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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 December 2018 (13.12.2018)

WIPO | PCT

(10) International Publication Number
WO 2018/224731 A1

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

(21) International Application Number:
PCT/FI2018/050419

(22) International Filing Date:
04 June 2018 (04.06.2018)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
20175513 05 June 2017 (05.06.2017) FI

(71) Applicant: **LS CANCERDIAG OY** [FI/FI]; Kuutsalontie 26 B, 00910 Helsinki (FI).

(72) Inventors: **NYSTRÖM, Minna**; Kuutsalontie 26 B, 00910 Helsinki (FI). **PUSSILA, Marjaana**; Laurinniityntie 2 C 48, 00440 Helsinki (FI).

(74) Agent: **BERGGREN OY**; P.O. Box 16, (Eteläinen Rautatiekatu 10 A), 00101 Helsinki (FI).

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

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

Published:
— *with international search report (Art. 21(3))*
— *with sequence listing part of description (Rule 5.2(a))*

WO 2018/224731 A1

(54) Title: A METHOD FOR DETERMINING WHETHER A SUBJECT IS AT RISK TO DEVELOP CANCER AND TOOLS RELATED THERETO

(57) Abstract: The present invention relates to the fields of diagnostics and more specifically predicting or estimating occurrence of diseases such as cancer. Still, the present invention relates to a method for determining whether a subject is at risk to develop cancer. And furthermore, the present invention relates to a kit comprising tools to determine the expression level of at least three genes associated with chromosomal segregation. Still furthermore the present invention relates to use of a kit of the present invention for determining whether a subject is at risk to develop cancer.

A method for determining whether a subject is at risk to develop cancer and tools related thereto

FIELD OF THE INVENTION

The present invention relates to the fields of diagnostics and more specifically predicting or estimating occurrence of diseases such as cancer. Still, the present invention relates to a method for determining whether a subject is at risk to develop cancer. And furthermore, the present invention relates to a kit comprising tools to determine the expression level of at least three genes associated with chromosomal segregation. Still furthermore the present invention relates to use of a kit of the present invention for determining whether a subject is at risk to develop cancer.

BACKGROUND OF THE INVENTION

Colorectal cancer (CRC) is the third most common cancer and the fourth most common cause of cancer-related deaths worldwide(1). The incidence rates increase significantly with age and interactions between genetic and environmental factors, including diet, are suggested to play a critical role in its etiology(2, 3). Cancer development always includes lack of genomic integrity in cells and different types of genomic instability, such as chromosomal instability (CIN) and microsatellite instability (MIN, also called MSI) are thought to reflect distinct initiating mechanisms in cancer(4). Three different pathways leading to genomic instability in colon cancer have been described. Most CRCs represent CIN, where chromosomes fail to trigger the spindle assembly checkpoint (SAC) leading to aberrant chromosome segregation. In recent years, many new genes have been reported, whose mutations and expression changes disturb chromosomal stability causing aneuploidy and/or comprehensive loss of heterozygosity (LOH) and alterations in chromosome structures(5-6). About 15% of sporadic CRCs and over 95% of CRCs in Lynch syndrome (LS), the most common inherited colon cancer syndrome, represent MSI caused by a defective DNA mismatch repair (MMR) mechanism(8). MMR deficiency causes accumulation of point mutations in the genome and especially in short repeat sequences called microsatellites, and is thought to be the driver defect in MSI carcinomas(10). The third pathway, CIMP (CpG Island Methylator Phenotype), characterized by global genome hypermethylation and tumor suppressor gene silencing, is seen in 20-30% of CRCs(8).

Colon cancer research focuses mainly on tumor characteristics, such as genomic instability, which can be utilized in treatment design. Recent findings have revealed that CIN and MSI pathways are not mutually exclusive(5, 7, 16), suggesting that also tumors with distinct features and instabilities may share initiative genomic aberrations while different tumor characteristics reflect subsequent alterations during cancer development.

Despite progressive cancer research, right now there are no very efficient tools and methods for determining the risk of developing cancer, e.g. colon cancer, from a sample of a subject before occurrence of cancer by utilizing other means than genetic testing revealing susceptibility to inherited diseases.

E.g. an EpiProColon test is available in the market allowing screening for colorectal cancer. The Septin9 blood test permits detection of the biomarker mSEPT9 and its epigenetic modification in blood plasma. The methylated gene SEPT9 is found in colorectal cancer but not in the healthy colon and therefore Septin9 blood test cannot be used for determining the risk of cancer before onset of cancer.

BRIEF DESCRIPTION OF THE INVENTION

An object of the present invention is thus to provide a method and tools for effective and specific cancer risk tests. The object of the invention is achieved by utilizing a specific combination of genes associated with chromosomal segregation.

The aim of the present invention was to find a panel of genes contributing to colon cancer development, and which panel could be efficiently used for evaluating the risk of cancer. Indeed, it has now been found that expression levels of specific genes can be used for determining the risk of cancer.

The present invention solves the problem of the field i.e. lack of very efficient and specific pre-malignant or pre-cancer test methods concerning colon cancer without results of genetic tests revealing susceptibility to inherited diseases. On the other hand the results of genetic tests revealing susceptibility to inherited diseases do not have any effect on the method of the present invention. Also, the present invention enables testing of colon samples for determining the risk of colon cancer independent of specific characteristics of colon cancers (e.g. including but not limited to one or more selected from the group consisting of CIN, MSI, microsatellite stability

(MSS), aneuploidy, LOH, CIMP), i.e. a method and test of the present invention are suitable for determining the risk of any colon cancer.

There is a so called "field defect" in the proximal portion of the colon, which is not detectable histologically from the normal mucosa and which predisposes to cancer. Genome-wide transcriptome analysis described in the present disclosure revealed that the normal mucosa expression profiles of the CRC mice is different from the profiles of the normal mice and formed a distinct cluster. Indeed, differences of expressions of very specific genes predispose to cancer. Thus, the present invention provides a method and tools for determining whether a subject is at risk to develop cancer. Optionally, patients with said differences can further be monitored with colonoscopy and thus, the development of cancer can be prevented at a very early stage.

Detection of field defects by a method and tools of the present invention can be exploited for determining a cancer risk of a specific tissue (colon) and site (site of a sample) or for following up said specific tissue or site. Gene expressions and changes thereof may thus be determined from specific sites of an organ system. Compared to the present invention prior art methods and tests (e.g. gene tests, metabolomics) reveal a general risk of a subject for a disease, not a site specific risk.

The present invention makes it possible e.g. to utilize the information achieved by a method and tools of the present invention for detecting cancer development or preventing cancer. Thus, the present invention enables screening of subjects, follow up and colonoscopies of cancer prone subjects and development of new preventive methods. As an example a person found to be at risk to develop cancer may change his life style or diet and thus may prevent or delay onset of colon cancer.

In the present invention a mouse model was used to study cancer preceding expression changes in colon mucosa, Mlh1 protein expression and MSI status in tumors, and the effect of inherited predisposition (*Mlh1*$^{+/-}$) and Western-style diet (WD) on those. A long term feeding experiment with either a healthy rodent diet AIN-93G (AIN) or Western-style diet (WD) modified from AIN was conducted. WD was used to ensure the development of colon carcinomas, since it has previously been shown to cause CRCs in mice even without any predisposing mutation or carcinogen treatment. Carcinomas developed mainly in WD fed mice. Interestingly, neither wildtype

Mlh1$^{+/+}$ nor heterozygote *Mlh1*$^{+/-}$ (B6.129-*Mlh1*^{tm1Rak}) mice lacked the Mlh1 protein or showed MSI in CRCs, while *Mlh1* RNA expression was already significantly decreased in the mucosa. Instead, CRC mice showed a distinct expression profile with shortage of *Mlh1* and/or several other chromosomal segregation gene-specific transcripts in mucosa and aberrant mitosis in tumors. (*Mlh1*$^{+/-}$ mice represent the mouse counterpart of Lynch syndrome.)

The mouse model of the present disclosure provided a valuable tool to study the process of carcinogenesis from the earliest changes in colon mucosa until tumor development and characterization. Moreover, the use of an animal model enabled to distinguish gene expression changes and sort out the ones that signal carcinogenesis.

The present invention relates to a method for determining whether a subject is at risk to develop cancer, wherein the method comprises: determining in a colon mucosa sample from a subject the expression level of at least three genes associated with chromosomal segregation; and determining the risk of cancer using the determined expression levels of the genes.

Also, the present invention relates to a kit comprising tools to determine the expression level of at least three genes associated with chromosomal segregation, one or more control samples, and reagents for performing said method.

And still, the present invention relates to a kit for use in a method according to the present invention comprising tools to determine the expression level of at least three genes associated with chromosomal segregation, and optionally one or more control samples and/or optionally reagents for performing said method.

And still, the present invention relates to use of a kit of the present invention for determining whether a subject is at risk to develop cancer.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows results of Mlh1 protein expression and loss of heterozygosity analyzes. (Figure 1A) An example of a colon carcinoma showing positive Mlh1 expression analyzed by immunohistochemistry (mouse E402, tubular adenocarcinoma). (Figure 1B) Four CRCs found in the heterozygote *Mlh1*$^{+/-}$ mice showing that the

normal *Mlh1* allele (350bp) was still present in tumors. (Figure 1C) In *Mlh1* heterozygote mice one of the *Mlh1* alleles is mutated by replacing the exon 2 with a neomycin cassette. Loss of Mlh1 heterozygosity was analyzed using the genotyping primers M001, M002, and M003, which produce two different length fragments, 350 bp and 500 bp, that separate the normal (M001/M003) and the mutated allele (M001/M002), respectively.

Figure 2 shows number of colon tumors and carcinomas in different age and diet groups. Aging and Western-style diet increased the total number of (Figure 2A) colon tumors and (Figure 2B) carcinomas. AIN (AIN-93G control diet), WD (Western-style diet).

Figures 3A and 3B show genome wide expression profiles in normal colon mucosa. MDS plots created with the 100 most differentially expressed genes between CRC (grey) mice and mice without cancer (black) (Figures 3A and 3B). The expression profiles of all six mice which developed carcinoma up to 18 mo of age form a distinct cluster in the plot.

Figure 4 shows the expression levels of (at least three) 10 chromosomal segregation-specific genes in colon mucosa. The expression levels (gene expression values after ComBat processing) are described as (Figure 4A) a line chart and (Figure 4B) expression values. In the carcinoma mice (E249, E314, E329, E333, E338, E347), two mice with similar expression profiles with CRC mice (E325 and E332) and the average levels of 74 mice without cancer.

Figure 5 reveals abnormal mitoses in mouse colon carcinomas. Representative pictures of abnormal mitoses (arrows) in (Figure 5A) serrated adenocarcinoma (mouse E347) and (Figure 5B) tubular adenocarcinoma (mouse E333), and a normal mitosis (arrow head).

Figure 6 shows a table revealing characteristics of mice and their tumors concerning methods of the present invention.

Figure 7 shows expressions of ten genes associated with chromosomal segregation in mouse individuals. Carcinoma mice are presented as E314, E333, E329, E338, E249 and E347. The columns representing expressions of genes are presented in the same order as the names of the genes mentioned in the figure.

Figure 8 shows results of differential gene expression analysis: carcinoma mice versus non-carcinoma mice. Shrink T scores (expression differences) and P-values are highlighted.

DETAILED DESCRIPTION OF THE INVENTION

Surprisingly, the genome-wide expression profiling experiment of the present disclosure demonstrates that very specific cancer preceding changes occur and they can be detected already in normal colon mucosa. These changes form a field-defect in histologically normal mucosa and trigger colorectal cancer (e.g. MMR-proficient, chromosomally unstable colorectal cancer). Very importantly the present invention demonstrates that cancer preceding changes are already seen in histologically normal colon mucosa and that decreased expression of at least three, four or five specific chromosomal segregation genes form a field-defect in mucosa.

As used herein “a field defect” refers to a field of pre-malignant tissue in which a new cancer is likely to arise. Field defects are histologically normal under the microscope. Field defect (also termed field cancerization, field change, field change cancerization, field carcinogenesis, cancer field effect or premalignant field defect) is a biological process in which large areas of cells at a tissue surface or within an organ are affected by carcinogenic alterations. The process arises from exposure to an injurious environment, often over a lengthy period. In the colon, it has been described as the process whereby colonic epithelial cells acquire pro-tumorigenic alterations that are insufficient to cause morphological change but which predispose to tumor (60).

Indeed, one embodiment the present invention relates to a method for determining a specific combination of cancer preceding changes or biomarkers indicative of cancer development in a sample, wherein the method comprises determining in a colon mucosa sample from a subject the expression levels of at least three genes associated with chromosomal segregation; and thereby determining the risk or development of cancer.

Altogether 80 mice were included in genome-wide expression profiling of the present disclosure and only two of the 74 mice without cancer shared the expression

profile of the CRC mice related to chromosomal segregation, indicating that the aberrant expression of this gene set signals carcinogenesis in colon mucosa.

The present invention concerns a method for determining whether a subject is at risk to develop cancer or whether cancer preceding changes are found in a sample of a subject. In said method at least the expression level of at least three (i.e. three or more) genes associated with chromosomal segregation is determined from a colon mucosa sample of a subject and the risk of cancer is determined by using said expression levels of the genes.

Methods of studying expression levels of specific genes or polynucleotides are known to a person skilled in the art and include but are not limited to northern blotting (for detecting specific RNA molecules present within an RNA mixture), reverse transcription polymerase chain reaction (RT-PCR) and quantitative reverse transcription polymerase chain reaction (RT-qPCR) (for detecting and quantifying mRNA), serial analysis of gene expression (SAGE) (utilizing a library of short sequence tags which can each be used to detect a transcript, the transcript can be determined by assessing how many times each tag is detected), a DNA microarray (for determining expression levels with a solid surface with attached collection of microscopic DNA spots), RNA sequencing (for measuring the sequence of RNA molecules, e.g. shotgun sequencing of cDNA molecules acquired from RNA through reverse transcription or technologies for sequencing RNA molecules so that the primary sequence and abundance of each RNA molecule can be determined). Any one or any combination of said methods of studying expression levels of specific genes or polynucleotides may be utilized in the present invention. In a specific embodiment of the present invention RNA sequencing or RT-qPCR is utilized to study the expression levels of specific genes.

As used herein “expression level” refers to amount of RNA copied from the DNA by transcription in the nucleus by RNA polymerase. RNA transcribed from the DNA is complementary to the template 3' → 5' DNA strand. Transcription of eukaryotic genes results in a primary transcript of RNA (pre-mRNA), which first has to undergo a series of modifications to become a mature mRNA. These modifications include but are not limited to 5' capping, 3' cleavage and polyadenylation, and RNA splicing. In one embodiment of the invention the expression level of the specific genes is the level of mRNAs.

