-620, SW-837 or Colo-205 cell lines, which can be purchased from LGC Standards GmbH, Germany. Thus, where expression level of a Wnt target gene biomarker is significantly higher than the threshold expression level values based on the positive reference sample, the patient's cancer is indicated as having a decreased risk of developing metastasis and/or tumor relapse.

In some embodiments, Wnt target gene biomarker expression profiles are compared to those in both positive and negative reference samples.

In some embodiments an expression level reference sample is a population of nucleic acids (e.g., mRNAs, aRNAs, cDNAs, or aRNAs) derived from a cancer biopsy sample within which the sequences of at least four Wnt target gene biomarkers are represented, and for which sensitivity to a demethylating agent has been determined. In some embodiments, the population of nucleic acids is derived from patient tumor cells cultivated in culture. In other embodiments, the population is derived directly from a biopsy without a cell culture step.

Methods for obtaining RNA from biological samples (e.g., tissues or cells) including linear aRNA amplification from single cells include, e.g., Luzzi et al. (2005), Methods Mol. Biol., 293:187-207. Further, diverse kits for high quality RNA purification are available commercially,

e.g., from Qiagen (Valencia, Calif.), Invitrogen (Carlsbad, Calif.), Clontech (Palo Alto, Calif.), and Stratagene (La Jolla, Calif.).

Method using methylation status of APCDD1, LGR5, DKK1, ASCL2 and AXIN2

As shown in the example section, the methods of the invention may involve detecting hypermethylation of the nucleic acid of certain genes. Surprisingly, the inventors found that the methylation status of any of the Wnt target gene biomarkers predicts the development of metastasis and/or tumor relapse with higher sensitivity that the expression level of the same gene(s). For example, using the methylation status of LGR5 instead of the expression level of LGR5 improves the sensitivity of the prediction of tumor relapse. Therefore, in a preferred embodiment, methylation of at least one of the Wnt target gene biomarkers is determined and the condition for the determination of the risk of developing metastasis and/or tumor relapse, the prediction of the response to a treatment, the identification and/or selection of a colorectal cancer patient, the selection of a suitable treatment regimen or the treatment of a patient according to methods of the invention is based on hypermethylation of at least one of the of the Wnt target gene biomarkers.

Various techniques for assessing the methylation status of a gene are known in the art and may be used in connection with the present invention.

Methylation status

The method according to the invention may involve detecting the methylation status of a nucleic acid in a sample of a subject. In certain embodiments of the invention, the detected methylation status is hypermethylation, an increase in the methylation of cytosines in CpG sites. The increase can be about 1%, about 5%, about 10%, about 20%, about 25%, about 30%, about 35%, about 40%, about 45%, about 50%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, about 90%, about 95%, or more than about 95% greater than the extent of methylation of cytosines typically expected or observed for the CpG site or sites being evaluated.

Methylation status Reference Samples

In certain embodiments of the invention, the detected methylation status is determined in part or wholly on the basis of a comparison with a control. The control can be a value or set of values related to the extent and/or pattern of methylation in a control sample. In certain embodiments of the invention, such a value or values may be determined, for example, by calculations, using algorithms, and/or from previously acquired and/or archived data. In certain embodiments of the invention, the value or set of values for the control is derived from experiments performed on samples or using a subject. For example, control data can be derived from experiments on samples derived from comparable tissues or cells, such as normal tissue adjacent to tumors.

In certain embodiments of the invention, a control comprising DNA that is mostly or entirely demethylated, at one or more of the CpG sites being analyzed, is used. Such a control might be obtained, for example, from mutant tissues or cells lacking methyltransferase activity and/or from tissues or cells that have been chemically demethylated. For example, controls may be obtained from tissues or cells lacking activity of methyltransferases Dnmtl, Dnmt2, Dnmt3a, Dnmt3b, or combinations thereof. Agents such as 5-aza-2'-deoxycytidine may be used to chemically demethylate DNA.

In certain embodiments of the invention, a control comprising DNA that is mostly or entirely methylated, at one or more of the CpG sites being analyzed, is used. Such a control might be obtained, for example, from cells or tissues that are known or expected to be mostly or entirely methylated at the CpG site or sites of interest. Such a control could also be obtained by cells or tissues in which methylation levels have been altered and/or manipulated, for example, by overexpression of methyltransferases (such as enzymes Dnmtl, Dnmt2, Dnmt3a, Dnmt3b, any of the bacterial 5-CpG methyltransferases, or combinations thereof). In certain embodiments of the invention, samples used to obtain control values are processed and/or manipulated in the same manner as the samples being evaluated. In a preferred embodiment, said control comprises a colorectal cancer cell line with a known methylation status. In preferred embodiments, said control sample comprises cell from HCT-15 (ATCC number CCL-225), HCT-116 (ATCC number CCL-247), SW-620 (ATCC number CCL-227), SW-837 (ATCC number CCL-235) or Colo-205 (ATCC number CCL-222) cell lines, which can be purchased from LGC Standards GmbH, Germany. In a preferred embodiment, a sample comprising HCT-116 cell line is used as a positive control for the detection of hypermethylation of LGR5, APCCD1 or ASCL2. In another preferred embodiment, the SW837 cell line is used as a positive control for detecting hypermethylation of APCDD1 and ASCL2. In another preferred embodiment, the Colo205 cell line is used as a positive control for detecting hypermethylation of APCDD1 and DKK1. In some embodiments, negative control samples are generated by treating positive control samples with a demethylating agent, preferably 5-aza-2'-deoxycytidine.

Detection of methylation status

Any of variety of techniques to detect methylation status can be used in the practice of inventive methods described herein. The following descriptions provide some examples of such techniques, and are not intended to limit the types of techniques that can be used with such methods. As will be understood of one of ordinary skill in the art, variations to the described techniques can also be used in accordance with inventive methods described herein. In certain embodiments of the invention, two or more methods of detecting methylation status are used together or in combination. In certain embodiments of the invention, the technique or techniques used to detect methylation

status is or are quantitative. For example, such methods may provide estimates of the percentage of DNA molecules in a sample that are methylated at one or more particular CpG sites.

In the present invention, any nucleic acid sample, in purified or nonpurified form, can be used, provided it contains or is suspected of containing, a nucleic acid sequence containing a target locus (e.g., CpG-containing nucleic acid). One nucleic acid region capable of being differentially methylated is a CpG site, a sequence of nucleic acid with an increased density relative to other nucleic acid regions of the dinucleotide CpG. The CpG doublet occurs in vertebrate DNA at only about 20% of the frequency that would be expected from the proportion of G*C base pairs. In certain regions, the density of CpG doublets reaches the predicted value; it is increased by ten-fold relative to the rest of the genome. CpG sites have an average G*C content of about 60%, compared with the 40% average in bulk DNA. The sites take the form of stretches of DNA typically about one to two kilobases long.

In many genes, the CpG sites begin just upstream of a promoter and extend downstream into the transcribed region. Methylation of a CpG site at a promoter usually suppresses expression of the gene. The sites can also surround the 5' region of the coding region of the gene as well as the 3' region of the coding region. Thus, CpG sites can be found in multiple regions of a nucleic acid sequence including 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.

In general, the CpG-containing nucleic acid is DNA. However, the method according to the present invention may employ, for example, samples that contain DNA, or DNA and RNA containing mRNA, wherein DNA or RNA may be single-stranded or double-stranded, or a DNA-RNA hybrid may be included in the sample.

A mixture of nucleic acids may also be used. The specific nucleic acid sequence to be detected may be a fraction of a larger molecule or can be present initially as a discrete molecule, so that the specific sequence constitutes the entire nucleic acid. It is not necessary that the sequence to be studied be present initially in a pure form; the nucleic acid may be a minor fraction of a complex mixture, such as contained in whole human DNA. Nucleic acids contained in a sample used for detection of methylated CpG sites may be extracted by a variety of techniques such as that described by Sambrook, et al. (Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, N.Y., 1989).

A nucleic acid can contain a regulatory region which is a region of DNA that encodes information or controls transcription of the nucleic acid. Regulatory regions include at least one promoter. A "promoter" is a minimal sequence sufficient to direct transcription, and renders promoter-dependent gene expression controllable for cell-type specific, tissue-specific, or inducible by external signals or agents. Promoters may be located in the 5' or 3' region of the gene. The number of nucleic acids in all or part of promoter regions can be used to measure CG-site

methylation. Moreover, it is generally recognized that methylation of the target gene promoter proceeds naturally from the outer boundary inward. Therefore, an early stage of cell conversion can be detected by analyzing methylation in these outer areas of the promoter region.

Methods for Detection of Methylation

Methylation-Specific PCR

When genomic DNA is treated with bisulfite, cytosine in the 5'-CpG'-3 region remains intact, if it was methylated, but the cytosine changes to uracil, if it was unmethylated. Accordingly, based on the base sequence converted after bisulfite treatment, PCR primer sets corresponding to a region having the 5'-CpG-3' base sequence are constructed. Herein, the constructed primer sets are two kinds of primer sets: a primer set corresponding to the methylated base sequence, and a primer set corresponding to the unmethylated base sequence. When genomic DNA is converted with bisulfite and then amplified by PCR using the above two kinds of primer sets, the PCR product is detected in the PCR mixture employing the primers corresponding to the methylated base sequence, if the genomic DNA was methylated, but the genomic DNA is detected in the PCR mixture employing the primers corresponding to the unmethylated, if the genomic DNA was unmethylated. This methylation can be quantitatively analyzed by agarose gel electrophoresis.

Real-Time Methylation Specific PCR

Real-time methylation-specific PCR is a real-time measurement method modified from the methylation-specific PCR method and comprises treating genomic DNA with bisulfite, designing PCR primers corresponding to the methylated base sequence, and performing real-time PCR using the primers. Methods of detecting the methylation of the genomic DNA include two methods: a method of detection using a TaqMan probe complementary to the amplified base sequence; and a method of detection using SYBR green. Thus, the real-time methylation-specific PCR allows selective quantitative analysis of methylated DNA. Herein, a standard curve is plotted using an in vitro methylated DNA sample, and a gene containing no 5'-CpG-3' sequence in the base sequence is also amplified as a negative control group for standardization to quantitatively analyze the degree of methylation.

Pyrosequencing

The pyrosequencing method is a quantitative real-time sequencing method modified from the bisulfite sequencing method. Similarly to bisulfite sequencing, genomic DNA is converted by bisulfite treatment, and then, PCR primers corresponding to a region containing no 5'-CpG-3' base sequence are constructed. Specifically, the genomic DNA is treated with bisulfite, amplified using the PCR primers, and then subjected to real-time base sequence analysis using a sequencing primer.

The degree of methylation is expressed as a methylation index by analyzing the amounts of cytosine and thymine in the 5'-CpG-3' region.

PCR Using Methylated DNA-specific binding protein, quantitative PCR, And DNA Chip Assay

When a protein binding specifically only to methylated DNA is mixed with DNA, the protein binds specifically only to the methylated DNA. Thus, either PCR using a methylation-specific binding protein or a DNA chip assay allows selective isolation of only methylated DNA. Genomic DNA is mixed with a methylation-specific binding protein, and then only methylated DNA was selectively isolated. The isolated DNA is amplified using PCR primers corresponding to the promoter region, and then methylation of the DNA is measured by agarose gel electrophoresis.

In addition, methylation of DNA can also be measured by a quantitative PCR method, and methylated DNA isolated with a methylated DNA-specific binding protein can be labeled with a fluorescent probe and hybridized to a DNA chip containing complementary probes, thereby measuring methylation of the DNA. Herein, the methylated DNA-specific binding protein may be, but not limited to, McrBt.

Methylation-Sensitive Restriction Enzyme

Detection of differential methylation can be accomplished by bringing a nucleic acid sample into contact with a methylation-sensitive restriction endonuclease that cleaves only unmethylated CpG sites.

In a separate reaction, the sample is further brought into contact with an isoschizomer of the methylation-sensitive restriction enzyme that cleaves both methylated and unmethylated CpG-sites, thereby cleaving the methylated nucleic acid.

Specific primers are added to the nucleic acid sample, and the nucleic acid is amplified by any conventional method. The presence of an amplified product in the sample treated with the methylation-sensitive restriction enzyme but absence of an amplified product in the sample treated with the isoschizomer of the methylation-sensitive restriction enzyme indicates that methylation has occurred at the nucleic acid region assayed. However, the absence of an amplified product in the sample treated with the methylation-sensitive restriction enzyme together with the absence of an amplified product in the sample treated with the isoschizomer of the methylation-sensitive restriction enzyme indicates that no methylation has occurred at the nucleic acid region assayed.

As used herein, the term "methylation-sensitive restriction enzyme" refers to a restriction enzyme (e.g., SmaI) that includes CG as part of its recognition site and has activity when the C is methylated as compared to when the C is not methylated. Non-limiting examples of methylation-sensitive restriction enzymes include MspI, HpaII, BssHII, BstUI and NotI. Such enzymes can be

used alone or in combination. Examples of other methylation-sensitive restriction enzymes include, but are not limited to SacII and EagI.

The isoschizomer of the methylation-sensitive restriction enzyme is a restriction enzyme that recognizes the same recognition site as the methylation-sensitive restriction enzyme but cleaves both methylated and unmethylated CGs. An example thereof includes MspI.

Primers of the present invention are designed to be "substantially" complementary to each strand of the locus to be amplified and include the appropriate G or C nucleotides as discussed above. This means that the primers must be sufficiently complementary to hybridize with their respective strands under polymerization reaction conditions. Primers of the present invention are used in the amplification process, which is an enzymatic chain reaction (e.g., PCR) in which that a target locus exponentially increases through a number of reaction steps. Typically, one primer is homologous with the negative (-) strand of the locus (antisense primer), and the other primer is homologous with the positive (+) strand (sense primer). After the primers have been annealed to denatured nucleic acid, the nucleic acid chain is extended by an enzyme such as DNA Polymerase I (Klenow), and reactants such as nucleotides, and, as a result, + and - strands containing the target locus sequence are newly synthesized. When the newly synthesized target locus is used as a template and subjected to repeated cycles of denaturing, primer annealing, and extension, exponential synthesis of the target locus sequence occurs. The resulting reaction product is a discrete nucleic acid duplex with termini corresponding to the ends of specific primers employed.

The amplification reaction is PCR which is commonly used in the art. However, alternative methods such as real-time PCR or linear amplification using isothermal enzyme may also be used. In addition, multiplex amplification reactions may also be used.

Bisulfite Sequencing Method

Another method for detecting a methylated CpG-containing nucleic acid comprises the steps of: bringing a nucleic acid-containing sample into contact with an agent that modifies unmethylated cytosine; and amplifying the CpG-containing nucleic acid in the sample using CpG-specific oligonucleotide primers, wherein the oligonucleotide primers distinguish between modified methylated nucleic acid and non-methylated nucleic acid and detect the methylated nucleic acid. The amplification step is optional and desirable, but not essential. The method relies on the PCR reaction to distinguish between modified (e.g., chemically modified) methylated DNA and unmethylated DNA. Such methods are described in U.S. Pat. No. 5,786,146 relating to bisulfite sequencing for detection of methylated nucleic acid.