As used herein “decreased expression level" or “down-regulated expression” refers to either a lack of expression (no presence of said expression product RNA) or less expression of a gene or polynucleotide of interest (resulting in lower amount of said expression product RNA) compared to a control level. Lack of expression or decreased expression can be proved for example by any one of the methods described above concerning methods of studying expression levels of specific genes or polynucleotides or any other suitable method known to a person skilled in the art.

In one embodiment of the invention the presence, absence or level of expression of the genes associated with chromosomal segregation is determined. In one embodiment the method is an *in vitro* method.

As used herein “a gene” refers to a DNA polynucleotide sequence encoding a specific polypeptide. As used herein "a polynucleotide" refers to any polynucleotide, such as single or double-stranded DNA (genomic DNA or cDNA) or RNA, comprising a nucleic acid sequence encoding a polypeptide in question, or a conservative sequence variant or fragment thereof.

In connection with polynucleotides, the term "conservative sequence variant" refers to nucleotide sequence modifications, which do not significantly alter biological properties of the encoded polypeptide. Conservative nucleotide sequence variants include variants arising from the degeneration of the genetic code and from silent mutations. Nucleotide substitutions, deletions and additions are also contemplated.

The term "variant" as used herein refers to a sequence having minor changes in the amino acid or nucleic acid sequence as compared to a given sequence. Such a variant may occur naturally e.g. as an allelic variant within the same strain, species or genus, or it may be generated by mutagenesis or other gene modification. It may comprise amino acid or nucleic acid substitutions, deletions or insertions, but it still functions in substantially the same manner as the given polypeptide.

As used herein “a fragment” refers to any part of a gene or polynucleotide. In a specific embodiment a fragment of a gene or polynucleotide encodes a polypeptide having activity of a full length polypeptide.

Herein, the terms "polypeptide" and "protein" are used interchangeably to refer to polymers of amino acids of any length.

In one embodiment of the invention the expression levels of the genes associated with chromosomal segregation in the sample when compared to a normal level are indicative of said subject being at risk to develop cancer.

In another embodiment of the invention decreased expression levels of the genes associated with chromosomal segregation in the sample when compared to a normal level are indicative of said subject being at risk to develop cancer.

As used herein “a normal level” refers to reference values. Within the reference (i.e. cut off) values (reference interval) the result is still "normal". A large number of healthy subjects are studied for reference values. Based on the results obtained, the reference values are calculated mathematically so that almost all healthy results are within these values: 95% of healthy people are within the reference range, but about 5% of healthy results may be higher or lower than the reference value. Reference values may vary by age and gender. In one embodiment of the invention the method further comprises calculating reference values from healthy subjects.

In another embodiment of the invention the expression levels of specific genes associated with chromosomal segregation in the subject are compared to normal expression levels of the same genes, and a decrease in the expression levels of said genes in the sample relative to the normal expression level is indicative of an increased risk of developing cancer. In one embodiment a decrease in the expression levels of specific genes is a significant decrease. As used herein “a significant decrease” refers to a decrease in the expression level, which is statistically significant ($p \leq 0.5$). Statistical methods suitable for the present invention are any common statistical methods known to a person skilled in the art. In a specific embodiment of the invention the statistical method for determining a decrease or significant decrease in the expression level includes but is not limited to a t-test, modified t-test, Shrinkage t-test or Fischer’s exact test.

The present invention surprisingly reveals that at least three or more (e.g. three, four, five, six, seven, eight, nine, ten, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 or more) genes associated with chromosomal segregation can be utilized for estimating the risk of a subject for developing cancer.

In a specific embodiment the specificity of the method is selected from the group consisting of 50%, 55%, 60%, 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% and 100% specificity.

In some embodiments of the invention the method may further comprise one or more of the following:
- providing a colon mucosa sample
- isolating total RNA of the sample
- cDNA conversion of the RNA sample
- detecting the expression level of a gene or genes by utilizing at least one primer or probe, which hybridizes (e.g. under stringent conditions) to said gene(s).

The combination of total RNA isolation or cDNA conversion and detection of the expression level of a gene may be performed either separately or simultaneously.

Results of the gene expression methods (e.g. RNA sequencing results) may optionally be confirmed by pathological analysis. Pre-cancerous histologically normal cells show normal mitotic activity and/or low numbers of atypical mitosis. Cancer cells (e.g. carcinoma cells) show increased mitotic activity and abundant numbers of atypical mitoses.

During cell division the spindle assembly checkpoint, which is the major target of mitotic alterations, maintains the genome stability by delaying cell division until all chromosomes are accurately aligned in the spindle. Aberrant expression of mitotic genes leads to mitotic abnormalities including centrosome defects and improper spindle checkpoint leading to chromosomal instability and tumor formation in multiple tissues including colon.

As used herein “chromosomal segregation” refers to a process of a cell wherein in mitosis two sister chromatids formed as a consequence of DNA replication separate from each other and migrate to opposite poles of the nucleus. As used herein “associated with chromosomal segregation” refers to the situation wherein a polypeptide encoded by a specific polynucleotide or gene participates either directly or indirectly in chromosomal segregation. As used herein “directly” refers e.g. to a situation wherein the polypeptide itself has a function in chromosomal segregation. As used herein “indirectly” refers e.g. to a situation wherein a polypeptide has a role for another polypeptide, which directly has a function in chromosomal segregation.

According to the present invention it may be possible or advantageous to determine expressions of at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9 or at least 10 genes presented in Figures 4A and B.

In one embodiment of the invention the expression levels of at least three genes associated with chromosomal segregation *Bub1* (BUB1, mitotic checkpoint serine/threonine kinase), *Mis18a* (MIS18 kinetochore protein A) and *Tpx2* (TPX2 microtubule associated) are determined. In another embodiment at least three genes associated with chromosomal segregation are selected from the following groups Bub1, Mis18a and Rad9a; Bub1, Tpx2 and Rad9a; Bub1, Mis18a and Pms2; Bub1, Tpx2 and Pms2; Bub1, Rad9a and Pms2. Optionally any other genes or any combination thereof (e.g. selected from the group consisting of *Bub1*, *Mis18a*, *Tpx2*, *Rad9a*, *Pms2*, *Mlh1*, *Cenpe*, *Ncapd3*, *Odf2* and *Dclre1b*) may be determined in addition to the combination of said three genes.

In another embodiment of the invention the expression levels of at least four genes associated with chromosomal segregation *Bub1*, *Mis18a*, *Tpx2* and *Rad9a* (RAD9 checkpoint clamp component A), or *Bub1*, *Mis18a*, *Tpx2* and *Pms2* (postmeiotic segregation increased 2) are determined. In another embodiment at least four genes associated with chromosomal segregation are selected from the following groups: Bub1, Mis18a, Rad9a and Pms2; Bub1, Tpx2, Rad9a and Pms2. Optionally any other genes or any combination thereof (e.g. selected from the group consisting of *Bub1*, *Mis18a*, *Tpx2*, *Rad9a*, *Pms2*, *Mlh1*, *Cenpe*, *Ncapd3*, *Odf2* and *Dclre1b*) may be determined in addition to the combination of said four genes.

In a further embodiment of the invention the expression levels of at least five genes associated with chromosomal segregation *Bub1* (BUB1, mitotic checkpoint serine/threonine kinase), *Mis18a* (MIS18 kinetochore protein A), *Tpx2* (TPX2 microtubule associated), *Rad9a* (RAD9 checkpoint clamp component A) and *Pms2* (postmeiotic segregation increased 2) are determined. Optionally one, two, three, four, five or more other genes (e.g. selected from the group consisting of *Mlh1*, *Cenpe*, *Ncapd3*, *Odf2* and *Dclre1b*) may be determined in addition to any combination of said five genes.

In another embodiment of the invention at least three, four, five, six, seven, eight, nine or ten genes associated with chromosomal segregation are selected from the

group consisting of *Bub1*, *Mis18a*, *Tpx2*, *Rad9a*, *Pms2*, *Mlh1*, *Cenpe*, *Ncapd3*, *Odf2* and *Dclre1b*. Any specific combinations of said genes are included within the scope of the invention. In a very specific embodiment the genes associated with chromosomal segregation have been selected from the following groups or combinations of genes:

at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Mlh1;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Cenpe;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Ncapd3;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Odf2;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Dclre1b;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Mlh1, Cenpe;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Mlh1, Cenpe, Ncapd3;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Mlh1, Cenpe, Ncapd3, Odf2;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Mlh1, Cenpe, Ncapd3, Odf2, Dclre1b;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Mlh1, Ncapd3;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Mlh1, Ncapd3, Odf2;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Mlh1, Ncapd3, Odf2, Dclre1b;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Mlh1, Odf2;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Mlh1, Odf2, Dclre1b;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Mlh1, Dclre1b;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Cenpe, Ncapd3;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Cenpe, Ncapd3, Odf2;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Cenpe, Ncapd3, Odf2, Dclre1b;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Cenpe, Odf2;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Cenpe, Odf2, Dclre1b;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Cenpe, Dclre1b;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Ncapd3, Odf2;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Ncapd3, Odf2, Dclre1b;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Ncapd3, Dclre1b;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Odf2, Dclre1b*.

In a still further specific embodiment the expression levels of at least ten genes associated with chromosomal segregation *Bub1* (BUB1, mitotic checkpoint serine/threonine kinase), *Mis18a* (MIS18 kinetochore protein A), *Tpx2* (TPX2 microtubule associated), *Rad9a* (RAD9 checkpoint clamp component A), *Pms2* (postmei-

otic segregation increased 2), *Mlh1* (mutL homologue 1), *Cenpe* (centromere protein E), *Ncapd3* (non-SMC condensing II complex subunit D3), *Odf2* (outer dens fiber of sperm tails 2) and *Dclre1b* (DNA cross-link repair 1B) are determined.

As used herein *Bub1* refers to a gene encoding BUB1 mitotic checkpoint serine/threonine kinase (e.g. Ensembl accession number ENSG00000169679). BUB1 serine/threonine-protein kinase plays a central role in mitosis. The encoded protein functions in part by phosphorylating members of the mitotic checkpoint complex and activating the spindle checkpoint. This protein also plays a role in inhibiting the activation of the anaphase promoting complex/cyclosome. This protein may also function in the DNA damage response. Alternate splicing results in multiple transcript variants.

As used herein *Mis18a* refers to a gene encoding MIS18 kinetochore protein A (e.g. Ensembl accession number ENSG00000159055). MIS18 kinetochore protein A is required for recruitment of CENPA to centromeres and normal chromosome segregation during mitosis.

As used herein *Tpx2* refers to a gene encoding a TPX2 microtubule associated protein (e.g. Ensembl accession number ENSMUSG00000010592). TPX2 microtubule associated is a RNA-binding protein, which is essential for gametogenesis in both males and females. Plays a central role during spermatogenesis. Acts by binding to the 3'-UTR of mRNA, specifically recognizing GUU triplets, and thereby regulating the translation of key transcripts.

As used herein *Rad9a* refers to a gene encoding RAD9, a checkpoint clamp component A (e.g. Ensembl accession number ENSG00000172613). A checkpoint clamp component A is a component of the 9-1-1 cell-cycle checkpoint response complex that plays a major role in DNA repair. The 9-1-1 complex is recruited to DNA lesion upon damage by the RAD17-replication factor C (RFC) clamp loader complex. It acts then as a sliding clamp platform on DNA for several proteins involved in long-patch base excision repair (LP-BER). The 9-1-1 complex stimulates DNA polymerase beta (POLB) activity by increasing its affinity for the 3'-OH end of the primer-template and stabilizes POLB to those sites where LP-BER proceeds. The 9-1-1 complex is necessary for the recruitment of RHNO1 to sites of double-

stranded breaks (DSB) occurring during the S phase. RAD9A possesses 3'->5' double stranded DNA exonuclease activity. Its phosphorylation by PRKCD may be required for the formation of the 9-1-1 complex.

As used herein *Pms2* refers to a gene encoding postmeiotic segregation increased 2 protein (e.g. Ensembl accession number ENSG00000122512). PMS2 is a component of the post-replicative DNA mismatch repair system (MMR). It heterodimerizes with MLH1 to form MutL alpha. DNA repair is initiated by MutS alpha (MSH2-MSH6) or MutS beta (MSH2-MSH6) binding to a dsDNA mismatch, then MutL alpha is recruited to the heteroduplex. Assembly of the MutL-MutS-heteroduplex ternary complex in presence of RFC and PCNA is sufficient to activate endonuclease activity of PMS2. It introduces single-strand breaks near the mismatch and thus generates new entry points for the exonuclease EXO1 to degrade the strand containing the mismatch. DNA methylation would prevent cleavage and therefore assure that only the newly mutated DNA strand is going to be corrected. MutL alpha (MLH1-PMS2) interacts physically with the clamp loader subunits of DNA polymerase III, suggesting that it may play a role to recruit the DNA polymerase III to the site of the MMR. PMS2 has also been implicated in DNA damage signaling, a process which induces cell cycle arrest and can lead to apoptosis in case of major DNA damages.

As used herein *Mlh1* refers to a gene encoding mutL homologue 1 (e.g. Ensembl accession number ENSG00000076242). MLH1 heterodimerizes with PMS2 to form MutL alpha, a component of the post-replicative DNA mismatch repair system. DNA repair is initiated by MutS alpha (MSH2-MSH6) or MutS beta (MSH2-MSH6) binding to a dsDNA mismatch, then MutL alpha is recruited to the heteroduplex. Assembly of the MutL-MutS-heteroduplex ternary complex in presence of RFC and PCNA is sufficient to activate endonuclease activity of PMS2. It introduces single-strand breaks near the mismatch and thus generates new entry points for the exonuclease EXO1 to degrade the strand containing the mismatch. DNA methylation prevents cleavage and therefore assures that only the newly mutated DNA strand is going to be corrected. The function of MutL alpha (MLH1-PMS2) has been described above under the PMS2. Furthermore, MLH1 heterodimerizes with MLH3 to form MutL gamma which plays a role in meiosis.

As used herein *Cenpe* refers to a gene encoding centromere protein E (e.g. Ensembl accession number ENSG00000138778). CENPE plays an important role in

chromosome congression, microtubule-kinetochore conjugation and spindle assembly checkpoint activation. It drives chromosome congression (alignment of chromosomes at the spindle equator resulting in the formation of the metaphase plate) by mediating the lateral sliding of polar chromosomes along spindle microtubules towards the spindle equator and by aiding the establishment and maintenance of connections between kinetochores and spindle microtubules. The transport of pole-proximal chromosomes towards the spindle equator is favored by microtubule tracks that are detyrosinated. CENPE plays an important role in the formation of stable attachments between kinetochores and spindle microtubules. The stabilization of kinetochore-microtubule attachment also requires CENPE-dependent localization of other proteins to the kinetochore including BUB1B, MAD1 and MAD2. CENPE plays a role in spindle assembly checkpoint activation (SAC) via its interaction with BUB1B resulting in the activation of its kinase activity, which is important for activating SAC. CENPE is necessary for the mitotic checkpoint signal at individual kinetochores to prevent aneuploidy due to single chromosome loss.

As used herein *Ncapd3* refers to a gene encoding non-SMC condensing II complex subunit D3 (e.g. Ensembl accession number ENSG00000151503). NCAPD3 is a regulatory subunit of the condensin-2 complex, a complex which establishes mitotic chromosome architecture and is involved in physical rigidity of the chromatid axis.

As used herein *Odf2* refers to a gene encoding outer dens fiber of sperm tails 2 (e.g. Ensembl accession number ENSG00000136811). ODF2 is a major component of sperm tail outer dense fibers (ODF). ODFs are filamentous structures located on the outside of the axoneme in the midpiece and principal piece of the mammalian sperm tail and may help to maintain the passive elastic structures and elastic recoil of the sperm tail. ODF2 may a modulating influence on sperm motility and functions as a general scaffold protein that is specifically localized at the distal/subdistal appendages of mother centrioles. ODF2 is a component of the centrosome matrix required for the localization of PLK1 and NIN to the centrosomes. ODF2 is required for the formation and/or maintenance of normal CETN1 assembly.

As used herein *Dclre1b* refers to a gene encoding DNA cross-link repair 1B protein (e.g. Ensembl accession number ENSG00000118655). DCLRE1B is 5'-3' exonuclease that plays a central role in telomere maintenance and protection during S-phase. It participates in the protection of telomeres against non-homologous end-joining (NHEJ)-mediated repair, thereby ensuring that telomeres do not fuse. DCLRE1B

also plays a key role in telomeric loop (T loop) formation by being recruited by TERF2 at the leading end telomeres and by processing leading-end telomeres immediately after their replication via its exonuclease activity: generates 3' single-stranded overhang at the leading end telomeres avoiding blunt leading-end telomeres that are vulnerable to end-joining reactions and expose the telomere end in a manner that activates the DNA repair pathways. Together with TERF2, DCLRE1B is required to protect telomeres from replicative damage during replication by controlling the amount of DNA topoisomerase (TOP1, TOP2A and TOP2B) needed for telomere replication during fork passage and prevent aberrant telomere topology. Also DCLRE1B is involved in responses to DNA damage.

Said polypeptides or proteins encoded by the mentioned genes refer to not only human proteins but also to any other homologue from any animal. Also, any variants of said polypeptides or proteins are included (e.g. multiple transcript variants resulting from alternative splicing. Proteins encoded by said genes are described in scientific articles and are well known to a skilled person.

Now interestingly a very specific combination of genes has been found relevant for determining the risk of cancer. In one embodiment decreased expression levels of very specific genes have been found to precede cancer. As an example, in the study of the present disclosure expression of *Mlh1* decreased similarly in samples of subjects having Lynch syndrome and subject not having Lynch syndrome. Decreased expression of *Mlh1* and/or *Pms2* was shown to predict cancer in subjects with or without an inherited predisposition for cancer.

In the present disclosure, figure 4 shows the expression levels of (at least three) 10 chromosomal segregation-specific genes in colon mucosa. Expression of each gene in each sample obtained from the normal mucosa of colorectal cancer mice was independently compared to the average expression level of the same genes in mucosa samples of non-carcinoma (healthy) mice. If expressions of only one or two genes (i.e. markers) are determined, then the normal mucosa of colorectal cancer mice and non-carcinoma mice cannot be differentiated from each other at a reliable level based on the expression data. Classification of mucosa samples to i) cancer predisposing mucosa or ii) non-cancer predisposing mucosa becomes possible with at least three markers (clustering is at a reliable level) and further improves with at least four or more markers (see Figures 7 and 8, and Table 1).

In the present invention the expression of at least three specific genes is determined from a colon mucosa sample. In a specific embodiment of the invention the sample is a histologically normal colon mucosa sample. As used herein "the histologically normal mucosa" refers to the colonic mucosa, which is folded in parts of colon and is relatively thin within the folds. Histologically normal mucosa has the following characteristics. Lamina propria is compact and crypts of Lieberkühn are shallow and straight. Surface epithelium is smooth and the epithelial cells palisade evenly. Goblet cells are abundant in the surface epithelium and throughout crypts in the proximal colon, relatively diminishing in number towards the base of the crypts in the distal colon. Some leucocytes are present in the lamina propria: Lymphocytes are most abundant cell type with some eosinophils and occasional macrophages and mast cells. Neutrophils are generally absent. Tela submucosa is generally inconspicuous. The thickness of the muscularis externa varies. Said histologically normal mucosa may be detected visually e.g. by staining methods and/or microscope (such as a light microscope).

As used herein "mucosa" refers to a membrane that lines colon. It consists of one or more layers of epithelial cells overlying a layer of loose connective tissue and is mostly of endodermal origin. In a very specific embodiment the sample is a mucosa sample of the proximal colon. Colon i.e. the large intestine is the last part of the gastrointestinal tract and of the digestive system in vertebrates.

In one embodiment of the invention a sample of a colon mucosa is obtained from a subject for the method of the present invention. Said sample may be obtained by any method well known to a person skilled in the art including but not limited to colonoscopy, sigmoidoscopy and surgical operations. The most common way to obtain a biopsy of the colon is through a colonoscopy or sigmoidoscopy. The bowel is first cleansed by appropriate procedures. A probe (colonoscope) is then inserted through the rectum to the colon to visualize the interior. The flexible probe makes it possible to pass an instrument that can clip a tiny piece of tissue (approximately of the size of a pin). Intestinal mucosa sampling by colonoscopy is a widely-used procedure for various medical conditions. The associated risks are generally low, but may include bleeding, infection, or perforation.

"Cancer" refers to a cell or cells having abnormal growth with the potential to invade or spread to other parts of the body. Cancers are classified by the type of a cell to

be the origin of the tumor. Cancers include but are not limited to carcinomas, sarcomas, lymphomas and leukemias, germ cell tumors, and blastomas. Carcinoma refers to a cancer arising from epithelial cells. In a specific embodiment of the invention the cancer is a carcinoma. In a very specific embodiment of the invention the cancer is a colon carcinoma or a colon carcinoma of the proximal colon. In one embodiment the carcinoma is selected from the group consisting of tubular carcinoma, mucinous carcinoma, tubulovillous carcinoma, serrated carcinoma and adenocarcinoma.

In one embodiment of the invention a subject is a human or an animal such as a mammal. The human may be a child, an adolescent or an adult. Any animal, such as a pet, domestic animal or production animal may be a subject of the present invention.

In a specific embodiment of the invention a subject is in need of the method or tools of the present invention. As an example the subject may be susceptible of cancer. In one embodiment of the invention a subject is at a low or high risk for developing cancer. As used herein a subject at a high risk is a subject of having e.g. a disease or inherited mutation increasing the risk of cancer. In one embodiment a subject is a human patient diagnosed with an inherited mutation such as one predisposing to colon cancer (e.g. Lynch syndrome, familial adenomatous polyposis). In one embodiment the subject may have any symptoms (such as pain, fever) or e.g. may have suffered from a tumor or cancer, which has been cured. In another embodiment a risk of a human or an animal for developing cancer is not determined before said human or an animal is subject to the method of the present invention. In a very specific embodiment of the invention “determining whether a subject is at risk to develop cancer” refers to determining specific biomarkers indicative of cancer development or preceding cancer.

In another embodiment the method of the present invention may be carried out for screening subjects without any suspicion of cancer. Indeed, the method and tools of the present invention may be used for screening any subject and thus, the subject may be healthy and asymptomatic.

Before classifying a subject as suitable for the method of the present invention, the clinician may for example study any symptoms or assay any disease markers of the

subject. Based on the results either being normal or deviating from the normal, the clinician may suggest the method of present invention for the subject.

It is a further object of the present invention to provide a kit comprising the necessary reagents for performing a method according to the present invention. A kit according to the present invention includes standard reagents, such as tools to determine the expression level of at least five genes associated with chromosomal segregation, one or more control samples e.g. representing the normal expression level of the genes of interest, and reagents for performing determination of expression levels of said genes.

According to the present invention it may be possible or advantageous to determine expressions of at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9 or at least 10 genes presented in Figures 4A and B by tools included in a kit.

In one embodiment of the invention the kit comprises tools to determine the expression levels of at least three genes associated with chromosomal segregation *Bub1*, *Mis18a* and *Tpx2*. In another embodiment at least three genes associated with chromosomal segregation are selected from the following groups Bub1, Mis18a and Rad9a; Bub1, Tpx2 and Rad9a; Bub1, Mis18a and Pms2; Bub1, Tpx2 and Pms2; Bub1, Rad9a and Pms2. Optionally any other genes or any combination thereof (e.g. selected from the group consisting of *Bub1*, *Mis18a*, *Tpx2*, *Rad9a*, *Pms2*, *Mlh1*, *Cenpe*, *Ncapd3*, *Odf2* and *Dclre1b*) may be determined in addition to the combination of said three genes.

In one embodiment of the invention the kit comprises tools to determine the expression levels of at least four genes associated with chromosomal segregation *Bub1*, *Mis18a*, *Tpx2* and *Rad9a*, or *Bub1*, *Mis18a*, *Tpx2* and *Pms2*. In another embodiment at least four genes associated with chromosomal segregation are selected from the following groups: Bub1, Mis18a, Rad9a and Pms2; Bub1, Tpx2, Rad9a and Pms2. Optionally any other genes or any combination thereof (e.g. selected from the group consisting of *Bub1*, *Mis18a*, *Tpx2*, *Rad9a*, *Pms2*, *Mlh1*, *Cenpe*, *Ncapd3*, *Odf2* and *Dclre1b*) may be determined in addition to the combination of said four genes.

In a further embodiment of the invention the kit comprises tools to determine the expression levels of at least five genes associated with chromosomal segregation

Bub1 (BUB1, mitotic checkpoint serine/threonine kinase), *Mis18a* (MIS18 kinetochore protein A), *Tpx2* (TPX2 microtubule associated), *Rad9a* (RAD9 checkpoint clamp component A) and *Pms2* (postmeiotic segregation increased 2). Optionally one, two, three, four, five or more other genes (e.g. selected from the group consisting of *Mlh1*, *Cenpe*, *Ncapd3*, *Odf2* and *Dclre1b*) may be determined in addition to any combination of said five genes.

In another embodiment the kit comprises tools to determine the expression level of (at least) three, four, five, six, seven, eight, nine or ten genes associated with chromosomal segregation selected from the group consisting of *Bub1*, *Mis18a*, *Tpx2*, *Rad9a*, *Pms2*, *Mlh1*, *Cenpe*, *Ncapd3*, *Odf2* and *Dclre1b*. Any specific combinations of said genes are included within the scope of the invention. In a very specific embodiment the kit comprises tools to determine the expression level of at least the genes associated with chromosomal segregation selected from the following groups:

at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Mlh1;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Cenpe;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Ncapd3;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Odf2;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Dclre1b;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Mlh1, Cenpe;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Mlh1, Cenpe, Ncapd3;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Mlh1, Cenpe, Ncapd3, Odf2;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Mlh1, Cenpe, Ncapd3, Odf2, Dclre1b;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Mlh1, Ncapd3;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Mlh1, Ncapd3, Odf2;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Mlh1, Ncapd3, Odf2, Dclre1b;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Mlh1, Odf2;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Mlh1, Odf2, Dclre1b;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Mlh1, Dclre1b;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Cenpe, Ncapd3;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Cenpe, Ncapd3, Odf2;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Cenpe, Ncapd3, Odf2, Dclre1b;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Cenpe, Odf2;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Cenpe, Odf2, Dclre1b;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Cenpe, Dclre1b;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Ncapd3, Odf2;*

at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Ncapd3, Odf2, Dclre1b;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Ncapd3, Dclre1b;*
at least *Bub1*, *Mis18a*, *Tpx2*, *Rad9a, Pms2, Odf2, Dclre1b*.
In a still further embodiment the kit for comprises tools to determine the expression level of at least ten genes associated with chromosomal segregation *Bub1*, *Mis18a*, *Tpx2*, *Rad9a*, *Pms2*, *Mlh1*, *Cenpe*, *Ncapd3*, *Odf2* and *Dclre1b*.

In some embodiments of the present invention specific primers or probes hybridize to any region of the RNA (e.g. mRNA) representing the gene of interest and thus enable determination of the expression level of said gene.

In some embodiments one or more control samples may be obtained from any control subject such as a human or an animal depending of the nature of the method. Optionally positive control samples showing decreased or increased expression levels compared to control samples may also be utilized in the present invention. Also, a quality control of the method may optionally be comprised within the kit.

In a specific embodiment the kit comprises the reference levels (i.e. cut off levels) of suitable subjects.

Optionally, a kit according to the present invention may also comprise additional reagents necessary for performing the method of the present invention, such as necessary buffers and enzymes. Optionally, the kit may comprise further reagents such as one or more reagents selected from the group consisting of RNeasy Plus Mini Kit (e.g. Qiagen), SuperScript Vilo cDNA Synthesis Kit (e.g. Thermo Fisher Scientific), gene specific TaqMan assay reagents, TaqMan™ Gene Expression Master Mix (e.g. Thermo Fisher Scientific).

One example of a kit according to the present invention is given in the examples below comprising tools to determine expression levels of specific genes (e.g. probes and/or primers) as well all reagents and control samples necessary for carrying out determination of said expression levels.

In a specific embodiment the kit comprises instructions for carrying out a method for determining expression levels of specific genes or determining whether a subject is at risk to develop cancer.

In a very specific embodiment of the invention the kit comprises tools to determine the expression levels of at least three, four or five genes associated with chromosomal segregation, reagents for performing said method, the reference levels (i.e. cut off levels) of suitable subjects, instructions for carrying out a method for determining expression levels of specific genes or determining whether a subject is at risk to develop cancer and optionally one or more control samples.

The following examples are given to further illustrate embodiments of the present invention, but are not intended to limit the scope of the invention. It will be obvious to a person skilled in the art, as technology advances, that the inventive concept can be implemented in various ways. The invention and its embodiments are thus not limited to the examples described herein, but may vary within the scope of the claims.