In real-time embodiments, quantitation of (hyper)methylation may be on an absolute basis, or may be relative to a constitutively methylated DNA standard, or may be relative to an unmethylated DNA standard. Methylation status may be determined by using the ratio between the signal of the marker under investigation and the signal of a reference gene where methylation status

is known (such as β-actin for example), or by using the ratio between the methylated marker and the sum of the methylated and the non-methylated marker. Alternatively, absolute copy number of the methylated marker gene can be determined.

Suitable controls may need to be incorporated in order to ensure the method chosen is working correctly and reliably. Suitable controls may include assessing the methylation status of a gene known to be methylated. This experiment acts as a positive control to ensure that false negative results are not obtained. The gene may be one which is known to be methylated in the sample under investigation or it may have been artificially methylated. In one embodiment, the gene of interest may be assessed in normal cells, following treatment with SssI methyltransferase, as a positive control. Additionally or alternatively, suitable negative controls may be employed with the methods of the invention. Here, suitable controls may include assessing the methylation status of a gene known to be unmethylated or a gene that has been artificially demethylated. This experiment acts as a negative control to ensure that false positive results are not obtained. In one embodiment, the gene of interest may be assessed in normal cells as a negative control, in particular if the gene is unmethylated in normal tissues.

Substrates

After the target nucleic acid region has been amplified, the nucleic acid amplification product can be hybridized to a known gene probe attached to a solid support (substrate) to detect the presence of the nucleic acid sequence.

As used herein, the term "substrate", when used in reference to a substance, structure, surface or material, means a composition comprising a non-biological, synthetic, non-living, planar or round surface that is not heretofore known to comprise a specific binding, hybridization or catalytic recognition site or a plurality of different recognition sites or a number of different recognition sites which exceeds the number of different molecular species comprising the surface, structure or material. Examples of the substrate include, but are not limited to, semiconductors, synthetic (organic) metals, synthetic semiconductors, insulators and dopants; metals, alloys, elements, compounds and minerals; synthetic, cleaved, etched, lithographed, printed, machined and microfabricated slides, devices, structures and surfaces; industrial polymers, plastics, membranes silicon, silicates, glass, metals and ceramics; and wood, paper, cardboard, cotton, wool, cloth, woven and non-woven fibers, materials and fabrics; and amphibious surfaces.

It is known in the art that several types of membranes have adhesion to nucleic acid sequences. Specific non-limiting examples of these membranes include nitrocellulose or other membranes used for detection of gene expression such as polyvinylchloride, diazotized paper and other commercially available membranes such as GENESCREEN™, ZETAPROBET™ (Biorad), and NYTRANT™. Beads, glass, wafer and metal substrates are also included. Methods for

attaching nucleic acids to these objects are well known in the art. Alternatively, screening can be done in a liquid phase.

Hybridization Conditions

In nucleic acid hybridization reactions, the conditions used to achieve a particular level of stringency will vary depending on the nature of the nucleic acids being hybridized. For example, the length, degree of complementarity, nucleotide sequence composition (e.g., GC/AT content), and nucleic acid type (e.g., RNA/DNA) of the hybridizing regions of the nucleic acids can be considered in selecting hybridization conditions. An additional consideration is whether one of the nucleic acids is immobilized, for example, on a filter.

An example of progressively higher stringency conditions is as follows: 2×SSC/0.1% SDS at room temperature (hybridization conditions); 0.2×SSC/0.1% SDS at room temperature (low stringency conditions); 0.2×SSC/0.1% SDS at 42° C. (moderate stringency conditions); and 0.1×SSC at about 68° C. (high stringency conditions).

Washing can be carried out using only one of these conditions, e.g., high stringency conditions, or each of the conditions can be used, e.g., for 10-15 minutes each, in the order listed above, repeating any or all of the steps listed. However, as mentioned above, optimal conditions will vary depending on the particular hybridization reaction involved, and can be determined empirically. In general, conditions of high stringency are used for the hybridization of the probe of interest.

Label

The probe of interest can be detectably labeled, for example, with a radioisotope, a fluorescent compound, a bioluminescent compound, a chemiluminescent compound, a metal chelator, or an enzyme. Appropriate labeling with such probes is widely known in the art and can be performed by any conventional method.

Medical treatment using a demethylating agent

The inventors found that treatment of colorectal cancer cells with the demethylating agent 5-Azacytidine (5-Aza) resulted in marked upregulation of LGR5, APCDD1, DKK1, AXIN2 and ASCL2 specifically in those cell lines where methylation was evident (Figure 2C). For instance, APCDD1 expression was enhanced by 5-Aza in the lines where APCDD1 promoter methylation was clearly detectable.

Intriguingly, the treatment of colorectal cancer cell lines with a demethylating agent resulted in markedly decreased clonogenicity of these lines (Figure 3A). Importantly, this was a general observation as also primary isolated colon-cancer stem cell cultures treated with 5-Aza demonstrated significantly lower clonogenicity as determined by limiting dilution analysis,

suggesting that the fraction of cancer stem cells in these cultures decreased (Figure 3B). Also in an in vivo model system in which primary cancer stem cell-induced xenograft tumors were growing subcutaneously the inventors observed that 5-Aza treatment resulted in markedly suppressed tumor growth (Figure 3C). Importantly, analysis of these 5-Aza treated xenografts confirmed the efficacy of 5-Aza on re-expression of the Wnt target genes in vivo (Figure 3D), validating the methylation-dependent regulation of Wnt target genes also in this in vivo model.

Re-expression of Wnt target genes as a result of treatment with a demethylating agent has clinical relevance for a colorectal cancer patient wherein the Wnt target gene biomarkers are down regulated and/or hypermethylated. To show this, the inventors have determined the effect of 5-Aza on Wnt activity levels. In several colorectal cancer cell lines, 5-Aza-mediated expression of genes previously suppressed by methylation resulted in decreased Wnt activity (Figure 3E). This already indicates an effect of 5-Aza on the activity of the Wnt pathway, but a more direct proof was deduced from the observation that overexpression of either AXIN2 or APCDD1, both methylated in colorectal cancer, was sufficient to decrease Wnt signaling levels (Figure 3F). This was not only observed in colorectal cancer cell lines, but confirmed in a primary cancer stem cell culture as well (Figure 3F, right bars, Co100). It is therefore likely that treatment of a colorectal cancer patient who suffers from metastasis and/or a tumor relapse will benefit from a treatment with a demethylating agent, because said patient is likely to have a methylated gene selected from LGR5, APCDD1, DKK1, AXIN2 and ASCL2.

In another aspect, the invention provides a method of predicting the response in a colorectal cancer patient to treatment to colorectal cancer. The drug for the treatment is a demethylating agent. In preferred embodiments said demethylating agent is used in combination with another chemotherapy. The presence of hypermethylation and/or under expression indicates that the Wnt target genes of the invention are methylated to a higher extend, which enables intervention with a demethylating agent. Therefore, this presence indicates a more positive clinical response to said treatment, whereas the absence of methylation or a lower level of methylation compared to a control sample indicates an unsuccessful clinical response to the treatment. If a positive clinical response to treatment with a demethylating agent is determined, the patient is identified or selected for a treatment with said demethylating agent. In other cases, the patient is not selected for treatment with a demethylating agent, and one or more alternative drug or medical intervention may be more beneficial for the treatment of the colorectal cancer patient.

Therefore, the invention concerns a method of predicting the response in a colorectal cancer patient to treatment with a demethylating agent comprising providing a biological sample of said colorectal cancer patient, and determining in said biological sample the methylation status and/or the expression level of a gene selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2, and predicting a positive clinical response to said treatment, if

hypermethylation and/or under expression is determined in a gene selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2.

The demethylating agent may be any demethylating agent which is suitable for a medical treatment. Many demethylating agents are known in the art, including 5'aza-cytidine, 5-aza-2'-deoxycytidine (decitabine), zebularine [1-(β-D-ribofuranosyl(dihydro-pyrimidin-2-1)] and L-methionine.

Demethylating agents, according to the present invention, are preferably selected from the group consisting of 5'aza-cytidine, 5-aza-2'-deoxycytidine (decitabine), zebularine [1-(β-D-ribofuranosyl(dihydro-pyrimidin-2-1)], L-methionine, inhibitors of histone deacetylase (HDAC) such as, for instance, valproic acid or trichostatin A, apicidine, hydralazine, procainamide (pronestyl), antisense oligonucleotides directed against DNA methyltransferase messenger RNA, their admixtures and derivatives thereof. Preferably, the demethylating agents to be used in the present invention are 5'aza-cytidine, decitabine and zebularine.

The demethylating agent is preferably used in a pharmaceutically effective amount. The term "pharmaceutically effective amount" shall mean that amount of a drug or pharmaceutical agent that will elicit the biological or medical response of a tissue, system, animal or human that is being sought by a researcher or clinician. This amount can be a therapeutically effective amount. It is of course preferable that this therapeutically effective amount will achieve the goal of improvement in disease severity and the frequency of incidence over treatment of the agent by itself, and/or of amelioration of adverse side effects typically associated with alternative therapies.

The invention further provides a method for identifying and/or selecting a colorectal cancer patient with colorectal cancer or treated for colorectal cancer suitable for treatment with a demethylating agent comprising providing a biological sample of said colorectal cancer patient, and determining in said biological sample the methylation status and/or the expression level of a gene selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2, and identifying and/or selecting the colorectal cancer patient for treatment with said demethylated agent if hypermethylation and/or under expression is determined in a gene selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2.

Further provided is a method for selecting a suitable treatment regimen for colorectal cancer in a colorectal cancer patient comprising providing a biological sample of said colorectal cancer patient, and determining in said biological sample the methylation status and/or the expression level of a gene selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2, and selecting said demethylating agent for treatment if hypermethylation and/or under expression is determined in a gene selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2.

The invention therefore further provides a method of treating a patient suffering from colorectal cancer with a demethylating agent comprising providing a biological sample of said

colorectal cancer patient, and determining in said biological sample the methylation status and/or the expression level of a gene selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2, and selecting said demethylating agent for treatment if hypermethylation and/or under expression is determined in a gene selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2.

Primer for the detection of methylation in colorectal cancer

The invention further provides a primer or primer pair for determining the methylation status of at least one gene selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2, wherein the primer or primer pair comprises the nucleotide sequence or sequences selected from Seq ID No. 31, Seq ID No. 32, Seq ID No. 33, Seq ID No. 34, Seq ID No. 35, Seq ID No. 36, Seq ID No. 37, Seq ID No. 38, Seq ID No. 39, Seq ID No. 40, Seq ID No. 41, Seq ID No. 42, Seq ID No. 43, Seq ID No. 44, Seq ID No. 45, Seq ID No. 46, Seq ID No. 47, Seq ID No. 48, Seq ID No. 49, Seq ID No. 50, Seq ID No. 51, Seq ID No. 52, Seq ID No. 53, Seq ID No. 54, Seq ID No. 55, Seq ID No. 56, Seq ID No. 57 and Seq ID No. 58. Preferably, the forward and reverse primers which form a primer pair (as can be understood from the detailed list in the example section) are used together.

Kits for the detection of methylation in colorectal cancer

Still another aspect of the invention is a kit for assessing methylation in a test sample. The kit comprises optionally 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. The kit also comprises a pair of oligonucleotide primers that specifically hybridizes under amplification conditions to the methylated gene following treatment with a reagent, which gene is selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2.

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. In certain embodiments of the invention, guidance is provided as to how to interpret the detected methylation status in order to make a diagnosis or other assessment in relation to cancer.

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. Typically the kit will contain oligonucleotide primers which specifically hybridize to regions within 1 kb of the transcription start sites of the genes identified in Table 1. Typically the kit will contain both a forward and a reverse primer for a single gene. 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 or other chemical moieties 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. Other moieties may include detectable labels or specific binding moieties, such as biotin. 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 used to assess the presence of the marker.

In certain embodiments of the invention, a set of oligonucleotide primers is provided, each of which is designed to hybridize to sodium bisulfite-modified nucleic acids of one particular gene, e.g., LGR5, APCDD1, AXIN2, DKK1, or ASCL2. In certain embodiments of the invention, a single kit contains multiple oligonucleotide primers, such that methylation of CpG sites in more than one gene can be analyzed. For example, a kit may contain oligonucleotide primers designed to hybridize to sodium bisulfite-modified nucleic acids of a gene of interest as well as oligonucleotide primers designed to hybridize to sodium bisulfite -modified nucleic acids of a gene of interest.

In certain embodiments of the invention, the kit provides reagents and/or instructions for sodium bisulfite sequencing analysis of LGR5, APCDD1, AXIN2, DKK1, and/or ASCL2 nucleic acids. Such reagents may include any of, for example, sodium bisulfite, buffers and solutions, spin columns for separation and/or purification of nucleic acids, reaction tubes, etc.

In certain embodiments of the invention, kits further comprise a control or reference sample that comprises DNA that is mostly or entirely demethylated. The DNA in the reference sample may be demethylated globally (e.g., at all or most CpG sites), in particular genes or regions of genes, or at one or more particular CpG sites. In certain embodiments of the invention, kits further comprise a control or reference sample that comprises DNA that is mostly or entirely methylated. The DNA in the reference sample may be methylated globally (e.g., at all or most CpG sites), in particular genes or regions of genes, or at one or more particular CpG sites.

The above disclosure generally describes the present invention. Unless defined otherwise, all technical and scientific terms used herein have the same meaning as those commonly understood by one of ordinary skill in the art to which this invention belongs. 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.

EXAMPLE SECTION

Previously the inventors have shown that colon-cancer stem cells can be identified in primary human colorectal cancer using Wnt signaling intensity levels and can be isolated by employing a Wnt reporter construct (TOP-GFP, Figure 1A) (Vermeulen et al., 2010). Gene expression profiling of the TOP-GFPhigh human colon-cancer stem cells indicated high expression of the stem cell marker LGR5, as well as several other Wnt targets (APCDD1, LEF1), while typical intestinal differentiation markers (e.g. MUC2 and FABP1) displayed low expression (Figure 1A). Based on two primary isolated spheroid cultures the inventors generated a colon-cancer stem cell signature comprising of 187 genes most differentially expressed between the cancer stem cells and the more differentiated cells. This profile was subsequently validated in several independent freshly isolated colorectal cancers (Figure 5A).

Crucially, by employing gene set enrichment analysis, they found that this colon cancer stem cell gene expression signature was intimately associated with disease recurrence in a set of 90 stage II colorectal cancer patients that underwent intentionally curative surgery at the institute (AMC-AJCCII-90, see Table 2 for patient characteristics), a finding they confirmed in an unrelated, publically available dataset (Figures 1B and 5B). Similar results were obtained using two independent intestinal stem cell profiles that have previously been shown to relate to disease recurrence lending strength to the validity of both the cancer stem cell signature as well as to the patient set (Figures 1B and 5B) (Merlos-Suarez et al., 2011). Using a simple rank-sum approach to stratify the AMC-AJCCII-90 patients in two groups the inventors further established the prognostic power of the cancer stem cell profile and revealed that especially early relapses in stage II colorectal cancer patients were characterized by a strong resemblance to this signature (Figure 1C).