EXAMPLES

Materials and methods

Mice, experimental study and diets

Heterozygote B6.129-*Mlh1*^{tm1Rak} mice (*Mlh1*$^{+/-}$) strain 01XA2 (46) and the C57BL/6 strain were obtained from NCI-MMHCC; National Institutes of Health, Mouse Repository, NCI-Frederick, MD. Altogether 12 animals (equal numbers of sexes), the *Mlh1*$^{+/-}$ mice and their wild-type C57BL/6 mates, formed six breeder pairs which produced the mouse colony used in our study. Mice were genotyped (Fig. 1) using genomic DNA extracted from earmarks according to the protocol published in our previous work(23). The mice were bred and treated according to the study protocol approved by the National Animal Experiment Board in Finland (ESLH-2008-06502/Ym-23).

The *Mlh1* heterozygote mice and their homozygote wild type littermates were divided into two dietary groups at the age of 5 weeks. The mice were fed with either healthy rodent control diet AIN-93G (47) or Western-style diet modified from AIN (Harlan Teklad, Madison, WI)(23) to resemble, on the nutritional level, the diet consumed in human Western population (high fat and energy content, low amounts of fiber, calcium and vitamin D_3).(23) Twelve mice per each group (*Mlh1*$^{+/+}$ AIN, *Mlh1*$^{+/-}$ AIN, *Mlh1*$^{+/+}$ WD, *Mlh1*$^{+/-}$ WD) with equal representation of sexes, at time point (tp)

0 (5 weeks of age, *Mlh1*$^{+/-}$, *Mlh1*$^{+/+}$), tp1 (12 mo of age), tp2 (18 mo of age) and tp3 (21 mo of age), 168 mice in total were sacrificed and sampled.

Collection of tumors and normal colon mucosa samples

All observed colon tumors were collected under dissecting microscope and preserved as FFPE blocks. If a tumor was large enough (three to five mm in diameter), approximately half of it was embedded in O.C.T compound (VWR, Radnor, Pennsylvania) for cryo sampling. Histological studies, stainings and the grading of neoplasias were carried out at The Finnish Centre for Laboratory Animal Pathology (FCLAP), University of Helsinki, Finland. The neoplasias were graded as hyperplasias, adenomas and carcinomas according to criteria based on consensus rodent intestinal cancer nomenclature(48), (**Hyperplasia:** Epithelial hyperplasia was characterized by localized increase in mucosal thickness, long uniform crypts with increased cell proliferation, and no atypia. **Adenoma:** Adenomas were classified either tubular or tubulovillous/papillary adenomas, and the degree of dysplasia was graded in low- and high-grade dysplasia. In adenomas with low-grade dysplasia, distorted and irregularly distributed crypt structures were composed of proliferating epithelium showing nuclear crowding and pseudostratification. The nuclear features of low-grade dysplasia consisted of mild to moderate increase in nuclear size, granular chromatin and discernible nucleoli. In adenomas with high-grade dysplasia, the crypt structures were disorganized and often packed, with cribriform areas. The cells showed increased atypia, irregular nuclei with coarse chromatin and enlarged, conspicuous nucleoli. **Carcinoma:** In carcinomas, there were signs of invasion. Some of the cases were early cancers, with limited submucosal invasion. Invasion was characterized by displacement of malignant glands between the muscularis mucosa, submucosal vascular structures or by pushing border –type invasion with tumor extending to submucosal level or beyond. Longitudinal pieces (excluding the previously harvested tumorous sections), representing approximately one third of the proximal mouse colon were collected for normal mucosa. The mucosa was separated from the underlying submucosa and musculature under a dissecting microscope and samples for RNA extraction were stored in RNAlater (Qiagen, Hilden, Germany) at -80°C.

Transcriptome analysis of normal mucosa

Transcriptome analysis was performed using RNA-sequencing method (RNA-seq). Total RNA was prepared from 0, 12 and 18 mo old mice (14, 40 and 40 mice respectively) using the RNeasy Plus Kit (Qiagen, Hilden, Germany) with an extra DNase treatment (Qiagen, Hilden, Germany). The RNA concentration was measured by Qubit 1.0 (Thermo Fisher Scientific, Waltham, MA, USA) and RNA integrity with the Agilent 2100 Bioanalyzer (Agilent technologies, Santa Clara, CA). Only high quality RNA (RNA integrity number RIN ≥ 8) qualified for expression analysis.

RNA-seq method followed the single-cell tagged reverse transcription (STRT)(49) protocol with modifications(22). Briefly, 10 ng of total RNA was converted to cDNA and amplified to form an Illumina-compatible library. In total, 25 PCR cycles were used, but as four base-pair unique molecular identifiers were applied, only the absolute number of unique reads were included in the subsequent analysis. The samples were sequenced on a total of six lanes of Illumina HiSeq2000, further processed to fastq files by Casava 1.8.2 (both Illumina, San Diego, CA, USA). Quality control was performed using the STRTprep pipeline (https://github.com/shka/STRTprep)(22). The processed reads were aligned by TopHat2(50) to the mouse RefSeq mm9 reference genome. STRT captures sequences at the 5'-end of poly(A)+ RNAs and the aligned reads therefore tend to be distributed close to the 5'-end (start site) of genes. STRTprep counts only the aligned reads at the 5'-untranslated region of protein-coding genes, or within the proximal (500 bp) upstream region.

Normalizing the RNA-seq data

STRTprep pipeline generated a read count matrix, with genes as rows and samples as columns. Different sample library sizes were normalized using DESeq-style normalization(51). Next shifted log transformation ($x_{log} = \log(x + 1)$) was done to generate more Gaussian like data and the ComBat program(52) was used to filter batch effects. These preprocessing steps and alternative pipelines were evaluated by looking at the hierarchical clustering of samples and by plotting quantiles of expression values for each sample.

Tests for Differential Gene Expression

Since the analyzed data was not any more integer count values after ComBat normalization, we tested three T-test based methods, Voom-Limma, Cyber-T and

Shrinkage-T (53-55), for analysis of differential gene expression. All these methods add a prior to variance estimate. Shrinkage-T is the only method here that allows also testing with unequal variance. This turned out to be important, as the genes with strongest separation between the sample groups had small variance in the analyzed subset and medium variance among remaining samples.

The three methods were evaluated by viewing the separation of cancer samples from the remaining samples in the Multi-dimensional Scaling (MDS) plots with top-k genes, which were selected using the evaluated statistic. Parameter k was varied from 25 to few hundreds. Shrink-T showed the best separation in the generated plots across all values of k. Each methods ability to find correlations with Gene Ontology classes was also tested. We used T-test scores from each method separately as an input to enrichment analysis tool called GSZ (Gene Set Z-score) (56). Shrink-T again generated strongest results. We therefore applied only Shrink-T in the subsequent analyses.

Visualization of sample differences

To detect similarities and differences between the samples we used MDS that generates a small-dimensional visualization from the multidimensional data while trying to preserve the pairwise distances of samples from the multi-dimensional data. Plot-MDS distributed in the Limma package was used as a basis of the analysis (57), although modified so that we were able to use any selected score to pick the genes that were used to calculate pair-wise distances.

The activity of *Mlh1* was visualized with ComBat normalized data. Samples were grouped based on the sample types (genotype, diet and time-point) to highlight the sample differences.

Pathway analysis

To study the biological functions and pathways enriched among the top separating genes we used QIAGEN´s Ingenuity Pathway Analysis (IPA Software 7.0, Qiagen, Hilden, Germany). Here, we analyzed both the top 100 and top 300 genes, which were found to separate the normal mucosa expression patterns in carcinoma mice from the others. The settings for a core analysis were as follows: ingenuity knowledge base (genes and endogenous chemicals) with both direct and indirect

relationships, default network interaction settings (include endogenous chemicals, 35 molecules per network and 25 networks per analysis). Data sources were used with stringent confidence (experimentally observed and high predicted) and data obtained in all species was selected with a relaxed filter.

MSI and LOH analyses

The microsatellite instability status was analyzed from seven carcinomas (two *Mlh1*$^{+/+}$ WD, four *Mlh1*$^{+/-}$ WD, and one *Mlh1*$^{+/+}$ AIN mice) using four dinucleotide (D18Mit15, D14Mit15, D10Mit2, D7Mit91) and two mononucleotide (JH104, U12235) markers(58). Tumor DNA samples were extracted from the cryo-preserved colon carcinomas using laser micro-dissection for cutting (Zeiss PALM MicroBeam, Carl Zeiss Microscopy GmbH, Jena, Germany) and normal DNA control samples from the tails of the same mice with QIAamp DNA micro Kit, and DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany), respectively. The genomic DNA was amplified with 6-FAM labeled primers in 11.1X PCR master mix (59) using the following PCR protocol: 1 min at 96°C, 30 cycles of 20 s at 96°C, 30 s at 62°C, and 15 s at 70°C, and 7 min at 70°C. The fragments were analyzed with ABI3730xl capillary electrophoresis (Thermo Fisher Scientific, Waltham, MA, USA) and visualized with PeakScanner v1.0 (Thermo Fisher Scientific, Waltham, MA, USA).

The four colon carcinomas found in the heterozygote *Mlh1*$^{+/-}$ mice, of which the cryo sample was available, were also studied for loss of *Mlh1.* Loss of heterozygosity was analyzed using the genotyping primers M001 (TGT CAA TAG GCT GCC CTA GG, SEQ ID NO: 1), M002 (TGG AAG GAT TGG AGC TAC GG, SEQ ID NO: 2), and M003 (TTT TCA GTG CAG CCT ATG CTC, SEQ ID NO: 3), which produce two different length fragments, 350 bp and 500 bp, separating the normal (M001/M003) and the mutated allele (M001/M002), respectively(23) (Fig. 1). DNA was amplified with the 11.1X PCR master mix as described above and the fragments were visualized on 1% SB agarose gel.

Immunohistochemical analysis of Mlh1 protein expression in carcinomas

Formalin-fixed, paraffin embedded cancer tissue blocks were studied for Mlh1 expression. The 4 µm thick sections were deparaffinized and rehydrated and heat induced antigen retrieval was performed with 10 mM citrate buffer (pH 6). To detect Mlh1, the slides were incubated overnight at 4°C with the rabbit monoclonal antibody

ab92312 (1:1500) (clone EPR3894, Abcam, Cambridge, UK). Stainings were visualized using UltraVision Detection System anti-rabbit HRP/DAB (ThermoFisher Scientific, Waltham, MA USA) by manufacturer´s instructions. Analysis of staining patterns was conducted at The Finnish Centre for Laboratory Animal Pathology.

Analysis of mitoses in carcinomas

A Feulgen with Midori green background stain was used to visualize nuclear material and mitoses in six carcinoma samples (E249, E314, E329, E333, E338, and E347). The samples were deparaffinized and rinsed in 1 M HCl. Mild acid hydrolysis was accomplished by using 60°C 1 M HCl and DNA was stained purple in Schiff´s reagent for 45 min. After several bisulfite washes the samples were counterstained briefly with 1% Midori light green, dehydrated through alcohol series to xylene and mounted with xylene based mounting media. The stained samples were analyzed under light microscope (Zeiss Axio Imager.A2, Carl Zeiss Microscopy GmbH, Jena, Germany) and the mitoses in the malignant areas of carcinomas were compared to mitoses in samples (E305, E311, E322, E323, and E346) from healthy mice.

Statistical Analysis

Differential Expression Analysis (DEA) used modified t-tests (Limma, cyber-T and shrinkage t). With Limma and cyber-T we used their own p-value estimates. Shrinkage-T does not provide a p-value estimate, which were estimated by re-calculating Shrinkage-T with 1000 permutations for each gene separately. Normal distribution was fitted to the permutations and a one-tailed p-value was obtained from the cumulative distribution. Multiple testing correction was performed using False Discovery Rate. Importantly, we used DEA mainly to order the genes to most differentially regulated genes. All analysis was performed within the R-environment. Pathway enrichment analysis was done using IPA which uses Fisher's exact test to analyze over-representation of genes from the analyzed gene groups. Here, multiple testing correction was done using the Benjamin-Hochberg method.

Human samples

Biopsies of the colon mucosa were obtained by through a colonoscopy or sigmoidoscopy. The bowel was first cleansed by appropriate procedures. A probe (co-

lonoscope) was then inserted through the rectum to the colon to visualize the interior. The flexible probe made it possible to pass an instrument that can clip a tiny piece of tissue (approximately of the size of a pin). Intestinal mucosa sampling by colonoscopy is a widely-used procedure for various medical conditions. The associated risks are generally low, but may include bleeding, infection, or perforation.

Methods carried out with mice samples as described above are also utilized for human colon mucosa samples.

Results

Carcinomas developed mainly and earlier in Western-style diet fed mice

The feeding study was done with offsprings produced by crossing two isogenic strains, the heterozygote *Mlh1*$^{+/-}$ (B6.129-*Mlh1*^{tm1Rak}) and the wild-type *Mlh1*$^{+/+}$ (C57BL/6) mice, and selecting an equal number of both genotypes to the study. Half of the mice fed Western-style diet and half the control diet, AIN-93G. In all 168 mice, 24 mice at time point 0 and 48 mice at time points 12, 18 and 21 month (mo), were operated. WD, high in energy and animal fat and low in fiber and nutrients, seemed to be a severe risk factor for CRC, since approximately 80% of all colon tumors, 10 out of 13 colon adenocarcinomas and 14 out of 20 adenomas and hyperplasias, developed in WD fed mice (Figure 6 representing a table).

At time points 12, 18 and 21 mo, 80%, 78% and 64% of all tumors and 100%, 80% and 72% of CRCs were found in WD fed mice, respectively, indicating that Western-style diet also accelerates the progression of carcinogenesis. The overall number of colon tumors increased significantly with time (Fig. 2), being five at 12 mo (one adenocarcinoma, two adenomas, one hyperplasia), nine at 18 mo (five adenocarcinomas, two adenomas, two hyperplasias), and nineteen at the 21 mo time point (seven adenocarcinomas, five adenomas, seven hyperplasias). Tumors were approximately evenly distributed between different genotypes since heterozygote *Mlh1*$^{+/-}$ mice showed 0%, 40% and 43% of carcinomas and 50%, 75% and 42% of adenomas and hyperplasias at different time points (Figure 6 representing a table). However, 75% (15/20) of adenomas and hyperplasias at 18 mo indicate that *Mlh1* heterozygosity accelerates their progression. All the 13 carcinomas were found in the proximal part of colon and the majority of them were either tubular (54%) or mucinous (31%), two were tubulovillous and one carcinoma had serrated histology.

Mlh1 mutation carriers did not show MSI, LOH and loss of MMR protein in tumors

To check for the typical Lynch syndrome characteristics, seven carcinomas found in 18 and 21 mo old mice, four in the *Mlh1*$^{+/-}$ mice (E338, E347, E437, E444) and three in the *Mlh1*$^{+/+}$ mice (E402, E410, E421), were analyzed for MSI status and Mlh1 expression. Surprisingly, all CRCs showed Mlh1 expression (Fig. 1), indicating that irrespective of the inherited mutation in one *Mlh1* allele in the heterozygote mice, the normal allele was still present in the tumors. The presence of the normal allele was further confirmed with LOH study in all four *Mlh1*$^{+/-}$ carcinomas (Fig. 1). To study whether the detected Mlh1 protein was functional and MMR proficient, we analyzed the stability of six polymorphic microsatellite regions in the mouse genome. The markers and their amplified fragment sizes were as follows: D14Mit15 (148 bp, 150 bp), D18Mit15 (151 bp, 157 bp), D7Mit91 (139 bp, 147 bp), D10Mit2 (117 bp, 122 bp), JH104 (178 bp, 181 bp), and U12235 (79 bp, 83 bp). Altogether six out of seven CRCs were studied (E410 could not be amplified) and shown to be microsatellite stable, since no differences in the fragment lengths were observed between the tumor and corresponding normal DNA.

Mlh1 RNA expression was significantly decreased in normal mucosa of CRC mice

After finding that irrespective of the mouse genotype, functional Mlh1 was still expressed in carcinomas, our interest was to look for potential early drivers of tumorigenesis on a genome-wide scale. Genome-wide transcriptome analysis was performed from 80 normal colon mucosa samples operated from 12 mo and 18 mo old mice. Analysis was done with RNAseq using the single-cell tagged reverse transcription method (STRT)(21, 22). The 21 mo old mice were left out from the RNAseq study due to many health problems most probably because of their old age. Altogether 12 216 expressed transcripts were identified in the samples. First, we analyzed the *Mlh1* gene expression levels from the STRT data. In our previous study(23) we showed that in the beginning of the feeding experiment (at time point 0), the *Mlh1* heterozygote mice showed exactly 50% lower *Mlh1* expression than the *Mlh1*$^{+/+}$ mice. Contrary to varying *Mlh1* expression levels in mice in general, 5/6 mice (E249, E314, E329, E333, E338) who developed carcinoma showed remarkably low *Mlh1* expression in their normal colon mucosa (P=0.03). The mouse E347

whose carcinoma had serrated histology had similar *Mlh1* expression level as non-carcinoma mice on average (Figure 6 representing a table).

Expression profiles in normal mucosa formed a distinct cluster for CRC mice

After finding that carcinoma mice had extremely low levels of *Mlh1* transcripts in their mucosa, we next compared their genome-wide expression profiles with profiles of all other 12 and 18 mo old mice. The normal mucosa expression profiles of the six CRC mice were strikingly different from the profiles of the other mice and formed a distinct cluster as visualized by an MDS plot created with the 100 most altered/differentially regulated genes (Figures 3A and 3B). Altogether 86% of the top 300 differentially regulated genes in CRC mice were down-regulated and 14 % were up-regulated.

Pathway analysis and shortage of chromosomal segregation gene-specific transcripts suggest problems in cell cycle regulation and mitosis

To further understand the biological functions and pathways enriched among the top separating genes in CRC mice, the expression data were analyzed with Ingenuity Pathway Analysis (IPA). According to IPA, chromosome segregation (P=2.92×10^{-5}), aneuploidy of fibroblasts (P=5.31×10^{-4}), checkpoint control (P=*1.10*×10^{-4}), DNA replication checkpoint (P=1.88×10^{-4}) and morphology of mitotic spindle (P=6.45×10^{-5}) were among the most affected biological functions. In network analysis the most affected molecular and cellular functions included cell cycle (P=9.24×10^{-5}), cellular assembly and organization (P=9.24×10^{-5}), DNA replication, recombination and repair (P=9.24×10^{-5}), cell death and survival (P=7.30×10^{-5}), and cellular growth and proliferation (P=3.07×10^{-3}). The analysis was also repeated with different RNA-seq data preprocessing (all mouse samples without ComBat normalization). These results confirmed our findings on chromosome segregation (P=1.03×10^{-5}), aneuploidy of fibroblasts (P=4.57×10^{-4}) and checkpoint control (P=*4.29*×10^{-4}).

The IPA results strongly indicated that there are severe problems in cell cycle regulation and mitosis already in colon mucosa. In the six mice who developed carcinoma up to 18 mo, the most altered/differentially expressed genes that pointed to chromosome segregation and spindle assembly checkpoint (SAC) were *Bub1* (BUB1, mitotic checkpoint serine/threonine kinase), *Mis18a* (MIS18 kinetochore

protein A), *Tpx2* (TPX2 microtubule associated), *Rad9a* (RAD9 checkpoint clamp component A), *Pms2* (postmeiotic segregation increased 2), *Mlh1* (mutL homolog 1, along with MMR function also triggers checkpoint activation), *Cenpe* (centromere protein E), *Ncapd3* (non-SMC condensing II complex subunit D3), *Odf2* (outer dens fiber of sperm tails 2) and *Dclre1b* (DNA cross-link repair 1B). Five of these 10 genes, *Bub1*, *Mis18a*, *Tpx2*, *Rad9a* and *Pms2,* were strongly down regulated in all of the six carcinoma mice (Fig. 4). *Mlh1*, *Cenpe*, *Ncapd3*, *Odf2* and *Dclre1b* showed variable level of down regulation in two CRC mice (E347 and E249). In E249, *Mlh1* and *Dclr1b* showed, however, approximately 50% down regulation when compared to the average expression level in non-carcinoma mice (Fig. 4). In E347, whose carcinoma was histologically different from the others and showed typical serrated phenotype, *Mlh1*, *Cenpe*, *Ncapd3* and *Odf2* expression levels were equal to the non-carcinoma mice. Importantly, among all the 74 mice that did not develop colon carcinoma up to 18 mo, only two mice, E325 and E332, showed similar low expression of all the 10 genes (Fig. 4). Although, no colonic tumors were found in those mice, E325 had bloody feces and anemia, suggesting undefined mucosal pathology.

Abnormal mitoses and chromosomal instability in carcinomas

Undisturbed mitosis is a central requirement of the normal cell cycle and division. In cancer cells, mitoses are often aberrant, showing aneuploidy caused by unequal segregation of chromosomes and/or structural changes in chromosomes, both of which lead to chromosomal instability. To validate the RNA sequencing results, which suggested impaired cell cycle regulation and mitosis in CRC mice, all the 13 carcinomas were stained with feulgen and analyzed for mitotic aberrations. Although all the carcinomas were well-differentiated early cancers with limited submucosal invasion and relatively lenient cytological changes, they exhibited increased mitotic activity and abundant numbers of unbalanced/atypical mitoses in contrast to normal tissue samples (Fig. 5).

Conclusions

Mlh1 protein expression was studied in colon tumors and *Mlh1* gene expression in histologically normal mucosa. Approximately 70% of all tumors and 80% of colon carcinomas developed in WD fed mice indicating a strong diet effect on cancer predisposition. 33% of CRCs and even 75% of adenomas and hyperplasias were found

in *Mlh1*$^{+/-}$ mice up to 18 mo of age. Surprisingly, Mlh1 protein was present and there was no MSI in their cancers. Genome-wide expression profiling of histologically normal mucosa however showed that 5/6 mice who developed CRC up to 18 mo had significantly decreased mucosal *Mlh1* RNA expression. Only in the carcinoma mouse E347 the *Mlh1* expression level was similar to the average level of 74 mice without cancer.

Low *Mlh1* expression, although a prominent signal, seemed not to be an absolute requirement or sufficient alone to cause colon cancer since several mice without CRC had low *Mlh1* expression as well. In order to identify other genes and pathways involved in CRC development we compared the genome wide expression profiles in the six CRC mice with the profiles of 74 mice without CRC. Remarkably, the expression profiles of CRC mice formed a clearly distinct cluster (Figures 3A and B), indicating a field-defect in normal colon mucosa(25, 26). By network analysis of top 100 CRC mice separating genes, low *Mlh1* expression in normal mucosa from CRC mice was associated with significant down regulation of several cancer related genes and pathways (Table 1) and especially of chromosomal segregation genes, *Bub1*, *Mis18a*, *Tpx2*, *Rad9a*, *Pms2*, *Cenpe*, *Ncapd3*, *Odf2* and *Dclre1b*. Only two (E325 and E332) of the 74 mice without cancer shared the expression profile of the CRC mice related to chromosomal segregation (Fig. 4). Although no colonic tumors were found in those mice (possibly not detected during operation), carcinogenesis has probably been currently happening in their mucosa. For example E325 had bloody feces and anemia suggesting pathological problems in mucosa. Differing from the other CRC mice, E347, which did not show decrease in *Mlh1* expression, showed decrease only in the expressions of *Bub1*, *Mis18a*, *Tpx2*, *Rad9a*, *Pms2*, and *Dclre1b*, suggesting their remarkable importance in serrated carcinogenesis. Furthermore, in the mouse E249 the *Mlh1* and *Dclre1b* genes showed approximately 50% lower expression than was detected in the non-carcinoma mice on average. Here, the milder decrease reflects young age of onset.

Table 1. The 300 most regulated genes between carcinoma and non-carcinoma mice.

Symbol MGI ID Name Shrink-T

Qrsl1 MGI:1923813 glutaminyl-tRNA synthase (glutamine-hydrolyzing)-like 1 -7.91
Eftud1 MGI:2141969 elongation factor Tu GTP binding domain containing 1 -7.39
Rad9a MGI:1328356 RAD9 homolog A -6.64
Slc10a5 MGI:2685251 solute carrier family 10 (sodium/bile acid cotransporter family), member 5 -6.49

Gpr39 MGI:1918361 G protein-coupled receptor 39 -6.34
Vamp5 MGI:1858622 vesicle-associated membrane protein 5 -6.27
Tmem82 MGI:2384869 transmembrane protein 82 -6.23
Tmem180 MGI:1922396 transmembrane protein 180 -6.20
Preb MGI:1355326 prolactin regulatory element binding -6.04
Mtmr9 MGI:2442842 myotubularin related protein 9 -6.03
Ranbp3 MGI:1919060 RAN binding protein 3 -6.03
Pard6a MGI:1927223 par-6 family cell polarity regulator alpha -6.01
D6Wsu163e MGI:107893 DNA segment, Chr 6, Wayne State University 163, expressed -5.94
Sphk2 MGI:1861380 sphingosine kinase 2 -5.80
Ncapd3 MGI:2142989 non-SMC condensin II complex, subunit D3 -5.73
Sec16b MGI:2148802 SEC16 homolog B (S. cerevisiae) -5.64
Triqk MGI:3650048 triple QxxK/R motif containing -5.63
Rbfa MGI:1915981 ribosome binding factor A -5.61
Pgd MGI:97553 phosphogluconate dehydrogenase -5.60
Ms4a10 MGI:1917076 membrane-spanning 4-domains, subfamily A, member 10 -5.58
Tle4 MGI:104633 transducin-like enhancer of split 4, homolog of Drosophila E(spl) 5.57
Mfsd3 MGI:1916822 major facilitator superfamily domain containing 3 -5.53
Tomm40 MGI:1858259 translocase of outer mitochondrial membrane 40 homolog (yeast) -5.52
Zfp60 MGI:99207 zinc finger protein 60 -5.51
Sf3a2 MGI:104912 splicing factor 3a, subunit 2 -5.47
Tcaf1 MGI:1914665 TRPM8 channel-associated factor 1 -5.43
Zcrb1 MGI:1914447 zinc finger CCHC-type and RNA binding motif 1 -5.36
Car9 MGI:2447188 carbonic anhydrase 9 -5.28
Ankrd27 MGI:2444103 ankyrin repeat domain 27 (VPS9 domain) -5.27
Ktn1 MGI:109153 kinectin 1 -5.26
Tgds MGI:1923605 TDP-glucose 4,6-dehydratase -5.17
Vti1a MGI:1855699 vesicle transport through interaction with t-SNAREs 1A -5.15
Kbtbd4 MGI:1914386 kelch repeat and BTB (POZ) domain containing 4 -5.09
Tbrg4 MGI:1100868 transforming growth factor beta regulated gene 4 -5.07
Zkscan5 MGI:107533 zinc finger with KRAB and SCAN domains 5 -5.07
Atl3 MGI:1924270 atlastin GTPase 3 -5.05
Mis18a MGI:1913828 MIS18 kinetochore protein homolog A (S. pombe) -5.03
Rnf185 MGI:1922078 ring finger protein 185 -5.02
Odf2 MGI:1098824 outer dense fiber of sperm tails 2 -5.01
Rilpl2 MGI:1933112 Rab interacting lysosomal protein-like 2 -5.00
Lman2 MGI:1914140 lectin, mannose-binding 2 -4.99
Gle1 MGI:1921662 GLE1 RNA export mediator (yeast) -4.97
Pnck MGI:1347357 pregnancy upregulated non-ubiquitously expressed CaM kinase -4.94
Gnb5 MGI:101848 guanine nucleotide binding protein (G protein), beta 5 -4.92
Dlg3 MGI:1888986 discs, large homolog 3 (Drosophila) -4.91
C4bp MGI:88229 complement component 4 binding protein 4.91
Rnf123 MGI:2148796 ring finger protein 123 -4.89