Characteristics	90 patients Stage II
Male sex (%)	*42(46.7%)*
Median at age operation (range)	*73.4(34.6-95.1)*
T-stage	
-T3	*81(90%)*
-T4	*9(10%)*
Location	
- left	*38(42.2%)*
- right	*52(57.8%)*
Differentiation grade	
- Moderate/well	*61(67.7%)*
- Poor	*29(32.2%)*
3yr disease free survival (%)	*71(79.7%)*
median follow-up	*3.26(0.14-9.86)yr*

Table 2: patient characteristics of the AMC-AJCC-II database

In an attempt to identify genes that are most predictive regarding tumor relapse and metastasis formation the inventors employed cluster analysis. Unsupervised K-means clustering of the AMC-AJCCII-90 set using the TOP-GFPhigh/cancer stem cell-derived gene expression profile resulted in two distinct patient groups with a significant difference in relapse free survival as the Kaplan-Meier curve illustrates (Figures 1D and 1E). Gene tree analysis revealed that segregation of these two clusters is accompanied by generation of two major subgroups of genes. The majority of genes were up-regulated in the patient cluster (black) that was associated with poor prognosis. However, a clearly distinct subset of genes, at the bottom region of the gene-tree, inversely correlated with disease relapse (Figure 1D). Surprisingly, this gene cluster contained many Wnt target genes of which the expression is intimately linked to intestinal stem cells and cancer stem cells. (For a list of genes and their differential expression between the cluster see Tables 3 and 4). Similar results were obtained by employing different stem cell signatures, including the LGR5- and EphB2 intestinal stem cell signatures that have been described previously (Figures 5C and D) (Merlos-Suarez et al., 2011). In all cases high expression of Wnt target genes was associated with the favorable prognosis cluster (Figure 1D, 5C and 5D). The inventors confirmed the expression level differences between the two groups by qPCR (Figure 5E) and found that the expression level differences were not related to oncogenic mutations present within the different patients (Figure 6A). Multivariate Cox regression analysis indicated the TOP-GFPhigh/cancer stem cell-derived gene expression profile to be an independent prognostic factor that was much more predictive than the presence of different mutations (Figure 6A). Strikingly, the inventors also observed that the

expression of the Wnt target genes was not a simple reflection of patient-to-patient variation in Wnt signalling activity as measured by nuclear localized β-catenin levels (Figures 6B and 6C).

Wnt-low group			Wnt-high group		
HUGO	**Fold**	**pval**	**HUGO**	**Fold**	**pval**
CELP	4,79	7,75E-19	VMO1	-1,78	9,57E-09
DNASE1	4,34	5,29E-07	HOTAIR	-1,79	0,00173
UMODL1	4,02	4,19E-06	DLX1	-1,79	0,01
CEL	3,51	3,3E-15	TREML3	-1,8	2,74E-05
PLA2G12B	3,27	2E-15	DNAH2	-1,8	1,61E-07
MEGF11	3,02	0,00667	WDR49	-1,81	0,06
WNT11	2,88	8,74E-10	ARX	-1,82	3,45E-05
FLJ12334	2,82	1,12E-12	C5orf46	-1,82	3,97E-08
KRT40	2,76	1,15E-06	DMRT3	-1,83	0,00473
SLC13A3	2,71	4,45E-10	CLEC4G	-1,84	0,000471
TNMD	2,65	6,39E-07	CCDC148	-1,84	0,000584
KLHL34	2,64	1,49E-06	SOX8	-1,84	0,00114
TTTY2	2,6	0,00306	KCNA1	-1,87	2,41E-05
ABCC2	2,58	4,17E-08	C19orf51	-1,88	8E-11
GPR143	2,55	4,53E-15	ASB4	-1,88	0,025
DRD1	2,53	0,024	PAEP	-1,91	0,000214
SLC1A7	2,45	1,41E-07	OSR1	-1,92	2,11E-05
LY6G6D	2,4	1,79E-22	GAD1	-1,93	7,83E-07
WIF1	2,33	3,07E-05	MBL2	-2,04	0,00463
OGDHL	2,32	8,71E-05	SERPINB2	-2,04	2,83E-10
SULT2A1	2,31	0,000862	C4orf47	-2,12	3,24E-10
SLC22A11	2,28	0,000321	TRIM7	-2,12	1,06E-16
TH	2,27	0,00571	LOC283432	-2,12	0,00894
ACSL6	2,26	4,91E-12	FSTL5	-2,14	0,026
SMR3A	2,25	0,016	GJB6	-2,15	0,066
CDHR1	2,2	5,13E-13	SOST	-2,16	0,00222
FABP6	2,18	1,03E-12	BNC1	-2,17	0,00139
KNG1	2,17	0,000327	ANKRD45	-2,17	0,000797
XPNPEP2	2,17	1,78E-06	C9orf66	-2,17	3,11E-05
IL22	2,14	5,13E-06	HEATR7B1	-2,2	0,00304
SLC38A4	2,14	1,71E-09	WDR63	-2,2	3,91E-05
NCKAP5	2,13	5,27E-08	USF1	-2,21	1,04E-11
FREM2	2,08	4,44E-14	COL6A6	-2,22	0,000438
ZNF663	2,08	0,00263	CNGA3	-2,3	0,156
TMEM132C	2,08	0,00156	CALB1	-2,37	0,000011
ADAMTSL2	2,05	0,014	SLC26A9	-2,49	1,07E-13
DLX6-AS1	2,04	0,000674	SEMG1	-2,5	2,59E-07
MAGEA11	2,02	0,02	CHST6	-2,56	6,51E-05
LOC283859	2,01	6,02E-13	GJB5	-2,57	9,78E-10
PAH	2,01	5,94E-09	MTA2	-2,68	7,86E-07
C3orf32	1,95	4,3E-12	MUC5AC	-2,73	4,02E-07
EYA1	1,92	0,000959	ZIC2	-2,75	3,09E-12
PNLIPRP2	1,91	2,94E-06	TNNT1	-2,89	3,56E-13
C1orf99	1,91	0,052	TTC29	-3	0,0058
MYT1	1,88	0,0021	NXF3	-3,01	1,42E-10
KRT23	1,87	7,98E-11	SPRR1A	-3,06	0,0024
SERP2	1,87	1,36E-06	MSMB	-3,35	0,00128
SERPINA6	1,86	8,27E-05	ANXA10	-3,66	3,01E-09
FBP2	1,86	0,142			
TRIM54	1,86	0,00228			
LOC200772	1,85	0,0061			
C19orf30	1,85	0,022			
SHISA6	1,84	3,05E-05			
IZUMO2	1,84	4,09E-06			
NKD1	1,81	1,59E-10			
TNNC2	1,81	5,58E-15			
SYN3	1,81	8,53E-06			
LY6G6E	1,81	0,000138			
CYP4F8	1,78	0,0044			
C1orf161	1,78	3,1E-09			

Table 3: Complete set of differentially expressed genes between Wnt-high and low groups.

	Wnt high			Wnt low	
HUGO	Fold	pval	HUGO	Fold	pval
CELP	4,79	7,75E-19	CEACAM6	1,09	9,35E-05
CEL	3,51	3,3E-15	MTMR11	1,09	0,000249
ACSL6	2,26	4,91E-12	CEACAM5	1,09	6,22E-06
LOC283859	2,01	6,02E-13	ZNF503	1,09	0,001
SERP2	1,87	1,36E-06	SOX2	1,08	0,633
NKD1	1,81	1,59E-10	SORBS2	1,08	0,04
MYH7B	1,61	0,032	UBE2L6	-1,06	0,000879
HUNK	1,58	1,24E-22	SLC7A7	-1,06	0,000165
DEFA6	1,52	0,000207	HMOX1	-1,07	0,000596
CHP2	1,51	0,00274	FMOD	-1,07	0,000876
NOTUM	1,49	0,000231	DKK4	-1,07	0,714
DEFA5	1,47	0,000425	FOSL1	-1,08	0,00884
AREGB	1,46	5,17E-12	LMCD1	-1,08	5,29E-05
WNT5B	1,45	0,000116	RGS10	-1,08	7,47E-06
DPEP1	1,38	1,21E-09	SULF1	-1,08	0,00224
FABP1	1,37	1,63E-07	MLLT11	-1,08	0,00273
LRP4	1,33	3,52E-10	F13A1	-1,09	0,019
APCDD1	1,32	2,82E-14	TTC9	-1,09	0,404
CAB39L	1,31	2,16E-11	TGFBR3	-1,09	0,000624
HMGCS2	1,3	0,000721	TRIM29	-1,09	0,019
KCNA2	1,28	0,029	SPINK4	-1,09	0,25
NRXN3	1,28	4,34E-05	MX1	-1,1	0,00348
TFCP2L1	1,26	1,17E-09	ROR1	-1,1	0,015
CAPN6	1,26	0,101	IQGAP2	-1,11	1,38E-06
C1orf105	1,25	4,15E-05	FZD2	-1,11	4,11E-05
ADH6	1,25	0,000931	VSNL1	-1,12	0,000713
REEP1	1,25	1,18E-07	AGPAT9	-1,12	0,000341
SPINK1	1,24	7,89E-10	CXCR4	-1,12	3,95E-08
KRT20	1,24	3,63E-05	CD55	-1,13	8,89E-12
PLA2R1	1,21	0,00132	GLT25D2	-1,13	0,00379
SLC2A12	1,21	3,84E-05	SLC1A1	-1,14	2,71E-07
RBP1	1,2	1,02E-05	GCNT3	-1,14	0,000224
CST1	1,19	0,011	CAMK2N2	-1,14	0,00263
PROX1	1,19	1,44E-05	C8orf84	-1,15	0,012
PLAGL1	1,17	6,73E-13	PHACTR1	-1,16	0,00977
APOBEC1	1,17	0,000434	ZBTB7C	-1,17	0,022
CYP2S1	1,16	4E-07	ETV1	-1,19	4,05E-06
NFE2	1,16	0,406	LEMD1	-1,22	0,00983
PTCH1	1,15	2,19E-12	MUC2	-1,22	0,000557
ARL14	1,13	0,00414	IFI6	-1,24	5,16E-08
NEURL1B	1,12	1,92E-09	MAOB	-1,25	0,00166
GPA33	1,12	0,000823	FRAS1	-1,27	1,49E-05
IRX5	1,12	0,193	SDR16C5	-1,29	1,91E-05
CHST13	1,12	0,00419	REG4	-1,34	5,3E-07
LGR5	1,1	0,025	IRX3	-1,35	0,00118
CCL24	1,1	0,102	LOC100128079	-1,59	0,016
MPP1	1,1	4,66E-06	DLX1	-1,79	0,01
CACHD1	1,1	0,022	ASB4	-1,88	0,025
			GAD1	-1,93	7,83E-07

Table 4: Genes differentially expressed between Wnt high and low group present in the colon stem cell signature

Intrigued by this counter-intuitive finding, which indicates that high Wnt target gene expression is rather linked to favorable and not to poor prognosis, the inventors repeated this analysis with a more defined set of Wnt target genes, previously identified by over-expression of dominant negative TCF4 (dnTCF4) in colorectal cancer cell lines (Van der Flier et al., 2007 van de Wetering et al., 2002). The clear majority of genes in this dnTCF4 signature are also markedly lower expressed in the poor prognosis patient cluster (Figures 1F and 1G). Surprisingly, even single Wnt target genes, selected from the group consisting of LGR5, DKK1, ASCL2, AXIN2 and APCDD1 can identify poor prognosis patients based on their low expression levels both in our patient set as well as in publically available datasets (Figure 6D).

The finding that high expression of a gene selected from the group consisting of LGR5, DKK1, ASCL2, AXIN2 and APCDD1 is associated with good prognosis in colorectal cancer immediately challenges the conventional interpretation as to why (cancer) stem cell-associated profiles define poor prognosis in cancer. It is generally believed that association with a cancer stem cell profile reflects the number of cancer stem cell-like cells in the malignancy. However, when the inventors defined the fraction of cells positive for CD133 using FACS staining of several freshly isolated colon cancer specimens, which so far is the best studied and validated means to identify colon-cancer stem cells, they could not correlate the number of cancer stem cells to the overall expression of cancer stem cell-associated Wnt target genes within these tumors (Figure 6E). In addition, also the fraction of colorectal cancer cells positive for nuclear β-catenin, which has been used before as trait to identify colon-cancer stem cells, does not correlate significantly with Wnt target gene expression in the patient dataset used by the inventors (Figure 6F). These findings indicate that cancer stem cell numbers in colorectal cancer are not causal determinants in the patient stratification obtained with the cancer stem cell-associated expression signature. Additionally, the lack of correlation between nuclear β-catenin levels and Wnt target gene expression indicates that additional regulatory mechanisms are in place to regulate Wnt target gene expression.

Using gene expression analysis, the inventors determined the expression of APCDD1, LGR5, DKK1, ASCL2 and AXIN2 at multiple stages during the adenoma-carcinoma sequence. Comparison of normal tissue with adenoma tissue revealed a marked increase in expression of most Wnt target genes (Figure 2A). Strikingly, evaluation of expression of APCDD1, LGR5, DKK1, ASCL2 and AXIN2 in colorectal cancer samples indicated a downregulation in the majority of patients compared to the adenoma stage (Figure 2A). This was also observed in an independent dataset containing normal, adenoma and carcinoma samples (Figures 7E and 7F). The relevance of this suppressed expression of APCDD1, LGR5, DKK1, ASCL2 and AXIN2 was immediately evident as patients that developed a tumor relapse displayed the lowest expression levels (Figure 2A, open triangles). This confirms the findings that a low gene expression of APCDD1, LGR5, DKK1, ASCL2 and AXIN2 is related to poor prognosis. Under expression of APCDD1, LGR5,

DKK1, ASCL2 and AXIN2 during tumor progression was not dependent on additional mutations or the level of Wnt signaling activity (Figure 6).

Treatment of colorectal cancer cells with the demethylating agent 5-Azacytidine (5-Aza) resulted in marked upregulation of LGR5, APCDD1, DKK1, AXIN2 and ASCL2 specifically in those cell lines where methylation was evident (Figure 2C). For instance, APCDD1 expression was enhanced by 5-Aza in the lines where APCDD1 promoter methylation was clearly detectable.

Intriguingly, the treatment of colorectal cancer cell lines with a demethylating agent resulted in markedly decreased clonogenicity of these lines (Figure 3A). Importantly, this was a general observation as also primary isolated colon-cancer stem cell cultures treated with 5-Aza demonstrated significantly lower clonogenicity as determined by limiting dilution analysis, suggesting that the fraction of cancer stem cells in these cultures decreased (Figure 3B). Also in an in vivo model system in which primary cancer stem cell-induced xenograft tumors were growing subcutaneously the inventors observed that 5-Aza treatment resulted in markedly suppressed tumor growth (Figure 3C). Importantly, analysis of these 5-Aza treated xenografts confirmed the efficacy of 5-Aza on re-expression of the Wnt target genes in vivo (Figure 3D), validating the methylation-dependent regulation of Wnt target genes also in this in vivo model. To establish the consequences of re-expression of the Wnt target gene biomarkers, the inventors determined the effect of 5-Aza on Wnt activity levels. In several lines 5-Aza-mediated expression of genes previously suppressed by methylation resulted in decreased Wnt activity (Figure 3E). This already indicates an effect of 5-Aza on the activity of the Wnt pathway, but a more direct proof can be deduced from the observation that overexpression of either AXIN2 or APCDD1, both methylated in colorectal cancer, was sufficient to decrease Wnt signaling levels (Figure 3F). This was not only observed in colorectal cancer cell lines, but was confirmed in a primary cancer stem cell culture as well (Figure 3F, right bars, Co100).

Wnt target gene methylation predicts prognosis in colorectal cancer

In order to determine whether also in patient material suppression of a stem cell-associated Wnt expression program is dependent on methylation, the inventors determined the relative methylation levels of the Wnt target genes in the AMC-AJCCII-90 patient set by either methylation-specific PCR or bisulphate sequencing (Figure 4). Intriguingly, also in the primary human tumors low Wnt target gene expression was associated with increased methylation of the promoter regions of these genes (Figure 4A). More importantly, the inventors found that methylation levels of already a small subset of these cancer stem cell-associated Wnt target genes resulted in a highly predictive association with disease recurrence and metastasis as determined by unsupervised cluster analysis using the relative methylation levels (Figures 4B and 4C) These findings further support the invention that methylation-dependent tuning of the Wnt target gene biomarkers is related to disease progression and increased risk for recurrent disease.

Gene expression array and patients cohorts

Two different colorectal cancer patient series were used for this study. The first one is consisting of colorectal cancer patient material collected in the Academic Medical Center (AMC) in Amsterdam, The Netherlands. We focused on a set of 90 AJCC stage II patients that underwent intentionally curative surgery in the years 1997-2006 (AMC-AJCCII-90). Patient samples were collected in accordance with the guidelines of the Medical Ethical Committee of the AMC and national legislation. Extensive medical records are kept from these patients and long-term clinical follow-up is available for the large majority. Both paraffin embedded as well as fresh frozen tissue is available from all these patients for analysis. To generate expression profiles RNA was extracted using Trizol in accordance with the manufacturer's protocol (Invitrogen, Breda, NL). RNA concentration was determined with NanoDrop ND-1000, and quality was determined using the RNA 6000 Nano assay on the Agilent 2100 Bioanalyzer (Agilent Technologies, Amstelveen, NL). Affymetrix microarray analysis, fragmentation of RNA, labelling, hybridization to Human Genome U133 Plus 2.0 microarrays, and scanning were performed in accordance with the manufacturer's protocol (Affymetrix, Santa Clara, CA). Microarray data can be viewed online (http://www.ncbi.nlm.nih.gov/geo/index.html) under GEO accession number (to be determined). The second patients set is composed of two merged cohorts that form a metacohort of 345 colorectal cancer patients and was described elsewhere (Merloz-suarez et al.). In addition, a separate panel of normal and adenoma fresh-frozen tissues was obtained and RNA extracted using Trizol.

Mutation analysis, RNA isolation, qPCR, rtPCR, methylation analysis and primer sequences

P53 mutations were analyzed by sequencing with specific sets of primers.

The K-RAS and the B-RAF mutations were determined in tumor-derived cDNA by PCR and subsequent sequencing. All sequencing reactions were performed using the Big Dye Terminator (BDT, Applied Biosystems, Nieuwerkerk a/d IJssel, NL). For MSI analysis we used the MSI Analysis System Version 1.2 (Promega, Leiden, NL) according to the manufacturer's instructions.

Primer	*Sequence*
Seq ID No. 1 (P53-exon1, Forward)	GCTTTCCACGACGGTGACA
Seq ID No. 2 (P53-exon1, Reverse)	TTGTTGAGGGCAGGGGAGTA
Seq ID No. 3 (P53-P33, Forward)	TGTCATCTTCTGTCCCTTCCC
Seq ID No. 4 (P53-P33, Reverse)	GATGGTGGTACAGTCAGAGC
Seq ID No. 5 (P53-P31, Forward)	TTGCGTGTGGAGTATTTGGA
Seq ID No. 6 (P53-P31, Reverse)	GCAAGCAAGGGTTCAAAGACC
Seq ID No. 7 (K-RAS, Forward)	CTGTGACATGTTCTAATATAGTCA
Seq ID No. 8 (K-RAS, Reverse)	GAATGGTCCTGCACCAGTAA

Seq ID No. 9 (B-RAF, Forward)	TGATTTTTGTGAATACTGGGAAC
Seq ID No. 10 (B-RAF, Reverse)	TGCTTGCTCTGATAGGAAAATG

To generate expression profiles or analysis RNA expression by rtPCR or qPCR, RNA was extracted using Trizol in accordance with the manufacturer's protocol (Invitrogen, Breda, NL). RNA concentration was analysed on a NanoDrop ND-1000. For microarray RNA quality was determined using the RNA 6000 Nano assay on the Agilent 2100 Bioanalyzer (Agilent Technologies, Amstelveen, NL). Affymetrix microarray analysis, fragmentation of RNA, labelling, hybridization to Human Genome U133 Plus 2.0 microarrays, and scanning were performed in accordance with the manufacturer's protocol (Affymetrix, Santa Clara, CA). Microarray data can be viewed online (http://www.ncbi.nlm.nih.gov/geo/index.html) under GEO accession number (to be determined).

For real time PCR, RNA was extracted using Trizol in accordance with the manufacturer's protocol (Invitrogen, Breda, NL). RNA concentration was determined with Real time RT–PCR was performed with SYBR green (Roche) in accordance with the manufacturer's instructions on a Light Cycler 480 (Roche). Primer sequences are described elsewhere (Vermeulen et al., 2010) or listed below:

Primer		*Sequence*
Seq ID No. 11	β-actin, Forward	CAGAAGGATTCCTATGTGGGCGA
Seq ID No. 12	β-actin, Reverse	TTCTCCATGTCGTCCCAGTTGGT
Seq ID No. 13	LGR5, Forward	AAT CCC CTG CCC AGT CTC
Seq ID No. 14	LGR5, Reverse	CCC TTG GGA ATG TAT GTC AGA
Seq ID No. 15	APCDD1, Forward	CATCCAGACAGCAGGTCTCA
Seq ID No. 16	APCDD1, Reverse	GGGCCTGACCTTACTTCACA
Seq ID No. 17	AXIN, Forward	CTCCTTATCGTGTGGGCAGT
Seq ID No. 18	AXIN2, Reverse	CTTCATCCTCTCGGATCTGC
Seq ID No. 19	ASCL2, Forward	GTGAAGCTGGTGAACTTGGGC
Seq ID No. 20	ASCL2, Reverse	CAGCGTCTCCACCTTGCTCA
Seq ID No. 21	Oct4a, Forward	ACGACCATCTGCCGCTTTG
Seq ID No. 22	Oct4a, Reverse	GCTTCCTCCACCCACTTCTG
Seq ID No. 23	SOX2, Forward	AAATGGGAGGGGTGCAAAAGAGGA
Seq ID No. 24	SOX2, Reverse	CAGCTGTCATTTGCTGTGGGTGATG
Seq ID No. 25	NANOG, Forward	GCCGAAGAATAGCAATGGTG
Seq ID No. 26	NANOG, Reverse	TGGTGGTAGGAAGAGTAGAGG
Seq ID No. 27	DKK1, Forward	GCCCCGGGAATTACTGCAAAAATG
Seq ID No. 28	DKK1, Reverse	CCGGAGACAAACAGAACCTTCTTGTC

Seq ID No. 29	KLF4, Forward	CCAATTACCCATCCTTCCTG
Seq ID No. 30	KLF4, Reverse	CGATCGTCTTCCCCTCTTTG

For methylation-specific PCR genomic DNA from colorectal cancer cell lines and primary colorectal cancer tissue was extracted and 500ng of gDNA was used for bi-sulfite conversion with the Epitect bisulfite kit (Qiagen, Venlo, NL) according to the manufacturer's instructions. Methylation-specific PCR was carried out using primer sets specific for either methylated or unmethylated CpG sites. For primary colorectal cancer, a nested PCR was used to enhance the product prior to quantitative MSP with SYBR green using methylation-insensitive primer sets. Plasma blood lymphocyte (PBL) gDNA was treated with Methylase (M.Sss1, Bioke, Leiden, NL). Untreated and treated samples were bi-sulfite converted and used as negative and positive controls respectively. Primer sequences can be found in supplemental material and methods.

Primer Sequence

Seq ID No. 31	APCDD1_MSP_M, Forward	GTGAGTTTTCGAGGGTTATTCGAGC
Seq ID No. 32	APCDD1_MSP_M, Reverse	CCTCTCCCAAAACTAAAACGACAAACG
Seq ID No. 33	APCDD1_MSP_U, Forward	ATGGTGAGTTTTTGAGGGTTATTTGAGTG
Seq ID No. 34	APCDD1_MSP_U, Reverse	CCCTCTCCCAAAACTAAAACAACAAACA
Seq ID No. 35	LGR5_MSP_M, Forward	TGGATTCGTAGAGAAATGGTTTC
Seq ID No. 36	LGR5_MSP_M, Reverse	ATACCTACATTTACAAAAATTCGCC
Seq ID No. 37	LGR5_MSP_U, Forward	TTTGGATTTGTAGAGAAATGGTTTT
Seq ID No. 38	LGR5_MSP_U, Reverse	ATACCTACATTTACAAAAATTCACC
Seq ID No. 39	Apcdd1, Forward nested	AGGTTTTAGAGTAGGATTGGAAATGT
Seq ID No. 40	Apcdd1, Reverse nested	ACCCCCTCTCCCAAAACTAA
Seq ID No. 41	AXIN2, Forward nested	TTGTATATAGTTTAGYGGTTGGG
Seq ID No. 42	AXIN2, Reverse nested	AAATCTAAACTCCCTACACACTT
Seq ID No. 43	DKK1, Forward nested	TTYGGTTTTGTTGTTTTTTTTTTAAGGGG
Seq ID No. 44	DKK1, Reverse nested	AACCRTCACTTTACAAACCTAAATCCC
Seq ID No. 45	DKK1_MSP_M, Forward	GTCGGAATGTTTCGGTTCGC
Seq ID No. 46	DKK1_MSP_M Reverse	CTAAATCCCCACGAAACCGTACCG
Seq ID No. 47	DKK1_MSP_U, Forward	GGGGTTGGAATGTTTTGGGTTTGT
Seq ID No. 48	DKK1_MSP_U Reverse	ACCTAAATCCCCACAAAACCATACCA
Seq ID No. 49	ASCL2, Forward nested	ATTCCTCTACCTACACCTTCCTAC
Seq ID No. 50	ASCL2, Reverse nested	GTTTGGAAGTTTAAGTTTATTAGTTTTA
Seq ID No. 51	ASCL2_MSP_M, Forward	GCGATTGGTTAAACGGGTGGTTTAC
Seq ID No. 52	ASCL2_MSP_M Reverse	AACGCGACCCTAACGACAACACG
Seq ID No. 53	ASCL2_MSP_U, Forward	TGATTGGTTAAATGGGTGGTTTATGA

Seq ID No. 54	ASCL2_MSP_U Reverse	AACACAACCCTAACAACAACACAAC
Seq ID No. 55	LGR5 Forward	CTGCCTGCAATCTACAAGGT
Seq ID No. 56	LGR5 Reverse	CCCTTGGGAATGTATGTCAGA
Seq ID No. 57	LGR5 Forward nested	ACTCTAAATCACAAATATCCACTCC
Seq ID No. 58	LGR5 Reverse nested	TGGTTAGGTTTGTTAGGGTTTTTAG

Nomenclature of the above listed primers is as follows: "MSP" stands for Methylation specific PCR, "M" stands for methylated and "U" for unmethylated.

Generation of the colon-cancer stem cell signature

Gene expression levels were measured on the 10% highest and lowest TOP-GFP fraction of cells in two independent colon cancer stem cell cultures (Vermeulen et al., 2010) using Human Genome U133 Plus 2.0 microarrays. The cancer stem cell signature was derived by selecting genes of raw data value >30 that were most differentially regulated between TOP-GFPHigh and TOP-GFPLow by at least 2-logfold change from each line and with a p-value < 0.05. Both orientations were taken into account to generate a list of 187 genes (134 up-regulated genes and 53 genes downregulated in TOP-GFPHigh compared to TOP-GFPLow). The intestinal stem cell-EphB2, Lgr5, and dnTCF4 signature were described elsewhere ((Merlos-Suarez et al., 2011; van de Wetering, 2002) respectively).

Gene set enrichment analysis

Gene set enrichment analysis was performed as previously described (Subramanian et al., 2005). Genes are ranked according to their association with a given phenotype. Genes at the top of the rank associates positively with the phenotype while genes at the bottom of the rank associates negatively. The gene set enrichment analysis determine whether genes present in a defined signature are either randomly distributed throughout the ranks (no association of the signature with the phenotype), present at the top of the rank (positively associated with the phenotype) or present at the bottom of the rank (negatively associated with the phenotype). The occurrence of disease recurrence was used as a variable. The output of the gene set enrichment analysis comprises the Enrichment score (ES), the normalized enrichment score (NES) and the False discovery rate (FDR). Clustering and survival analysis Unsupervised K-means cluster analysis as well as Kaplan-Meier survival curves were generated in the different expression data sets with the different gene signatures using the software package R2 (https://r2.amc.nl), a web based microarray analysis application developed by JK (manuscript in preparation). For single genes survival prediction, the median expression value of each gene was used as a cut-off to generate two groups of 45 patients having either a low or high expression relative expression. P-value was calculated using the log-rank test.

Prediction power of signature and multivariate analysis

To evaluate the predictive power of the cancer stem cell signature, the inventors selected the 134 up-regulated genes. Every individual gene was ranked according to their expression in each patient. The rank score for all the genes per patient was summed to define the rank for each patient. A high expression of a gene is translated into a low rank score. Patients that have an overall high expression of the genes in the signature have an overall low rank score and therefore are highly associated with the cancer stem cell profile. A Kaplan-Meier survival curve was generated to plot the relapse free survival of patients having a high (n=30) versus low (n=60) correlation with the cancer stem cell profile. For multivariate analysis, the Cox proportional hazard model was used. All p-values are two sided. Statistical analysis was performed in SPSS.

Cell culture and generation of the colon-CSC signature

The generation and culture of colon CSCs has been previously described (Vermeulen et al., 2010). Colorectal cancer cell lines (DLD-1, HCT-15, HCT-116, SW620, SW837 and Colo205) were purchased at LGC Standards GmbH, Germany. All lines were maintained in DMEM (supplemented with 10% FCS/1% glutamine) except Colo205 which was cultured in RPMI-1640 (10% FCS/1% glutamine). Gene expression levels were measured on the 10% highest and lowest TOP-GFP fraction of cells in two independent colon CSC cultures (Vermeulen et al., 2010) using Human Genome U133 Plus 2.0 microarrays. The CSC signature was derived by selecting genes of raw data value >30 that were most differentially regulated between TOP-GFPHigh and TOP-GFPLow by at least 2-logfold change from each line and with a p-value < 0.05. Both orientations were taken into account to generate a list of 187 genes (134 up-regulated genes and 53 genes down-regulated in TOP-GFPHigh compared to TOP-GFPLow), that was subsequently validated on several independent CSC cultures (GEO accession number of the array data is). The ISC-EphB2, Lgr5, and dnTCF4 signature were described elsewhere (Merlos-Suarez et al., 2011; van de Wetering, 2002).

REFERENCES

Barker,N., van Es,J.H., Kuipers,J., Kujala,P., van den,B.M., Cozijnsen,M.,Haegebarth,A., Korving,J., Begthel,H., Peters,P.J., and Clevers,H. (2007). Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature 449, 1003-1007.

Bertucci,F., Cervera,N., and Birnbaum,D. (2007). A gene signature in breast cancer. N. Engl. J. Med. 356, 1887-1888.