Serp2 MGI:1919911 stress-associated endoplasmic reticulum protein family member 2 4.88
Foxa3 MGI:1347477 forkhead box A3 -4.84
Rpusd1 MGI:1919186 RNA pseudouridylate synthase domain containing 1 -4.81
Gm5803 MGI:3645633 predicted gene 5803 -4.80
Fbxl15 MGI:1915681 F-box and leucine-rich repeat protein 15 -4.78
Sfxn3 MGI:2137679 sideroflexin 3 -4.75
Enthd2 MGI:1926027 ENTH domain containing 2 -4.74
Smagp MGI:2448476 small cell adhesion glycoprotein 4.74
Mt2 MGI:97172 metallothionein 2 4.72
Slc44a2 MGI:1915932 solute carrier family 44, member 2 -4.72
Pla2g4a MGI:1195256 phospholipase A2, group IVA (cytosolic, calcium-dependent) -4.71
Bola3 MGI:1925903 bolA-like 3 (E. coli) -4.68
Ifrd1 MGI:1316717 interferon-related developmental regulator 1 -4.68
Dclre1b MGI:2156057 DNA cross-link repair 1B, PSO2 homolog (S. cerevisiae) -4.66
C330007P06RMikGI:1924894 RIKEN cDNA C330007P06 gene -4.65
Mospd2 MGI:1924013 motile sperm domain containing 2 -4.64
Hspa13 MGI:1309463 heat shock protein 70 family, member 13 -4.63
Atp6v0e2 MGI:1923502 ATPase, H+ transporting, lysosomal V0 subunit E2 4.63
Cstf2t MGI:1932622 cleavage stimulation factor, 3' pre-RNA subunit 2, tau -4.61
Ttll1 MGI:2443047 tubulin tyrosine ligase-like 1 -4.60
Casp12 MGI:1312922 caspase 12 -4.59
Pms2 MGI:104288 postmeiotic segregation increased 2 (S. cerevisiae) -4.57
Bcl2a1a MGI:102687 B cell leukemia/lymphoma 2 related protein A1a -4.56
Trim25 MGI:102749 tripartite motif-containing 25 -4.56
Hist1h2bj MGI:2448388 histone cluster 1, H2bj -4.55
Rabgap1 MGI:2385139 RAB GTPase activating protein 1 -4.52
Fkbp8 MGI:1341070 FK506 binding protein 8 -4.49
Creb3l3 MGI:2384786 cAMP responsive element binding protein 3-like 3 4.48
Bub1 MGI:1100510 budding uninhibited by benzimidazoles 1 homolog (S. cerevisiae) -4.48
Tpx2 MGI:1919369 TPX2, microtubule-associated protein homolog (Xenopus laevis) -4.47
Ints10 MGI:1918135 integrator complex subunit 10 -4.43
Dph6 MGI:1913882 diphthamine biosynthesis 6 -4.43
Gdpd1 MGI:1913819 glycerophosphodiester phosphodiesterase domain containing 1 4.42
2010003K11RMikGI:1917111 RIKEN cDNA 2010003K11 gene -4.40
Cxcl16 MGI:1932682 chemokine (C-X-C motif) ligand 16 -4.38
Mrpl10 MGI:1333801 mitochondrial ribosomal protein L10 -4.38
Slc25a12 MGI:1926080 solute carrier family 25 (mitochondrial carrier, Aralar), member 12 4.38
Urah MGI:1916142 urate (5-hydroxyiso-) hydrolase 4.36
Nfyc MGI:107901 nuclear transcription factor-Y gamma -4.32
1700037C18RMikGI:1920511 RIKEN cDNA 1700037C18 gene -4.29
Ddx23 MGI:1921601 DEAD (Asp-Glu-Ala-Asp) box polypeptide 23 -4.27
Cenpe MGI:1098230 centromere protein E -4.23

P4ha2 MGI:894286 procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase), alpha II polypeptide 4.22
Apol9b MGI:1919148 apolipoprotein L 9b -4.21
Ccar2 MGI:2444228 cell cycle activator and apoptosis regulator 2 -4.20
Plscr3 MGI:1917560 phospholipid scramblase 3 -4.20
Slco2b1 MGI:1351872 solute carrier organic anion transporter family, member 2b1 -4.16
Kdm8 MGI:1924285 lysine (K)-specific demethylase 8 -4.15
Tmem164 MGI:2148020 transmembrane protein 164 -4.12
Efcab11 MGI:1926017 EF-hand calcium binding domain 11 -4.12
Wrap53 MGI:2384933 WD repeat containing, antisense to Trp53 -4.10
Cenpn MGI:1919405 centromere protein N -4.09
Atf6b MGI:105121 activating transcription factor 6 beta -4.09
Hspbp1 MGI:1913495 HSPA (heat shock 70kDa) binding protein, cytoplasmic co-chaperone 1 -4.08
Slfn2 MGI:1313258 schlafen 2 -4.08
Cenpa MGI:88375 centromere protein A -4.07
4931414P19RMikGI:1921609 RIKEN cDNA 4931414P19 gene -4.06
Arnt MGI:88071 aryl hydrocarbon receptor nuclear translocator -4.06
Dolk MGI:2677836 dolichol kinase -4.06
Tlr4 MGI:96824 toll-like receptor 4 -4.04
Dlst MGI:1926170 dihydrolipoamide S-succinyltransferase (E2 component of 2-oxo-glutarate complex) -4.04
Vim MGI:98932 vimentin 4.04
Uprt MGI:2685620 uracil phosphoribosyltransferase (FUR1) homolog (S. cerevisiae) -4.03
2210016F16RMikGI:1917403 RIKEN cDNA 2210016F16 gene -4.02
Fahd2a MGI:1915376 fumarylacetoacetate hydrolase domain containing 2A -4.02
Ccdc23 MGI:1916466 small vasohibin binding protein 4.01
Pin4 MGI:1916963 protein (peptidyl-prolyl cis/trans isomerase) NIMA-interacting, 4 (parvulin) 3.99
Lxn MGI:107633 latexin 3.99
Arhgef3 MGI:1918954 Rho guanine nucleotide exchange factor (GEF) 3 -3.99
Zfp748 MGI:1916455 zinc finger protein 748 -3.98
Fubp1 MGI:1196294 far upstream element (FUSE) binding protein 1 -3.98
Ptcd3 MGI:1917206 pentatricopeptide repeat domain 3 -3.97
Pstpip2 MGI:1335088 proline-serine-threonine phosphatase-interacting protein 2 -3.97
Osgin1 MGI:1919089 oxidative stress induced growth inhibitor 1 -3.96
Zfp930 MGI:2675306 zinc finger protein 930 -3.96
Psmg4 MGI:1916916 proteasome (prosome, macropain) assembly chaperone 4 3.95
Cped1 MGI:2444814 cadherin-like and PC-esterase domain containing 1 -3.94
Rnase1 MGI:97919 ribonuclease, RNase A family, 1 (pancreatic) -3.93
Cpsf6 MGI:1913948 cleavage and polyadenylation specific factor 6 -3.93
Kif4 MGI:108389 kinesin family member 4 -3.92
Traip MGI:1096377 TRAF-interacting protein -3.92
Exoc2 MGI:1913732 exocyst complex component 2 -3.92
Hist1h4b MGI:2448420 histone cluster 1, H4b -3.92

Fancc MGI:95480 Fanconi anemia, complementation group C -3.91
2310007B03RMikGI:1919124 RIKEN cDNA 2310007B03 gene -3.90
Bud13 MGI:2443443 BUD13 homolog (yeast) -3.90
Wdr12 MGI:1927241 WD repeat domain 12 -3.90
Fadd MGI:109324 Fas (TNFRSF6)-associated via death domain -3.90
Stk4 MGI:1929004 serine/threonine kinase 4 -3.89
Lurap1l MGI:106510 leucine rich adaptor protein 1-like 3.88
Cln8 MGI:1349447 ceroid-lipofuscinosis, neuronal 8 3.87
Btrc MGI:1338871 beta-transducin repeat containing protein -3.87
Vps13c MGI:2444207 vacuolar protein sorting 13C (yeast) -3.87
Tra2b MGI:106016 transformer 2 beta homolog (Drosophila) 3.86
AA986860 MGI:2138143 expressed sequence AA986860 3.86
Pak1 MGI:1339975 p21 protein (Cdc42/Rac)-activated kinase 1 -3.85
Letmd1 MGI:1915864 LETM1 domain containing 1 -3.85
Pusl1 MGI:3047787 pseudouridylate synthase-like 1 -3.85
Fbxo38 MGI:2444639 F-box protein 38 -3.85
Rad54l MGI:894697 RAD54 like (S. cerevisiae) -3.84
Coq10b MGI:1915126 coenzyme Q10 homolog B (S. cerevisiae) -3.84
Nme3 MGI:1930182 NME/NM23 nucleoside diphosphate kinase 3 3.83
Ppp1r2 MGI:1914099 protein phosphatase 1, regulatory (inhibitor) subunit 2 -3.83
Ikbkg MGI:1338074 inhibitor of kappaB kinase gamma -3.83
Trmt1l MGI:1916185 tRNA methyltransferase 1 like -3.83
Ddrgk1 MGI:1924256 DDRGK domain containing 1 3.83
Zc3h12a MGI:2385891 zinc finger CCCH type containing 12A -3.82
Zfp65 MGI:107769 zinc finger protein 65 -3.82
Zfp808 MGI:3704127 zinc finger protein 80 -3.82
Nat8 MGI:1915646 N-acetyltransferase 8 (GCN5-related -3.81
Slc9a3r2 MGI:1890662 solute carrier family 9 (sodium/hydrogen exchanger), member 3 regulator 2 -3.81
Pls3 MGI:104807 plastin 3 (T-isoform) -3.80
Hist1h2an MGI:2448300 histone cluster 1, H2an -3.79
Ank3 MGI:88026 ankyrin 3, epithelial -3.75
Cd36 MGI:107899 CD36 antigen -3.74
Golph3l MGI:1917129 golgi phosphoprotein 3-like -3.74
Slc39a7 MGI:95909 solute carrier family 39 (zinc transporter), member 7 -3.74
Stk19 MGI:1860085 serine/threonine kinase 19 -3.74
Gcdh MGI:104541 glutaryl-Coenzyme A dehydrogenase -3.74
Tmx3 MGI:2442418 thioredoxin-related transmembrane protein 3 -3.73
Apol10a MGI:3036238 apolipoprotein L 10A -3.71
Mmgt1 MGI:2384305 membrane magnesium transporter 1 -3.71
Ptbp2 MGI:1860489 polypyrimidine tract binding protein 2 -3.71
Tuba1b MGI:107804 tubulin, alpha 1B -3.70
Egfl7 MGI:2449923 EGF-like domain 7 -3.70
Zfp956 MGI:2141515 zinc finger protein 956 -3.69
Ccnf MGI:102551 cyclin F -3.69
9030617O03MRikGI:2444813 RIKEN cDNA 9030617O03 gene -3.68
Slc39a1 MGI:1353474 solute carrier family 39 (zinc transporter), member 1 -3.68
Myd88 MGI:108005 myeloid differentiation primary response gene 88 -3.67

Ace2 MGI:1917258 angiotensin I converting enzyme (peptidyl-dipeptidase A) 2 -3.67
Kif22 MGI:109233 kinesin family member 22 -3.66
Trmt2a MGI:96270 TRM2 tRNA methyltransferase 2A -3.65
Gnb1 MGI:95781 guanine nucleotide binding protein (G protein), beta 1 -3.65
Dnajb1 MGI:1931874 DnaJ (Hsp40) homolog, subfamily B, member 1 -3.65
Creld1 MGI:2152539 cysteine-rich with EGF-like domains 1 -3.64
Gba2 MGI:2654325 glucosidase beta 2 -3.64
Pot1a MGI:2141503 protection of telomeres 1A -3.64
Sf3a1 MGI:1914715 splicing factor 3a, subunit 1 -3.63
Tmx2 MGI:1914208 thioredoxin-related transmembrane protein 2 -3.63
Gskip MGI:1914037 GSK3B interacting protein 3.61
Tnip3 MGI:3041165 TNFAIP3 interacting protein 3 -3.61
Cnot4 MGI:1859026 CCR4-NOT transcription complex, subunit 4 3.60
Uap1l1 MGI:2443318 UDP-N-acteylglucosamine pyrophosphorylase 1-like 1 -3.60
Ezh1 MGI:1097695 enhancer of zeste 1 polycomb repressive complex 2 subunit -3.59
Mettl8 MGI:2385142 methyltransferase like 8 -3.59
Mbtps1 MGI:1927235 membrane-bound transcription factor peptidase, site 1 -3.58
Cep95 MGI:2443502 centrosomal protein 95 -3.57
Smg9 MGI:1919247 smg-9 homolog, nonsense mediated mRNA decay factor (C. elegans) -3.57
Aurkaip1 MGI:1913327 aurora kinase A interacting protein 1 3.56
Ttf2 MGI:1921294 transcription termination factor, RNA polymerase II -3.56
Adck4 MGI:1924139 aarF domain containing kinase 4 -3.55
Hdgfrp2 MGI:1194492 hepatoma-derived growth factor, related protein 2 -3.55
Csde1 MGI:92356 cold shock domain containing E1, RNA binding -3.53
Clns1a MGI:109638 chloride channel, nucleotide-sensitive, 1A -3.53
Dpf3 MGI:1917377 D4, zinc and double PHD fingers, family 3 -3.52
H1f0 MGI:95893 H1 histone family, member 0 -3.52
Haus5 MGI:1919159 HAUS augmin-like complex, subunit 5 -3.52
Pitx2 MGI:109340 paired-like homeodomain transcription factor 2 -3.51
Wfdc1 MGI:1915116 WAP four-disulfide core domain 1 -3.51
Ubac1 MGI:1920995 ubiquitin associated domain containing 1 -3.51
Abcf1 MGI:1351658 ATP-binding cassette, sub-family F (GCN20), member 1 3.51
Sec24c MGI:1919746 Sec24 related gene family, member C (S. cerevisiae) -3.50
Galnt7 MGI:1349449 UDP-N-acetyl-alpha-D-galactosamine: polypeptide Nacetyl-galactosaminyltransferase 7 -3.50
Shc1 MGI:98296 src homology 2 domain-containing transforming protein C1 -3.49
P4hb MGI:97464 prolyl 4-hydroxylase, beta polypeptide -3.49
Skiv2l MGI:1099835 superkiller viralicidic activity 2-like (S. cerevisiae) -3.49
Mlh1 MGI:101938 mutL homolog 1 (E. coli) -3.49
Gmpr2 MGI:1917903 guanosine monophosphate reductase 2 -3.48
Zfp329 MGI:1921283 zinc finger protein 329 -3.48
Abi3bp MGI:2444583 ABI gene family, member 3 (NESH) binding protein -3.47
Pax4 MGI:97488 paired box 4 -3.47
Gnpat MGI:1343460 glyceronephosphate O-acyltransferase -3.47
Spdl1 MGI:1917635 spindle apparatus coiled-coil protein 1 -3.46
Eps8l3 MGI:2139743 EPS8-like 3 -3.45