Edge,S.B. and Compton,C.C. (2010). The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Ann. Surg. Oncol. 17, 1471-1474.

Fearon,E.R. and Vogelstein,B. (1990). A genetic model for colorectal tumorigenesis. Cell 61, 759-767.

Koinuma,K., Yamashita,Y., Liu,W., Hatanaka,H., Kurashina,K., Wada,T., Takada,S., Kaneda,R., Choi,Y.L., Fujiwara,S.I., Miyakura,Y., Nagai,H., and Mano,H. (2006). Epigenetic silencing of AXIN2 in colorectal carcinoma with microsatellite instability. Oncogene 25, 139-146.

Liu,R., Wang,X., Chen,G.Y., Dalerba,P., Gurney,A., Hoey,T., Sherlock,G., Lewicki,J., Shedden,K., and Clarke,M.F. (2007). The prognostic role of a gene signature from tumorigenic breast-cancer cells. N. Engl. J. Med. 356, 217-226.

Medema,J.P. and Vermeulen,L. (2010) Microenvironmental Regulation of Stem Cells in Intestinal Homeostasis and Cancer (Nature, in press)

Merlos-Suarez,A., Barriga,F.M., Jung,P., Iglesias,M., Cespedes,M.V., Rossell,D., Sevillano,M., Hernando-Momblona,X., da Silva-Diz,V., Munoz,P., Clevers,H., Sancho,E., Mangues,R., and Batlle,E. (2011). The intestinal stem cell signature identifies colorectal cancer stem cells and predicts disease relapse. Cell Stem Cell 8, 511-524.

Morin,P.J., Sparks,A.B., Korinek,V., Barker,N., Clevers,H., Vogelstein,B., and Kinzler,K.W.. (1997). Activation of beta-catenin-Tcf signaling in colon cancer by mutations in beta-catenin or APC. Science 275:1787-90.

O'Brien,C.A., Pollett,A., Gallinger,S., and Dick,J.E. (2007). A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. Nature 445, 106-110.

Rai,K., Sarkar,S., Broadbent,T.J., Voas,M., Grossmann,K.F., Nadauld,L.D., Dehghanizadeh,S., Hagos,F.T., Li,Y., Toth,R.K., Chidester,S., Bahr,T.M., Johnson,W.E., Sklow,B., Burt,R., Cairns,B.R., and Jones,D.A. (2010). DNA demethylase activity maintains intestinal cells in an undifferentiated state following loss of APC. Cell 142, 930-942.

Ricci-Vitiani,L., Lombardi,D.G., Pilozzi,E., Biffoni,M., Todaro,M., Peschle,C., and De,M.R. (2007). Identification and expansion of human colon-cancer-initiating cells. Nature 445, 111-115.

Shimomura,Y., Agalliu,D., Vonica,A., Luria,V., Wajid,M., Baumer,A., Belli,S., Petukhova,L., Schinzel,A., Brivanlou,A.H., Barres,B.A., and Christiano,A.M. (2010). APCDD1 is a novel Wnt inhibitor mutated in hereditary hypotrichosis simplex. Nature 464, 1043-1047.

Suzuki,H., Watkins,D.N., Jair,K.W., Schuebel,K.E., Markowitz,S.D., Chen,W.D., Pretlow,T.P., Yang,B., Akiyama,Y., van,E.M., Toyota,M., Tokino,T., Hinoda,Y., Imai,K., Herman,J.G., and Baylin,S.B. (2004). Epigenetic inactivation of SFRP genes allows constitutive WNT signaling in colorectal cancer. Nat. Genet. 36, 417- 422.

Todaro,M., Alea,M.P., Di Stefano,A.B., Cammareri,P., Vermeulen,L., Iovino,F., Tripodo,C., Russo,A., Gulotta,G., Medema,J.P., and Stassi,G. (2007). Colon cancer stem cells dictate tumor growth and resist cell death by production of interleukin-4. Cell Stem Cell 1, 389-402.

Toyota,M., Ahuja,N., Ohe-Toyota,M., Herman,J.G., Baylin,S.B., and Issa,J.P. (1999). CpG island methylator phenotype in colorectal cancer. Proc. Natl. Acad. Sci. U. S. A 96, 8681-8686.

van de Wetering,M., Sancho,E., Verweij,C., de Lau,W., Oving,I., Hurlstone,A., van der Horn,K., Batlle,E., Coudreuse,D., Haramis,A.P., Tjon-Pon-Fong,M., Moerer,P., van den Born,M., Soete,G., Pals,S., Eilers,M., Medema,R., and Clevers,H. (2002). The beta-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. Cell 111, 241-50.

van der Flier,L.G., van Gijn,M.E., Hatzis,P., Kujala,P., Haegebarth,A., Stange,D.E., Begthel,H., van den,B.M., Guryev,V., Oving,I., van Es,J.H., Barker,N., Peters,P.J., van de,W.M., and Clevers,H. (2009). Transcription factor achaete scute-like 2 controls intestinal stem cell fate. Cell 136, 903-912.

Vermeulen,L., De Sousa E Melo, van der Heijden,M., Cameron,K., de Jong,J.H., Borovski,T., Tuynman,J.B., Todaro,M., Merz,C., Rodermond,H., Sprick,M.R., Kemper,K., Richel,D.J., Stassi,G., and Medema,J.P. (2010). Wnt activity defines colon cancer stem cells and is regulated by the microenvironment. Nat. Cell Biol. 12, 468-476.

Vermeulen,L., Todaro,M., de Sousa,M.F., Sprick,M.R., Kemper,K., Perez,A.M., Richel,D.J., Stassi,G., and Medema,J.P. (2008). Single-cell cloning of colon cancer stem cells reveals a multi-lineage differentiation capacity. Proc. Natl. Acad. Sci. U. S. A 105, 13427-13432.

Merlos-Suarez,A., Barriga,F.M., Jung,P., Iglesias,M., Cespedes,M.V., Rossell,D., Sevillano,M., Hernando-Momblona,X., da Silva-Diz,V., Munoz,P., Clevers,H., Sancho,E., Mangues,R., and Batlle,E. (2011). The intestinal stem cell signature identifies colorectal cancer stem cells and predicts disease relapse. Cell Stem Cell 8, 511-524.

Subramanian,A., Tamayo,P., Mootha,V.K., Mukherjee,S., Ebert,B.L., Gillette,M.A., Paulovich,A., Pomeroy,S.L., Golub,T.R., Lander,E.S., and Mesirov,J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U. S. A 102, 15545-15550.

van de Wetering,M., Sancho,E., Verweij,C., de Lau,W., Oving,I., Hurlstone,A., van der Horn,K., Batlle,E., Coudreuse,D., Haramis,A.P., Tjon-Pon-Fong,M., Moerer,P., van den Born,M.,

Soete,G., Pals,S., Eilers,M., Medema,R., and Clevers,H. (2002). The betacatenin/ TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. Cell 111, 241-50.

Vermeulen,L., De Sousa E Melo, van der Heijden,M., Cameron,K., de Jong,J.H., Borovski,T., Tuynman,J.B., Todaro,M., Merz,C., Rodermond,H., Sprick,M.R., Kemper,K., Richel,D.J., Stassi,G., and Medema,J.P. (2010). Wnt activity defines colon cancer stem cells and is regulated by the microenvironment. Nat. Cell Biol. 12, 468-476.

CLAIMS

1. Method of determining a risk of developing metastasis and/or tumor relapse in a colorectal cancer patient comprising the steps of:
(a) determining in a biological sample from said patient the methylation status and/or the expression level of one or more genes selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2, and
(b) determining whether one or more of said genes is hypermethylated or under expressed, wherein hypermethylation and/or under expression of one or more of said genes indicates an increased risk of developing metastasis and/or tumor relapse.

2. Method according to claim 1, wherein hypermethylation and/or under expression is determined in at least two, preferably three, four or five genes of the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2.

3. Method according to claim 2, wherein hypermethylation or under expression of at least two, preferably three, four or five of said genes indicates an increased risk of developing metastasis and/or tumor relapse.

4. Method according to claim 1, wherein hypermethylation and/or under expression is determined in the genes selected from the group consisting of:
(a) LGR5, APCDD1, AXIN2, DKK1, and ASCL2,
(b) LGR5, APCDD1, AXIN2 and DKK1,
(c) LGR5, APCDD1, AXIN2, and ASCL2, and
(d) AXIN2 and ASCL2.

5. Method according to any one of claims 1-4, wherein said colorectal cancer patient is suffering from colorectal cancer in stage I, II or III, more preferably stage II or III.

6. Method of predicting the response in a colorectal cancer patient to treatment with a demethylating agent comprising:
(a) determining the risk of developing metastasis and/or tumor relapse in a colorectal cancer patient according to the method of any one of claims 1-5, and
(b) predicting a positive clinical response to said treatment, if an increased risk is determined in step (a).

7. Method according to claim 6, wherein said demethylating agent is 5-aza-cytidine, 5-aza-2'-deoxycytidine (decitabine), zebularine [1-(β-D-ribofuranosyl(dihydro-pyrimidin-2-1)] or L-methionine.

8. Method for identifying and/or selecting a colorectal cancer patient suitable for treatment with a demethylating agent comprising:
(a) determining the response in a colorectal cancer patient to treatment to colorectal cancer with a demethylating agent according to the method of claim 6 or 7, and
(b) identifying and/or selecting the colorectal cancer patient for treatment with said demethylating agent if a positive clinical response is determined in step (a).

9. Method for selecting a suitable treatment regimen for colorectal cancer in a colorectal cancer patient comprising:
(a) identifying and/or selecting a colorectal cancer patient suitable for treatment with a demethylating agent according to the method of claim 8, and
(b) selecting said demethylating agent for treatment if the colorectal cancer patient is identified and/or selected for treatment with said demethylating agent in step (a).

10. Method of treating a colorectal cancer patient with a demethylating agent comprising:
(a) determining in a biological sample from said patient the methylation status and/or the expression level of one or more genes selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2, and
(b) determining whether one or more of said genes is hypermethylated or under expressed,
(c) administering to said patient a therapeutically effect amount of a demethylating agent if one or more of said genes is hypermethylated or under expressed.

11. A kit comprising one or more primers for determining the methylation status of at least one gene selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2, wherein said one or more primers is an oligonucleotide having 18 to 40 nucleotides and having at least 80%, preferably at least 95% sequence identity to a nucleic acid sequence selected from Seq ID No. 31, Seq ID No. 32, Seq ID No. 33, Seq ID No. 34, Seq ID No. 35, Seq ID No. 36, Seq ID No. 37, Seq ID No. 38, Seq ID No. 39, Seq ID No. 40, Seq ID No. 49, Seq ID No. 50, Seq ID No. 55, Seq ID No. 56, Seq ID No. 57, and Seq ID No. 58.

12. A kit according to claim 11, comprising two or more primers for determining the methylation status of at least two genes selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2, wherein said two or more primers are oligonucleotides having 18 to 40

nucleotides and having at least 80%, preferably at least 95% sequence identity to a nucleic acid sequence selected from Seq ID No. 31, Seq ID No. 32, Seq ID No. 33, Seq ID No. 34, Seq ID No. 35, Seq ID No. 36, Seq ID No. 37, Seq ID No. 38, Seq ID No. 39, Seq ID No. 40, Seq ID No. 41, Seq ID No. 42, Seq ID No. 43, Seq ID No. 44, Seq ID No. 45, Seq ID No. 46, Seq ID No. 47, Seq ID No. 48, Seq ID No. 49, Seq ID No. 50, Seq ID No. 51, Seq ID No. 52, Seq ID No. 53, Seq ID No. 54, Seq ID No. 55, Seq ID No. 56, Seq ID No. 57, and Seq ID No. 58.

13. A kit for assessing methylation in a test sample, 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 one or more oligonucleotide primers and/or pair of oligonucleotide primers that specifically hybridizes under amplification conditions to a gene selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2, preferably one or more primer according to claim 11 or 12.

14 Use of the kit of any one of claims 11-13 for determining the methylation status of one or more genes selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2.

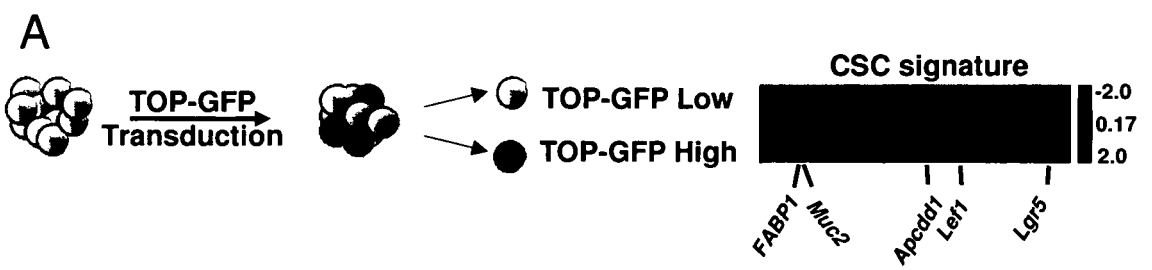
15 Use of the kit of claim any one of claims 11-13 for determining whether a sample is from an individual at risk of developing metastasis and/or tumor relapse.

16 Use of the kit of claim any one of claims 11-13 for determining a risk of developing metastasis and/or tumor relapse in a colorectal cancer patient.

17. A method for classifying a biological sample of a colorectal cancer patient, comprising determining in said sample the methylation status and/or the expression level of a gene selected from the group consisting of LGR5, APCDD1, AXIN2, DKK1, and ASCL2.

18. Method according to any one of claims 1-10, wherein the methylation level is determined using one or more primers as defined in claim 11 or 12.

Fig. 1


A
TOP-GFP
Transduction
TOP-GFP Low
TOP-GFP High
CSC signature
-2.0
0.17
2.0
FABP1
Muc2
Apcdd1
Lef1
Lgr5

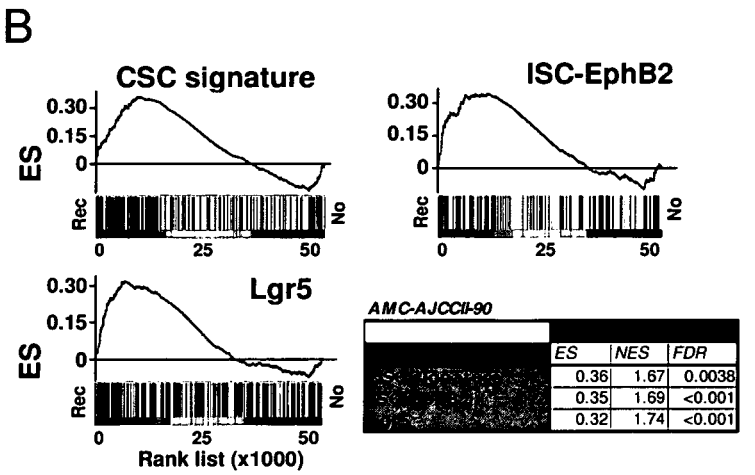
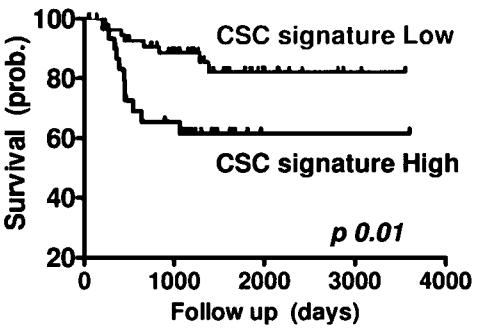
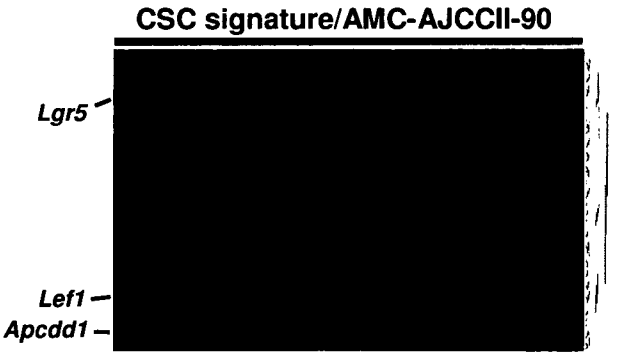

B
CSC signature
ISC-EphB2
Lgr5
ES
0.30
0.15
0
Rec
No
0
25
50
Rank list (x1000)
AMC-AJCCII-90
ES NES FDR
0.36 1.67 0.0038
0.35 1.69 <0.001
0.32 1.74 <0.001

Fig. 1 continued

C



D



E

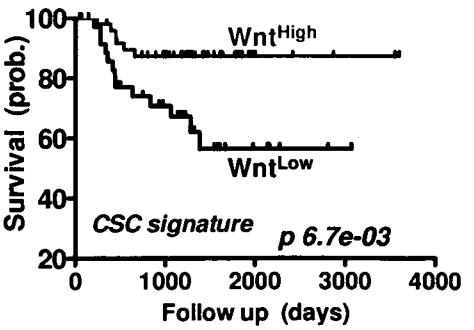


Fig. 1 continued

F

G

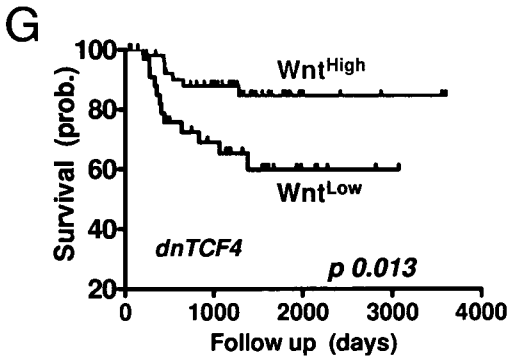


Fig. 2

A

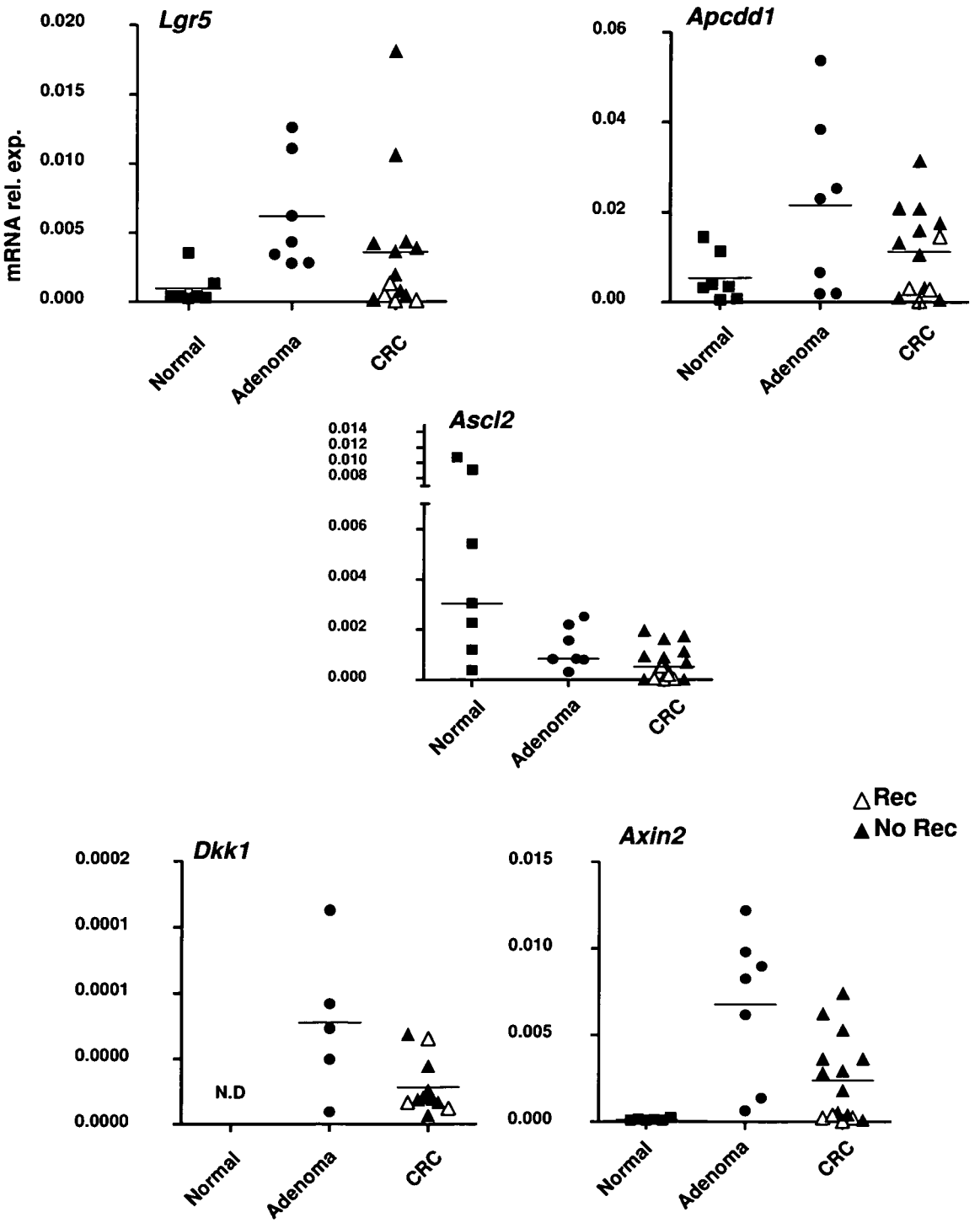
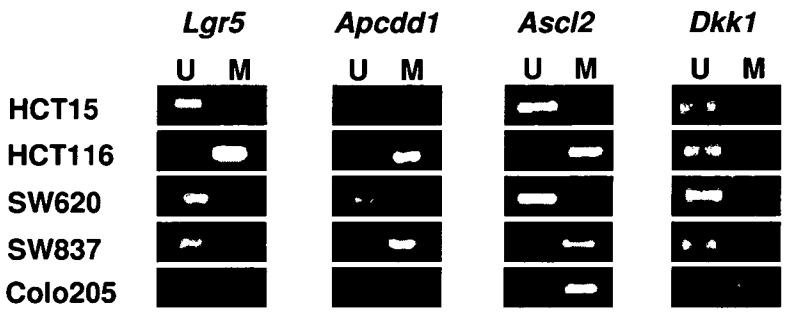

Lgr5
mRNA rel. exp.
0.020
0.015
0.010
0.005
0.000
Normal
Adenoma
CRC
Apcdd1
0.06
0.04
0.02
0.00
Normal
Adenoma
CRC
Ascl2
0.014
0.012
0.010
0.008
0.006
0.004
0.002
0.000
Normal
Adenoma
CRC
△Rec
▲No Rec
Dkk1
0.0002
0.0001
0.0001
0.0000
0.0000
N.D
Normal
Adenoma
CRC
Axin2
0.015
0.010
0.005
0.000
Normal
Adenoma
CRC

Fig. 2 continued

B



C

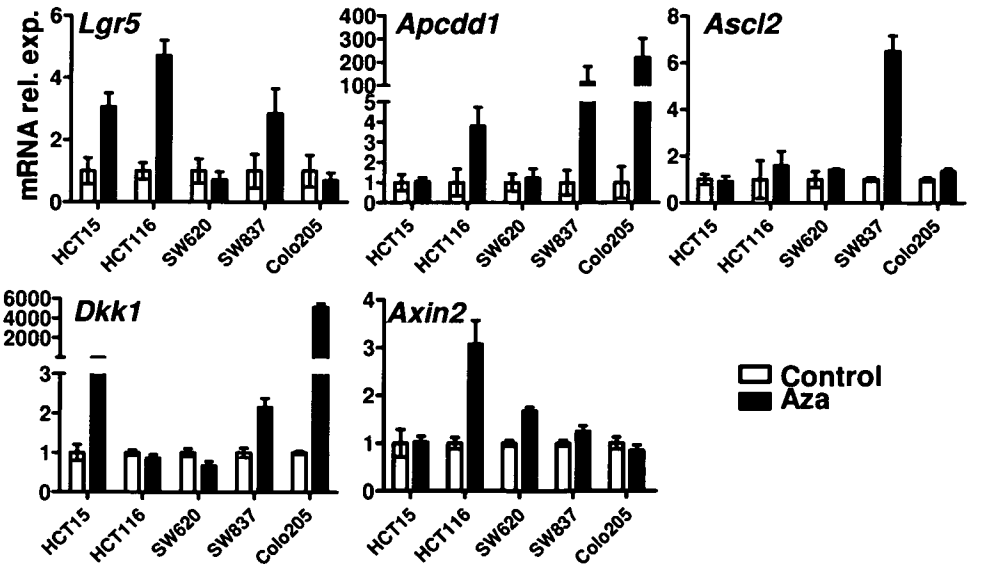
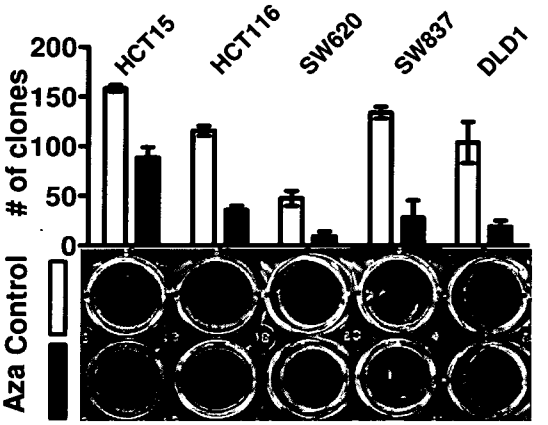


Fig. 3

A


HCT15
HCT116
SW620
SW837
DLD1
200
150
100
50
0
# of clones
Aza
Control

B

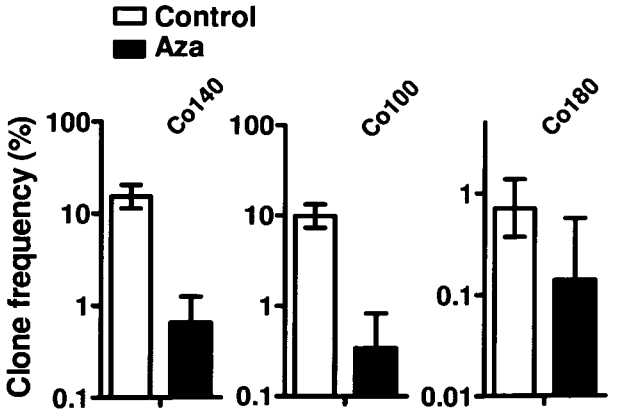
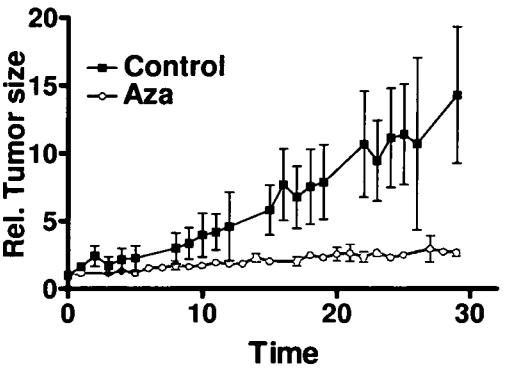

Control
Aza
Co140
Co100
Co180
100
10
1
0.1
0.01
Clone frequency (%)

Fig. 3 continued

C



D

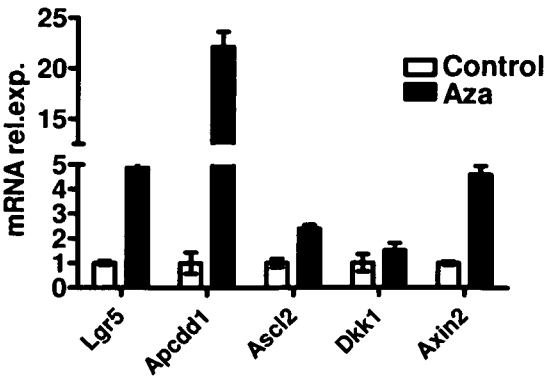
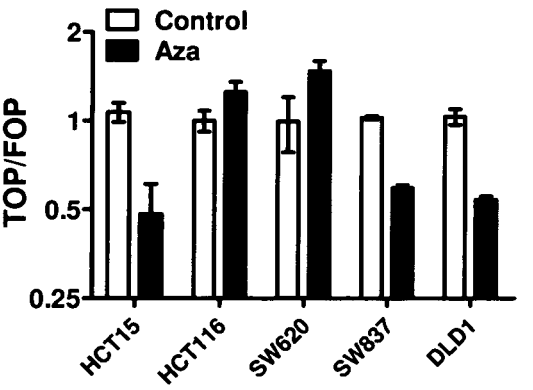