Clcn3 MGI:103555 chloride channel, voltage-sensitive 3 -3.45
Cdk5rap1 MGI:1914221 CDK5 regulatory subunit associated protein 1 3.45
Itgae MGI:1298377 integrin alpha E, epithelial-associated -3.44
Rnf121 MGI:1922462 ring finger protein 121 -3.44
Twf2 MGI:1346078 twinfilin, actin-binding protein, homolog 2 (Drosophila) -3.43
Polr3a MGI:2681836 polymerase (RNA) III (DNA directed) polypeptide A -3.43
Tmem176a MGI:1913308 transmembrane protein 176A -3.42
Ccnb2 MGI:88311 cyclin B2 -3.42
Mt1 MGI:97171 metallothionein 1 3.42
Pole MGI:1196391 polymerase (DNA directed), epsilon -3.40
Bckdk MGI:1276121 branched chain ketoacid dehydrogenase kinase -3.40
Clps MGI:88421 colipase, pancreatic 3.40
Tsr1 MGI:1915061 || MGI:2T1S4R415 6260 S| r|RNA accumulation -3.40
Gpalpp1 MGI:1914717 GPALPP motifs containing 1 -3.39
Stoml2 MGI:1913842 stomatin (Epb7.2)-like 2 3.39
Irgm1 MGI:107567 immunity-related GTPase family M member 1 -3.39
Acox3 MGI:1933156 acyl-Coenzyme A oxidase 3, pristanoyl -3.39
Rph3al MGI:1923492 rabphilin 3A-like (without C2 domains) -3.39
Fam111a MGI:1915508 family with sequence similarity 111, member A -3.38
Cit MGI:105313 citron -3.37
Slc5a8 MGI:2384916 solute carrier family 5 (iodide transporter), member 8 -3.37
Cacnb3 MGI:103307 calcium channel, voltage-dependent, beta 3 subunit -3.36
Cox4i2 MGI:2135755 cytochrome c oxidase subunit IV isoform 2 3.36
Dctn2 MGI:107733 dynactin 2 3.36
Mfap2 MGI:99559 microfibrillar-associated protein 2 -3.35
Cc2d2a MGI:1924487 coiled-coil and C2 domain containing 2A -3.35
Igf2bp2 MGI:1890358 insulin-like growth factor 2 mRNA binding protein 2 -3.35
Ophn1 MGI:2151070 oligophrenin 1 -3.35
Rcn3 MGI:1277122 reticulocalbin 3, EF-hand calcium binding domain -3.35
Ppan MGI:2178445 peter pan homolog (Drosophila) -3.34
Rcl1 MGI:1913275 RNA terminal phosphate cyclase-like 1 -3.34
Tssc1 MGI:1289332 tumor suppressing subtransferable candidate 1 -3.33
Nnmt MGI:1099443 nicotinamide N-methyltransferase -3.33
Slc25a40 MGI:2442486 solute carrier family 25, member 40 -3.32
Srsf10 MGI:1333805 serine/arginine-rich splicing factor 10 -3.32
Jup MGI:96650 junction plakoglobin -3.32
Tipin MGI:1921571 timeless interacting protein -3.32
Phgr1 MGI:1858382 proline/histidine/glycine-rich 1 3.32
Pcp4 MGI:97509 Purkinje cell protein 4 3.32
Tbl3 MGI:2384863 transducin (beta)-like 3 -3.31
Fancg MGI:1926471 Fanconi anemia, complementation group G -3.30
Laptm5 MGI:108046 lysosomal-associated protein transmembrane 5 -3.30
Kpna2 MGI:103561 karyopherin (importin) alpha 2 -3.30
Phlda2 MGI:1202307 pleckstrin homology-like domain, family A, member 2 -3.29
Wdr46 MGI:1931871 WD repeat domain 46 -3.29
Rsrp1 MGI:106498 arginine/serine rich protein 1 3.29
Irf2 MGI:96591 interferon regulatory factor 2 3.29
Zfp426 MGI:1920248 zinc finger protein 426 -3.29
Cdh1 MGI:88354 cadherin 1 -3.28

Leng1 MGI:1917007 leukocyte receptor cluster (LRC) member 1 -3.27
Htr2b MGI:109323 5-hydroxytryptamine (serotonin) receptor 2B -3.27
Rhbdd1 MGI:1924117 rhomboid domain containing 1 -3.26
Tbx3 MGI:98495 T-box 3 -3.26
Arhgap6 MGI:1196332 Rho GTPase activating protein 6 -3.26
Chd4 MGI:1344380 chromodomain helicase DNA binding protein 4 -3.25
Gm20594 MGI:5295700 predicted gene, 20594 3.24
Fcrls MGI:1933397 Fc receptor-like S, scavenger receptor -3.23
Slc35b3 MGI:1913978 solute carrier family 35, member B3 -3.23
Pigh MGI:99463 phosphatidylinositol glycan anchor biosynthesis, class H -3.23
Tor1aip1 MGI:3582693 torsin A interacting protein 1 -3.22
Grp MGI:95833 gastrin releasing peptide -3.22
Mia2 MGI:2159614 melanoma inhibitory activity 2 -3.22
Usb1 MGI:2142454 U6 snRNA biogenesis 1 -3.21
Cyb5r3 MGI:94893 cytochrome b5 reductase 3 -3.21
Arhgef38 MGI:1924919 Rho guanine nucleotide exchange factor (GEF) 38 -3.21
Dhps MGI:2683592 deoxyhypusine synthase -3.21
Mt3 MGI:97173 metallothionein 3 3.21
Atxn10 MGI:1859293 ataxin 10 3.21
Dimt1 MGI:1913504 DIM1 dimethyladenosine transferase 1-like (S. cerevisiae) -3.21
Fcer1g MGI:95496 Fc receptor, IgE, high affinity I, gamma polypeptide 3.19
Fam49a MGI:1261783 family with sequence similarity 49, member A -3.18
Hells MGI:106209 helicase, lymphoid specific -3.18
Snrpb2 MGI:104805 U2 small nuclear ribonucleoprotein B 3.18
Kdm5c MGI:99781 lysine (K)-specific demethylase 5C -3.18
Mzb1 MGI:1917066 marginal zone B and B1 cell-specific protein 1 3.17
Oma1 MGI:1914263 OMA1 homolog, zinc metallopeptidase (S. cerevisiae) -3.17
Slc9a8 MGI:1924281 solute carrier family 9 (sodium/hydrogen exchanger), member 8 -3.17
Abcc3 MGI:1923658 ATP-binding cassette, sub-family C (CFTR/MRP), member 3 -3.16

Furthermore, Figure 7 shows expressions of ten genes associated with chromosomal segregation in mouse individuals (carcinoma mice are presented as E314, E333, E329, E338, E249 and E347), and Figure 8 shows results of differential gene expression analysis: carcinoma mice versus non-carcinoma mice (Shrink T scores (expression differences) and P-values are highlighted).

In the present study, the mRNA expression was significantly decreased in five SAC associated genes, *Mlh1*, *Bub1*, *Rad9a*, *Dclre1b* and *Cenpe.* Of those, *Bub1* is a major player and activator in SAC and its haploinsufficiency (heterozygosity) is known to be responsible for chromosome segregation defects and aneuploidy(29). During mitosis Bub1 is required for the recruitment of other checkpoint and motor proteins, such as Cenpe, to a kinetochore(30). There is evidence suggesting that

inaccurate chromosome segregation with causal implication of Bub1 deficiency drives tumorigenesis through tumor-suppressor gene LOH(31), perfectly in line with our findings that the majority of the carcinoma mice distinguishing genes were tumor suppressor genes, which were down regulated (Table 1). *Cenpe*, a kinesin-like motor protein which is an efficient stabilizer of microtubule capture at kinetochores and hence essential for metaphase chromosome alignment(32), was strongly down regulated in the mice with carcinoma. While it plays an important role in the movement of chromosomes toward the metaphase plate during mitosis, it is also necessary for the mitotic checkpoint signal at the kinetochore to prevent chromosome loss(33). *Dclre1b* has a central role in telomere maintenance and protection during S-phase through its 5-3 exonuclease activity. Moreover, in case of spindle stress, Dclre1b like Bub1 is involved in prophase checkpoint(34, 35). *RAD9A*, a component of the 9-1-1 cell cycle checkpoint response complex, plays a major role in DNA repair and participates in multiple cell cycle checkpoints and apoptosis and its aberrant expression has been linked to tumorigenesis of multiple tissues(36). Interestingly, Rad9 also physically interacts with the MMR protein MLH1(37). The MMR mechanism is so essential for normal cell function that it may explain why even a small amount of MLH1 appears to be sufficient for MMR function, whereas its checkpoint activation role seems to require a full complement of the protein(38). It has been argued that the *MLH1* heterozygosity/haploinsufficiency may drive the development of cancer by accumulation of insertion/deletion mutations in other gatekeeper genes prior to MSI.(39) Indeed, cells with diminished amount of MLH1 protein may still be MMR proficient, although they show defects in DNA damage signaling(37). Consequently the damaged cells may not activate cell cycle checkpoints and enter apoptosis. Our observation that low mRNA expression of *Mlh1* in carcinoma mice was associated with down regulation of several other genes related to chromosome segregation and checkpoint control supports the proposition that already decreased amount of *Mlh1*, when MMR is still functional, has an important role in tumorigenesis.

Low expression of *Tpx2*, *Mis18a*, *Ncapd3* and *Odf2* reflects problems in formation of the nuclear spindle and chromosome segregation. *Tpx2*, *Ncpd3* and *Odf2*, a general scaffolding protein(40), are all involved in microtubules related processes in spindle formation.Tpx2 plays a role in microtubule organization and is involved in centrosome maturation(41). In fact, TPX2-depleted cells fail to form proper mitotic spindles(42). Recent findings suggest that TPX2 also plays an important role in promoting colon tumorigenesis(43). In the present study results support a driver role for *Tpx2*, since it was strongly down regulated in colon mucosa in all carcinoma

mice. Ncapd3 functions in the condensin II complex and is needed to establish the chromosomal architecture necessary for proper spindle assembly and chromosome segregation. Chromosome condensation and resolution are compromised when condensin is depleted(44). The MIS18 complex accumulates at the centromere during anaphase to early G1 phase, slightly ahead of the histone H3 variant CENPA, and is an absolute requirement for the localization of CENPA at centromeres. Importantly, Mis18a knockout causes severe chromosomal missegregation, lack of CENPA, and ultimately cell death(45). Here, along with *Mis18a*, *Cenpa* was significantly down regulated in the normal colon mucosa of the CRC mice (Table 1) supporting the finding of improper chromosome segregation.

One example of the method of the present invention

Normal mucosa samples were collected from mice as described above in the "materials and methods" section. Total RNA was prepared and converted to cDNA as described above in the chapter "Transcriptome analysis of normal mucosa".

After reverse transcription genes of interest i.e. at least three, four, five, six, seven, eight, nine or ten genes associated with chromosomal segregation selected from the group consisting of *Bub1*, *Mis18a*, *Tpx2*, *Rad9a*, *Pms2*, *Mlh1*, *Cenpe*, *Ncapd3*, *Odf2* and *Dclre1b* are amplified with gene specific primers (e.g. commercial primers). Optionally one or more control samples are utilized in the method.

Modified t-tests are utilized as described above in "Statistical Analysis" chapter.

When the expression levels of at least three, four, five, six, seven, eight, nine or ten genes associated with chromosomal segregation selected from the group consisting of *Bub1*, *Mis18a*, *Tpx2*, *Rad9a*, *Pms2*, *Mlh1*, *Cenpe*, *Ncapd3*, *Odf2* and *Dclre1b* are decreased compared to normal expression levels of the same genes, said decrease of said at least three, four, five, six, seven, eight, nine or ten genes is indicative of an increased risk of developing cancer.

Optionally, subjects found with said decreased expression levels of at least three, four, five, six, seven, eight, nine or ten genes associated with chromosomal segregation selected from the group consisting of *Bub1*, *Mis18a*, *Tpx2*, *Rad9a*, *Pms2*, *Mlh1*, *Cenpe*, *Ncapd3*, *Odf2* and *Dclre1b* are further examined with colonoscopy.

References

1. Brenner H, Kloor M, Pox CP. Colorectal cancer. Lancet. 2014;383(9927):1490-502.
2. Bingham S, Riboli E. Diet and cancer--the European Prospective Investigation into Cancer and Nutrition. Nat Rev Cancer. 2004;4(3):206-15.
3. Fearon ER. Molecular genetics of colorectal cancer. Annu Rev Pathol. 2011;6:479-507.
4. Abbas T, Keaton MA, Dutta A. Genomic instability in cancer. Cold Spring Harb Perspect Biol. 2013;5(3):a012914.
5. Al-Sohaily S, Biankin A, Leong R, Kohonen-Corish M, Warusavitarne J. Molecular pathways in colorectal cancer. J Gastroenterol Hepatol. 2012;27(9):1423-31.
6. Pino MS, Chung DC. The chromosomal instability pathway in colon cancer. Gastroenterology. 2010;138(6):2059-72.
7. Weinberg RA. The biology of cancer. New York: Garland Science; 2007. xix, 796 p., 48 p. p.
8. Colussi D, Brandi G, Bazzoli F, Ricciardiello L. Molecular pathways involved in colorectal cancer: implications for disease behavior and prevention. Int J Mol Sci. 2013;14(8):16365-85.
9. Peltomaki P. Update on Lynch syndrome genomics. Fam Cancer. 2016;15(3):385-93.
10. Boland CR, Goel A. Microsatellite instability in colorectal cancer. Gastroenterology. 2010;138(6):2073-87 e3.
11. Parsons MT, Buchanan DD, Thompson B, Young JP, Spurdle AB. Correlation of tumour BRAF mutations and MLH1 methylation with germline mismatch repair (MMR) gene mutation status: a literature review assessing utility of tumour features for MMR variant classification. J Med Genet. 2012;49(3):151-7.
12. Ahuja N, Issa JP. Aging, methylation and cancer. Histol Histopathol. 2000;15(3):835-42.
13. Toyota M, Issa JP. CpG island methylator phenotypes in aging and cancer. Semin Cancer Biol. 1999;9(5):349-57.
14. Thompson BA, Spurdle AB, Plazzer JP, Greenblatt MS, Akagi K, Al-Mulla F, et al. Application of a 5-tiered scheme for standardized classification of 2,360 unique mismatch repair gene variants in the InSiGHT locus-specific database. Nat Genet. 2014;46(2):107-15.
15. Valo S, Kaur S, Ristimaki A, Renkonen-Sinisalo L, Jarvinen H, Mecklin JP, et al. DNA hypermethylation appears early and shows increased frequency with dysplasia in Lynch syndrome-associated colorectal adenomas and carcinomas. Clin Epigenetics. 2015;7:71.
16. Goel A, Arnold CN, Niedzwiecki D, Chang DK, Ricciardiello L, Carethers JM, et al. Characterization of sporadic colon cancer by patterns of genomic instability. Cancer Res. 2003;63(7):1608-14.
17. Newmark HL, Yang K, Kurihara N, Fan K, Augenlicht LH, Lipkin M. Western-style diet-induced colonic tumors and their modulation by calcium and vitamin D in C57Bl/6 mice: a preclinical model for human sporadic colon cancer. Carcinogenesis. 2009;30(1):88-92.