Fig. 3 continued

E



F

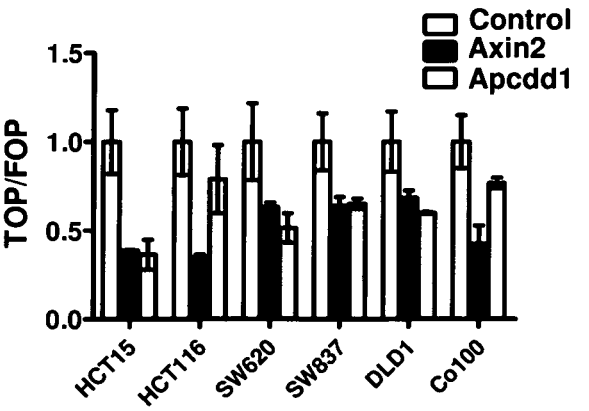


Fig. 4

A

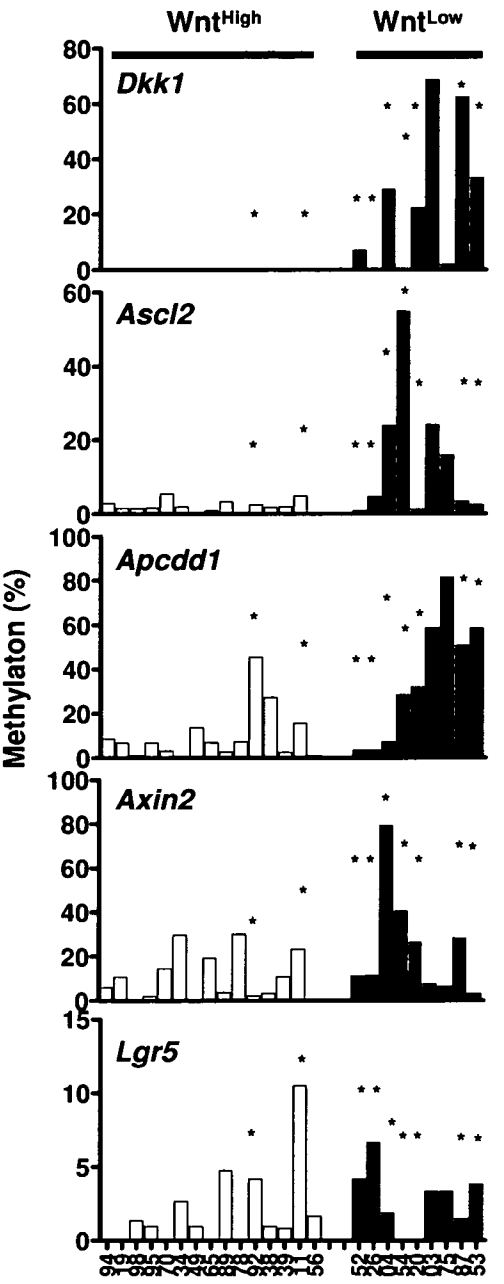
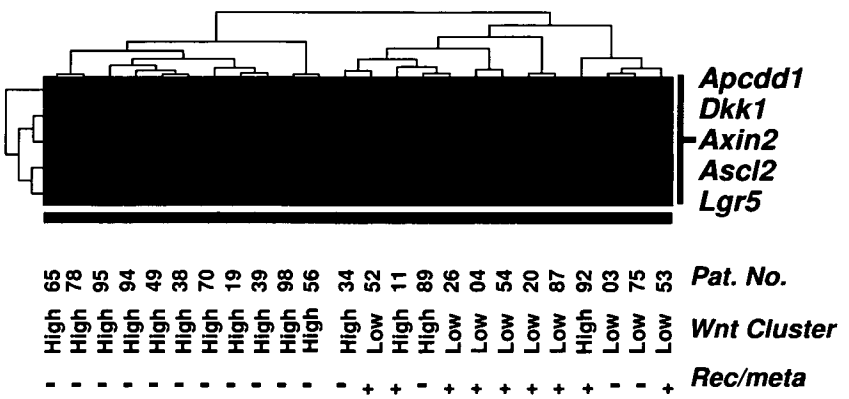


Fig. 4 continued

B



C

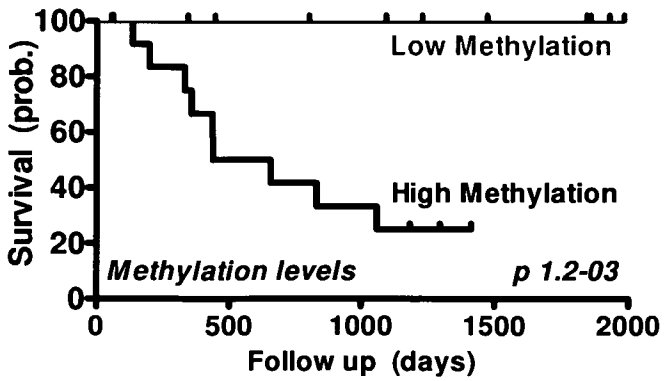
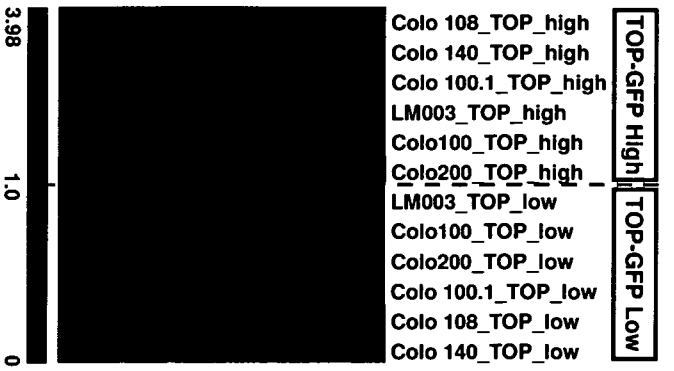


Fig. 5

A


3.98
1.0
0
Colo 108_TOP_high
Colo 140_TOP_high
Colo 100.1_TOP_high
LM003_TOP_high
Colo100_TOP_high
Colo200_TOP_high
LM003_TOP_low
Colo100_TOP_low
Colo200_TOP_low
Colo 100.1_TOP_low
Colo 108_TOP_low
Colo 140_TOP_low
TOP-GFP High
TOP-GFP Low

B

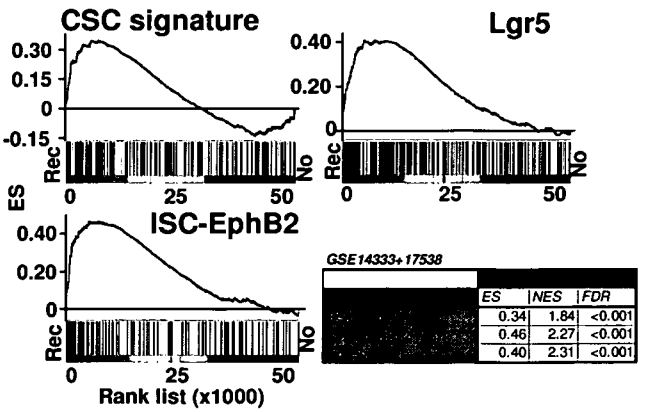
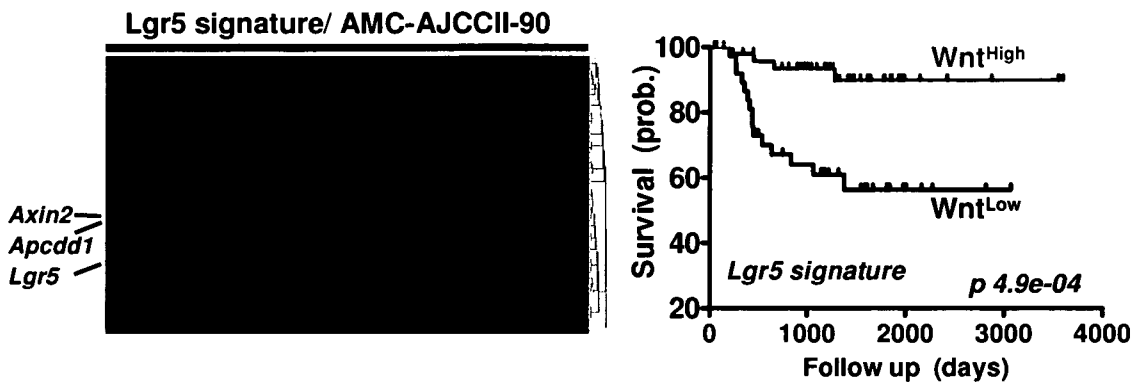

CSC signature
Lgr5
ISC-EphB2
ES
Rec
No
Rank list (x1000)
GSE14333+17538
ES NES FDR
0.34 1.84 <0.001
0.46 2.27 <0.001
0.40 2.31 <0.001

Fig. 5 continued

C



D

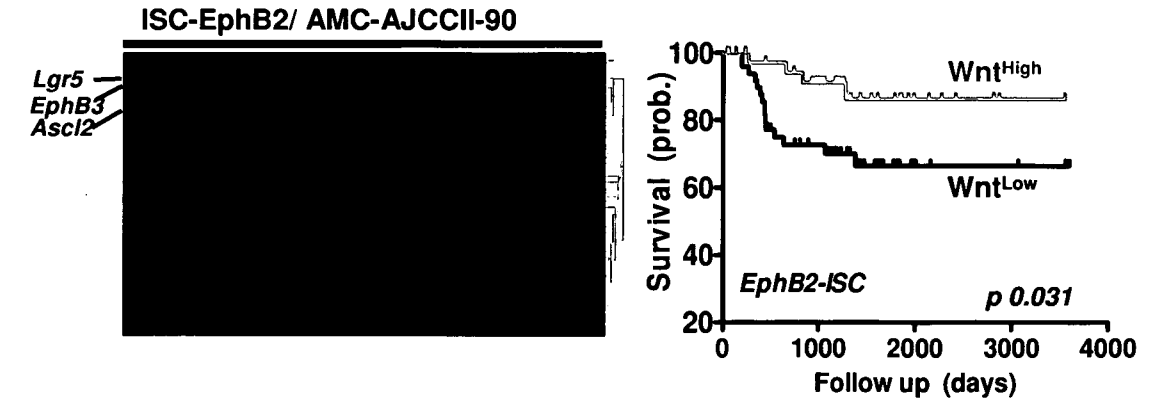


Fig. 5 continued

E

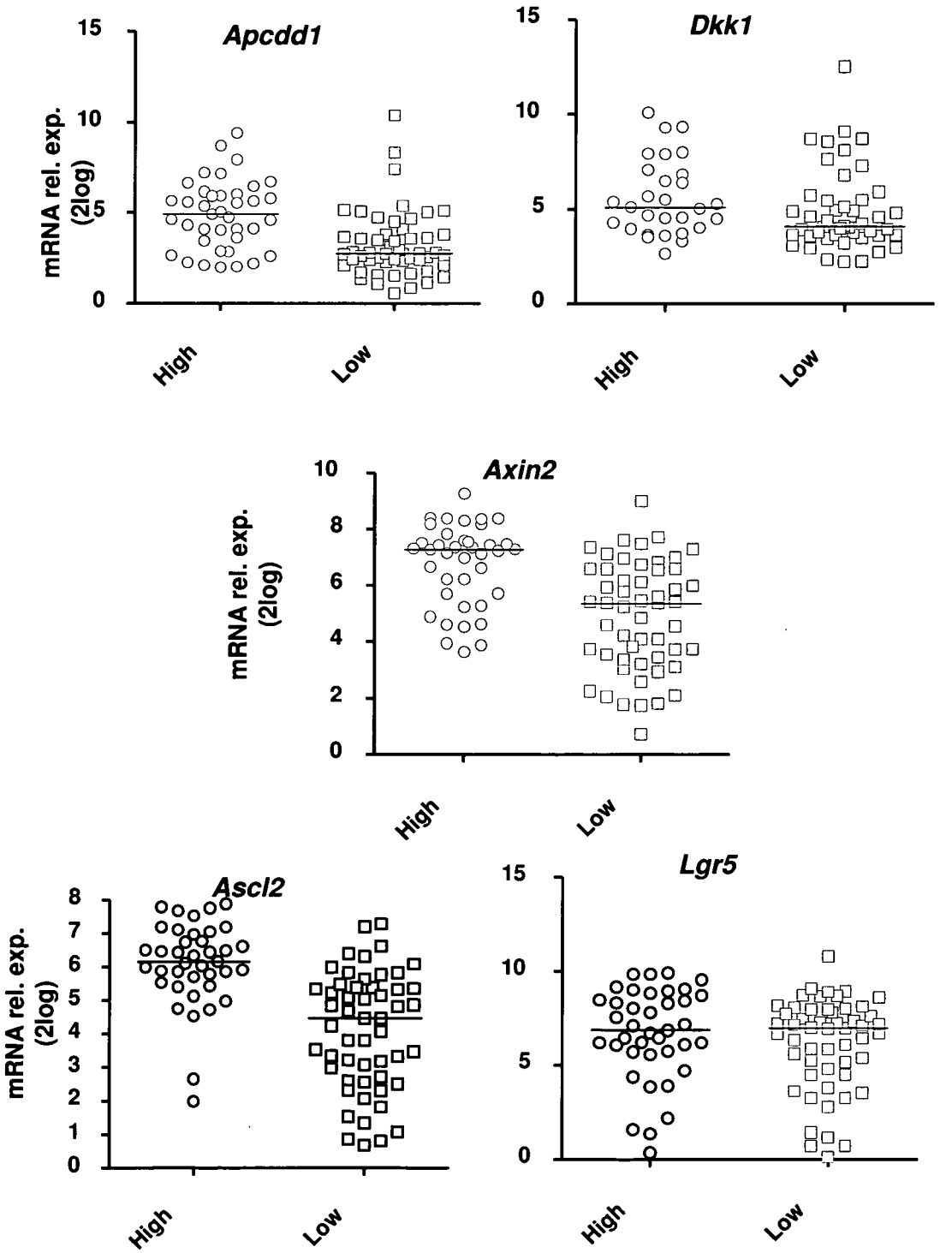


Fig. 6

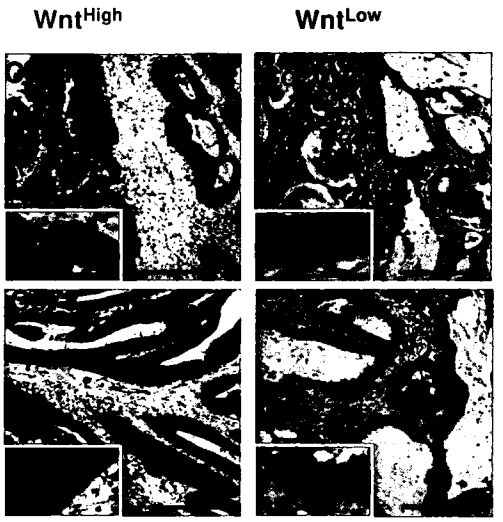
A

Multivariate Cox regression analysis

	p-value	HR	95% CI
Wnt targets			
High			
Low	0.002	5.9	1.89-18.3
p53			
Wt			
Mutant	0.546	1.35	0.50-3.66
Ras			
Wt			
Mutant	0.563	1.4	0.45-4.31
Raf			
Wt			
Mutant	0.62	1.38	0.38-5.08
MSI status			
MSS			
MSI	0.02	0.22	0.061-0.786

Fig. 6 continued

B



C

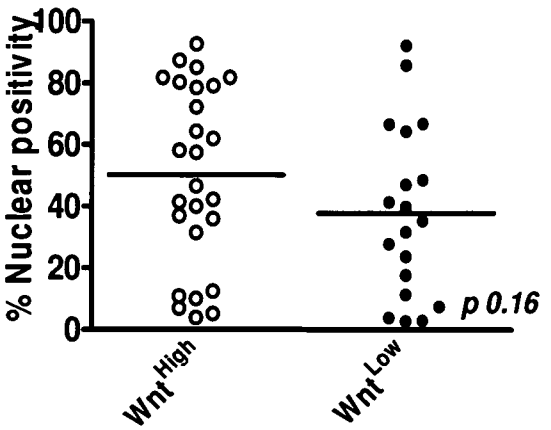


Fig. 6 continued

D

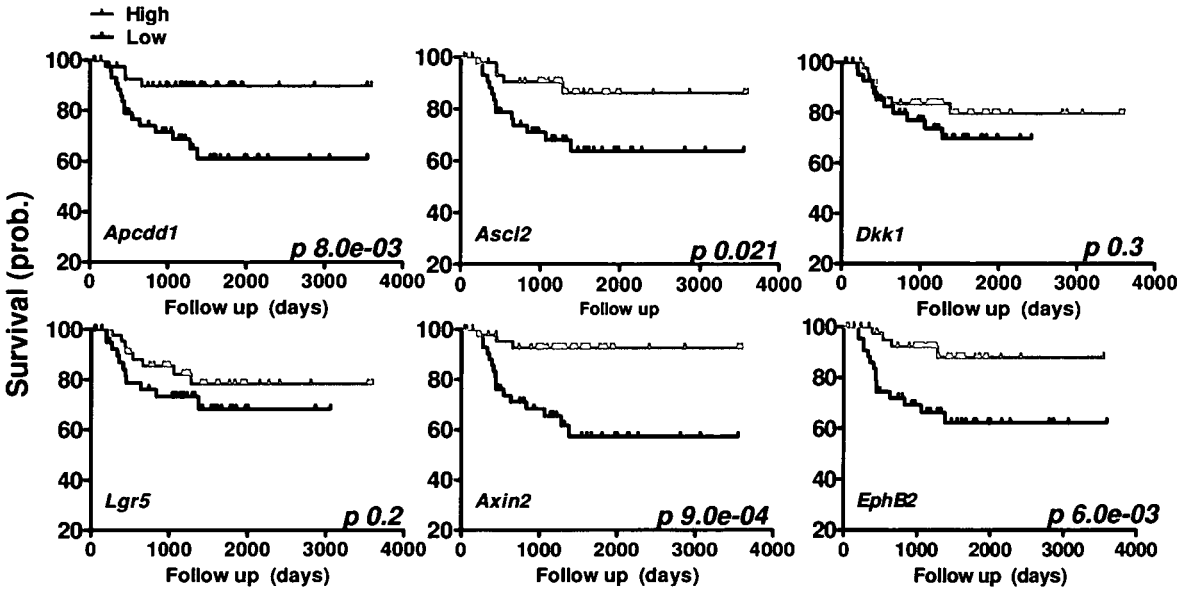
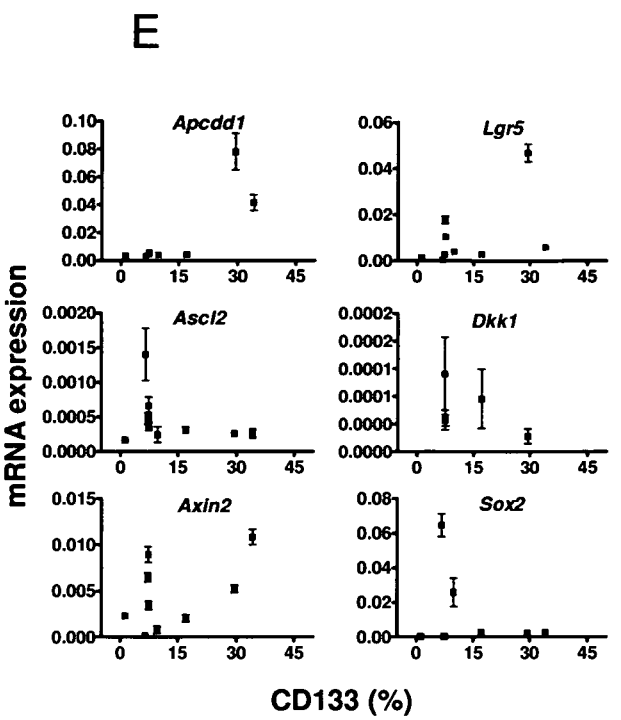


Fig. 6 continued



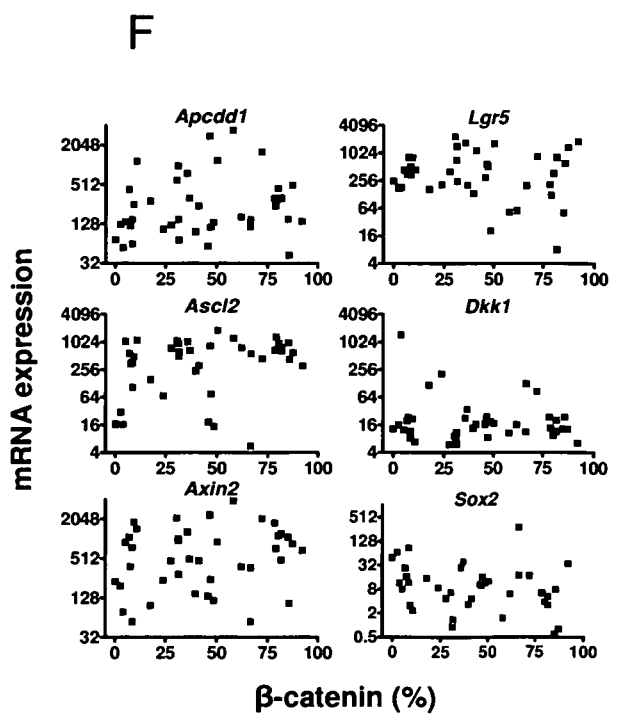


Fig. 7

A

		AMC-AJCCII-90			SIEBER_SMITH 345		
		ES	NES	FDR	ES	NES	FDR
[redacted]	[redacted]	-0.169	-0.88868	0.762583	**-0.32339**	**1.220367**	**0.072775**
[redacted]	[redacted]	**-0.234**	**-1.11745**	**0.214573**	**-0.47062**	**1.825061**	**<0.0001**
[redacted]	[redacted]	-0.195	-1.0248	0.45474	-0.29044	1.112615	0.251484
[redacted]	[redacted]	**-0.245**	**-1.20843**	**0.195024**	-0.24034	0.865386	0.784324
[redacted]	[redacted]	**-0.277**	**-1.27039**	**0.14704**	**-0.49223**	**1.819479**	**<0.0001**
[redacted]	[redacted]	**-0.271**	**1.429386**	**0.089712**	**-0.42836**	**1.640488**	**<0.0001**
[redacted]	[redacted]	**-0.281**	**-1.48283**	**0.182515**	**-0.41257**	**1.563696**	**<0.0001**
[redacted]	[redacted]	**-0.280**	**-1.44235**	**0.117198**	**-0.40845**	**1.530571**	**<0.0001**
[redacted]	[redacted]	**-0.249**	**-1.1732**	**0.214887**	0.176725	0.621088	1
[redacted]	[redacted]	-0.292	-1.13812	0.24833	-0.25963	-0.90986	0.826413

B

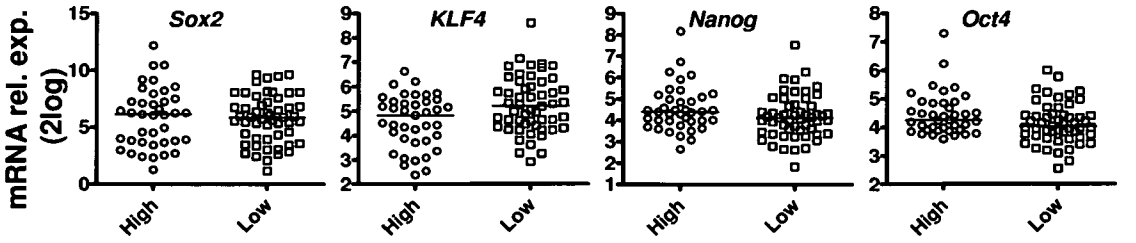
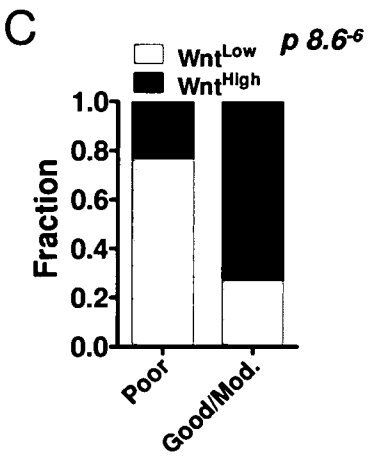


Fig. 7 continued



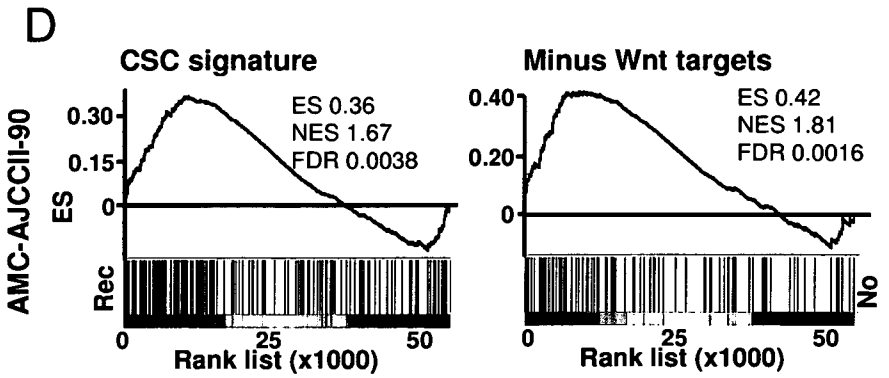
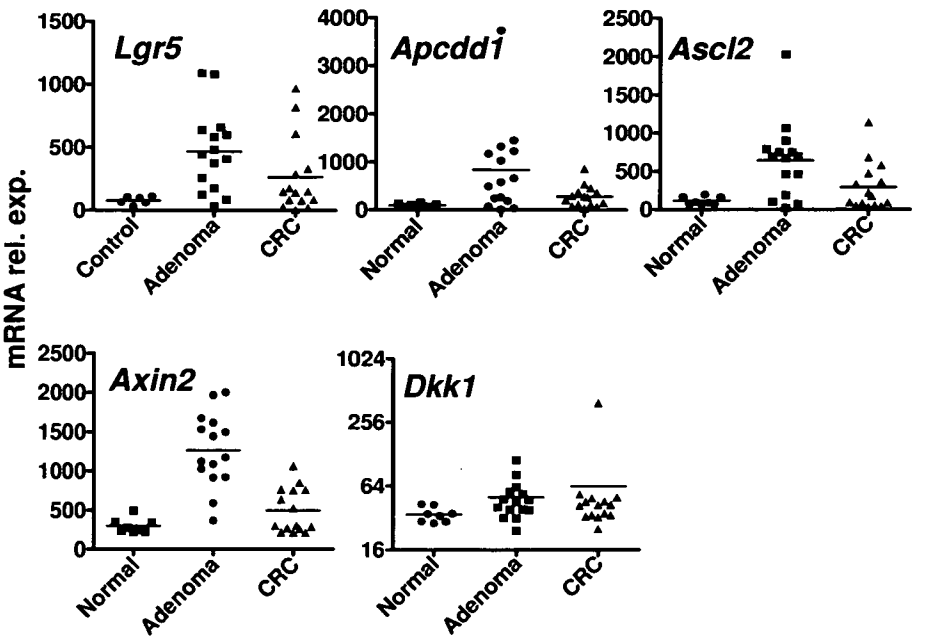
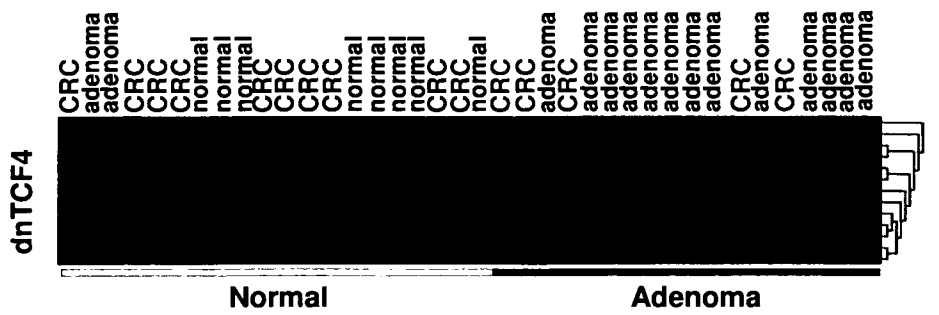


Fig. 7 continued

E



F



INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP2011/005555

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application and necessary to the claimed invention, the international search was carried out on the basis of:

   a. (means)

   [ ] on paper

   [X] in electronic form

   b. (time)

   [ ] in the international application as filed

   [X] together with the international application in electronic form

   [ ] subsequently to this Authority for the purpose of search

2. [ ] In addition, in the case that more than one version or copy of a sequence listing and/or table relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2011/005555

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12Q1/68
ADD.

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

B. FIELDS SEARCHED

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

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, BIOSIS, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	HAO JUN-MEI ET AL: "A five-gene signature as a potential predictor of metastasis and survival in colorectal cancer", JOURNAL OF PATHOLOGY, vol. 220, no. 4, March 2010 (2010-03), pages 475-489, XP002679451, ISSN: 0022-3417	2-5,7, 12,18
Y	abstract, p. 476, col. 1 and p. 477, col.1	1,6, 8-10,17
	----- -/--	

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

[X] See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
18 July 2012	27/08/2012

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

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2011/005555

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WALKER FRANCESCA ET AL: "LGR5 Is a Negative Regulator of Tumourigenicity, Antagonizes Wnt Signalling and Regulates Cell Adhesion in Colorectal Cancer Cell Lines", PLOS ONE, vol. 6, no. 7, E22733, July 2011 (2011-07) , pages 1-20, XP002679452, ISSN: 1932-6203 abstract	1,6, 8-10,17

Y	AGUILERA O ET AL: "Epigenetic inactivation of the Wnt antagonist DICKKOPF-1 (DKK-1) gene in human colorectal cancer", ONCOGENE, vol. 25, no. 29, July 2006 (2006-07), pages 4116-4121, XP002679453, ISSN: 0950-9232 abstract	1,6, 8-10,17

Y	HUGHES T A ET AL: "Regulation of axin2 expression at the levels of transcription, translation and protein stability in lung and colon cancer", CANCER LETTERS, vol. 233, no. 2, 28 February 2006 (2006-02-28), pages 338-347, XP027883753, NEW YORK, NY, US ISSN: 0304-3835 [retrieved on 2006-02-28] abstract	1,6, 8-10,17

X	AHN JOONG BAE ET AL: "DNA Methylation Predicts Recurrence From Resected Stage III Proximal Colon Cancer", CANCER, vol. 117, no. 9, May 2011 (2011-05), pages 1847-1854, XP002679454, ISSN: 0008-543X, DOI: 10.1002/CNCR.25737	2-5,7,12
Y	abstract	1,6, 8-10,17

X	FR 2 919 063 A1 (BIOMERIEUX SA [FR]) 23 January 2009 (2009-01-23)	2-5,7,12
Y	abstract	1,6, 8-10,17

X	WO 2009/143281 A2 (INTRADIGM CORP [US]; XIE FRANK Y [US]; YANG XIAODONG [US]) 26 November 2009 (2009-11-26) the whole document	11,13-16

	-/--	

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2011/005555

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2006/015742 A2 (UNIV FRIEDRICH ALEXANDER ER [DE]; BRUECKL WOLFGANG MICHAEL [DE]; WEIN) 16 February 2006 (2006-02-16)	2-5,7,12
Y	abstract	1,6, 8-10,17

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2011/005555

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
FR 2919063 A1	23-01-2009	AU 2008290466 A1	26-02-2009
		CA 2693655 A1	26-02-2009
		CN 101855555 A	06-10-2010
		EP 2171469 A2	07-04-2010
		FR 2919063 A1	23-01-2009
		JP 2010533851 A	28-10-2010
		US 2010129828 A1	27-05-2010
		WO 2009024691 A2	26-02-2009
WO 2009143281 A2	26-11-2009	NONE	
WO 2006015742 A2	16-02-2006	DE 102004037860 A1	16-03-2006
		EP 1776475 A2	25-04-2007
		US 2008311567 A1	18-12-2008
		WO 2006015742 A2	16-02-2006