18. Newmark HL, Yang K, Lipkin M, Kopelovich L, Liu Y, Fan K, et al. A Western-style diet induces benign and malignant neoplasms in the colon of normal C57Bl/6 mice. Carcinogenesis. 2001;22(11):1871-5.
19. Yang K, Kurihara N, Fan K, Newmark H, Rigas B, Bancroft L, et al. Dietary induction of colonic tumors in a mouse model of sporadic colon cancer. Cancer Res. 2008;68(19):7803-10.
20. Yang K, Yang W, Mariadason J, Velcich A, Lipkin M, Augenlicht L. Dietary components modify gene expression: implications for carcinogenesis. J Nutr. 2005;135(11):2710-4.
21. Islam S, Kjallquist U, Moliner A, Zajac P, Fan JB, Lonnerberg P, et al. Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. Genome Res. 2011;21(7):1160-7.
22. Krjutskov K, Katayama S, Saare M, Vera-Rodriguez M, Lubenets D, Samuel K, et al. Single-cell transcriptome analysis of endometrial tissue. Hum Reprod. 2016;31(4):844-53.
23. Pussila M, Sarantaus L, Dermadi Bebek D, Valo S, Reyhani N, Ollila S, et al. Cancer-predicting gene expression changes in colonic mucosa of Western diet fed Mlh1+/- mice. PloS one. 2013;8(10):e76865. Epub 2013/11/10.
24. Ollikainen M, Hannelius U, Lindgren CM, Abdel-Rahman WM, Kere J, Peltomaki P. Mechanisms of inactivation of MLH1 in hereditary nonpolyposis colorectal carcinoma: a novel approach. Oncogene. 2007;26(31):4541-9.
25. Park SK, Song CS, Yang HJ, Jung YS, Choi KY, Koo DH, et al. Field Cancerization in Sporadic Colon Cancer. Gut Liver. 2016;10(5):773-80.
26. Hawthorn L, Lan L, Mojica W. Evidence for field effect cancerization in colorectal cancer. Genomics. 2014;103(2-3):211-21.
27. Lara-Gonzalez P, Westhorpe FG, Taylor SS. The spindle assembly checkpoint. Curr Biol. 2012;22(22):R966-80.
28. Perez de Castro I, de Carcer G, Malumbres M. A census of mitotic cancer genes: new insights into tumor cell biology and cancer therapy. Carcinogenesis. 2007;28(5):899-912.
29. de Voer RM, Geurts van Kessel A, Weren RD, Ligtenberg MJ, Smeets D, Fu L, et al. Germline mutations in the spindle assembly checkpoint genes BUB1 and BUB3 are risk factors for colorectal cancer. Gastroenterology. 2013;145(3):544-7.
30. Johnson VL, Scott MI, Holt SV, Hussein D, Taylor SS. Bub1 is required for kinetochore localization of BubR1, Cenp-E, Cenp-F and Mad2, and chromosome congression. J Cell Sci. 2004;117(Pt 8):1577-89.
31. Sotillo R, Schvartzman JM, Benezra R. Very CIN-ful: whole chromosome instability promotes tumor suppressor loss of heterozygosity. Cancer Cell. 2009;16(6):451-2.
32. Putkey FR, Cramer T, Morphew MK, Silk AD, Johnson RS, McIntosh JR, et al. Unstable kinetochore-microtubule capture and chromosomal instability following deletion of CENP-E. Dev Cell. 2002;3(3):351-65.
33. Weaver BA, Bonday ZQ, Putkey FR, Kops GJ, Silk AD, Cleveland DW. Centromere-associated protein-E is essential for the mammalian mitotic checkpoint to prevent aneuploidy due to single chromosome loss. J Cell Biol. 2003;162(4):551-63.
34. Liu L, Akhter S, Bae JB, Mukhopadhyay SS, Richie CT, Liu X, et al. SNM1B/Apollo interacts with astrin and is required for the prophase cell cycle checkpoint. Cell Cycle. 2009;8(4):628-38.

35. Yan Y, Akhter S, Zhang X, Legerski R. The multifunctional SNM1 gene family: not just nucleases. Future Oncol. 2010;6(6):1015-29.
36. Lieberman HB, Bernstock JD, Broustas CG, Hopkins KM, Leloup C, Zhu A. The role of RAD9 in tumorigenesis. J Mol Cell Biol. 2011;3(1):39-43.
37. He W, Zhao Y, Zhang C, An L, Hu Z, Liu Y, et al. Rad9 plays an important role in DNA mismatch repair through physical interaction with MLH1. Nucleic Acids Res. 2008;36(20):6406-17.
38. Cejka P, Stojic L, Mojas N, Russell AM, Heinimann K, Cannavo E, et al. Methylation-induced G(2)/M arrest requires a full complement of the mismatch repair protein hMLH1. The EMBO journal. 2003;22(9):2245-54.
39. Wang L, Tsutsumi S, Kawaguchi T, Nagasaki K, Tatsuno K, Yamamoto S, et al. Whole-exome sequencing of human pancreatic cancers and characterization of genomic instability caused by MLH1 haploinsufficiency and complete deficiency. Genome Res. 2012;22(2):208-19.
40. Ishikawa H, Kubo A, Tsukita S, Tsukita S. Odf2-deficient mother centrioles lack distal/subdistal appendages and the ability to generate primary cilia. Nat Cell Biol. 2005;7(5):517-24.
41. Karsenti E, Vernos I. The mitotic spindle: a self-made machine. Science. 2001;294(5542):543-7.
42. De Luca M, Lavia P, Guarguaglini G. A functional interplay between Aurora-A, Plk1 and TPX2 at spindle poles: Plk1 controls centrosomal localization of Aurora-A and TPX2 spindle association. Cell Cycle. 2006;5(3):296-303.
43. Wei P, Zhang N, Xu Y, Li X, Shi D, Wang Y, et al. TPX2 is a novel prognostic marker for the growth and metastasis of colon cancer. J Transl Med. 2013;11:313.
44. Wignall SM, Deehan R, Maresca TJ, Heald R. The condensin complex is required for proper spindle assembly and chromosome segregation in Xenopus egg extracts. J Cell Biol. 2003;161(6):1041-51.
45. Baumann K. Keeping centromeric identity. Nat Rev Mol Cell Biol. 2012;13(6):340.
46. Edelmann W, Cohen PE, Kane M, Lau K, Morrow B, Bennett S, et al. Meiotic pachytene arrest in MLH1-deficient mice. Cell. 1996;85(7):1125-34. Epub 1996/06/28.
47. Reeves PG, Nielsen FH, Fahey GC, Jr. AIN-93 purified diets for laboratory rodents: final report of the American Institute of Nutrition ad hoc writing committee on the reformulation of the AIN-76A rodent diet. J Nutr. 1993;123(11):1939-51.
48. Washington MK, Powell AE, Sullivan R, Sundberg JP, Wright N, Coffey RJ, et al. Pathology of rodent models of intestinal cancer: progress report and recommendations. Gastroenterology. 2013;144(4):705-17.
49. Islam S, Kjallquist U, Moliner A, Zajac P, Fan JB, Lonnerberg P, et al. Highly multiplexed and strand-specific single-cell RNA 5' end sequencing. Nat Protoc. 2012;7(5):813-28.
50. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 2013;14(4):R36.
51. Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol. 2010;11(10):R106.
52. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007;8(1):118-27.

53. Law CW, Chen Y, Shi W, Smyth GK. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol. 2014;15(2):R29.
54. Baldi P, Long AD. A Bayesian framework for the analysis of microarray expression data: regularized t -test and statistical inferences of gene changes. Bioinformatics. 2001;17(6):509-19.
55. Opgen-Rhein R, Strimmer K. Accurate ranking of differentially expressed genes by a distribution-free shrinkage approach. Stat Appl Genet Mol Biol. 2007;6:Article9.
56. Mishra P, Toronen P, Leino Y, Holm L. Gene set analysis: limitations in popular existing methods and proposed improvements. Bioinformatics. 2014;30(19):2747-56.
57. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47.
58. Edelmann W, Yang K, Kuraguchi M, Heyer J, Lia M, Kneitz B, et al. Tumorigenesis in Mlh1 and Mlh1/Apc1638N mutant mice. Cancer Res. 1999;59(6):1301-7.
59. Kauppi L, May CA, Jeffreys AJ. Analysis of meiotic recombination products from human sperm. Methods Mol Biol. 2009;557:323-55.
60. Luo, Yu and Grady. Field cancerization in the colon: a role for aberrant DNA methylation? Gastroenterol Rep (Oxf). 2014 Feb; 2(1): 16–20.

Claims

1. Method for determining whether a subject is at risk to develop cancer, wherein the method comprises:

determining in a colon mucosa sample from a subject the expression levels of at least three genes associated with chromosomal segregation; and determining the risk of cancer using the determined expression levels of the genes.

2. The method according to claim 1, wherein the expression levels of the genes associated with chromosomal segregation in the sample when compared to a normal level are indicative of said subject being at risk to develop cancer.

3. The method according to claim 1 or 2, wherein the expression levels of said genes associated with chromosomal segregation in the subject are compared to normal expression levels of the same genes, and a decrease in the expression levels of said genes in the sample relative to the normal expression level is indicative of an increased risk of developing cancer.

4. The method according to any one of claims 1-3, wherein the expression levels of at least three genes associated with chromosomal segregation *Bub1* (BUB1, mitotic checkpoint serine/threonine kinase), *Mis18a* (MIS18 kinetochore protein A) and *Tpx2* (TPX2 microtubule associated) are determined.

5. The method according to any one of claims 1-4, wherein the expression levels of at least four genes associated with chromosomal segregation *Bub1*, *Mis18a*, *Tpx2* and *Rad9a* (RAD9 checkpoint clamp component A), or *Bub1*, *Mis18a*, *Tpx2* and *Pms2* (postmeiotic segregation increased 2) are determined.

6. The method according to any one of claims 1-5, wherein the expression levels of at least five genes associated with chromosomal segregation *Bub1*, *Mis18a*, *Tpx2*, *Rad9a* and *Pms2* are determined.

7. The method according to any one of claims 1-6, wherein the expression levels of at least ten genes associated with chromosomal segregation *Bub1* (BUB1, mitotic checkpoint serine/threonine kinase), *Mis18a* (MIS18 kinetochore protein A), *Tpx2* (TPX2 microtubule associated), *Rad9a* (RAD9 checkpoint clamp component A), *Pms2* (postmeiotic segregation increased 2), *Mlh1* (mutL homologue 1), *Cenpe* (centromere protein E), *Ncapd3* (non-SMC condensing II complex subunit D3),

Odf2 (outer dens fiber of sperm tails 2) and *Dclre1b* (DNA cross-link repair 1B) are determined.

8. The method according to any one of claims 1-7, wherein at least three, four, five, six, seven, eight, nine or ten genes associated with chromosomal segregation are selected from the group consisting of *Bub1*, *Mis18a*, *Tpx2*, *Rad9a*, *Pms2*, *Mlh1*, *Cenpe*, *Ncapd3*, *Odf2* and *Dclre1b*.

9. The method according to any one of claims 1 – 8, wherein the sample is a histologically normal colon mucosa sample.

10. The method according to any one of claims 1 – 9, wherein the sample is a mucosa sample of the proximal colon.

11. The method according to any one of claims 1 – 10, wherein the cancer is a carcinoma.

12. The method according to any one of claims 1 – 11, wherein the cancer is a colon carcinoma or a colon carcinoma of the proximal colon.

13. A kit for use in a method according to any one of claims 1-12 comprising tools to determine the expression levels of at least three genes associated with chromosomal segregation and optionally one or more control samples, and/or optionally reagents for performing said method.

14. The kit for use according to claim 13 comprising tools to determine the expression levels of at least three genes associated with chromosomal segregation *Bub1*, *Mis18a* and *Tpx2*.

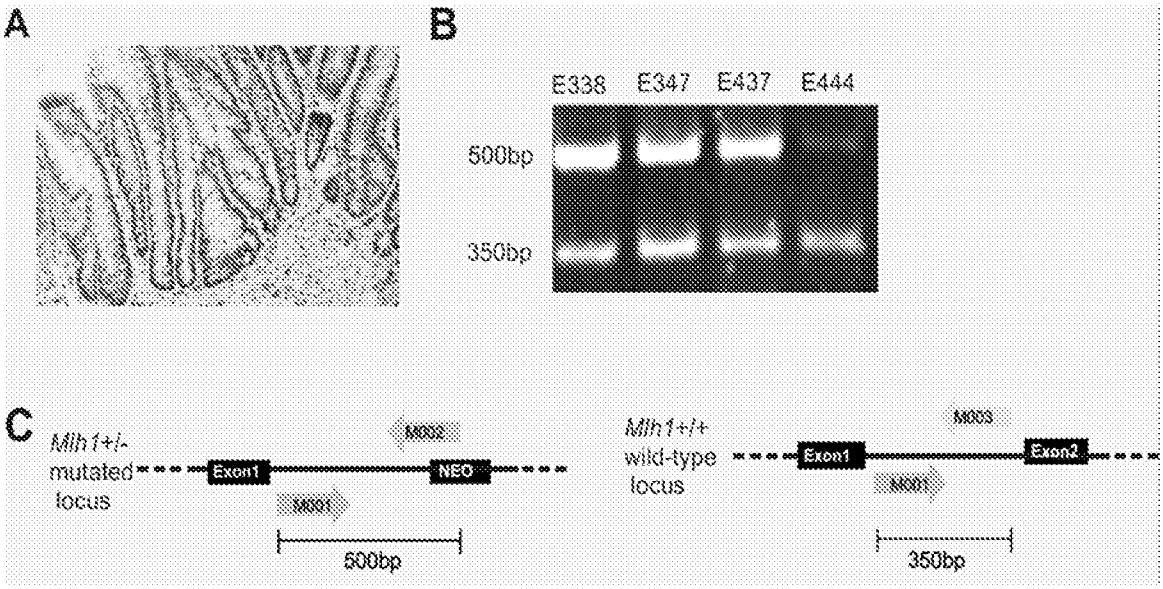
15. The kit for use according to any one of claims 13-14 comprising tools to determine the expression levels of at least four genes associated with chromosomal segregation *Bub1*, *Mis18a*, *Tpx2* and *Rad9a*, or *Bub1*, *Mis18a*, *Tpx2* and *Pms2*.

16. The kit for use according to any one of claims 13-15 comprising tools to determine the expression levels of at least five genes associated with chromosomal segregation *Bub1*, *Mis18a*, *Tpx2*, *Rad9a* and *Pms2*.

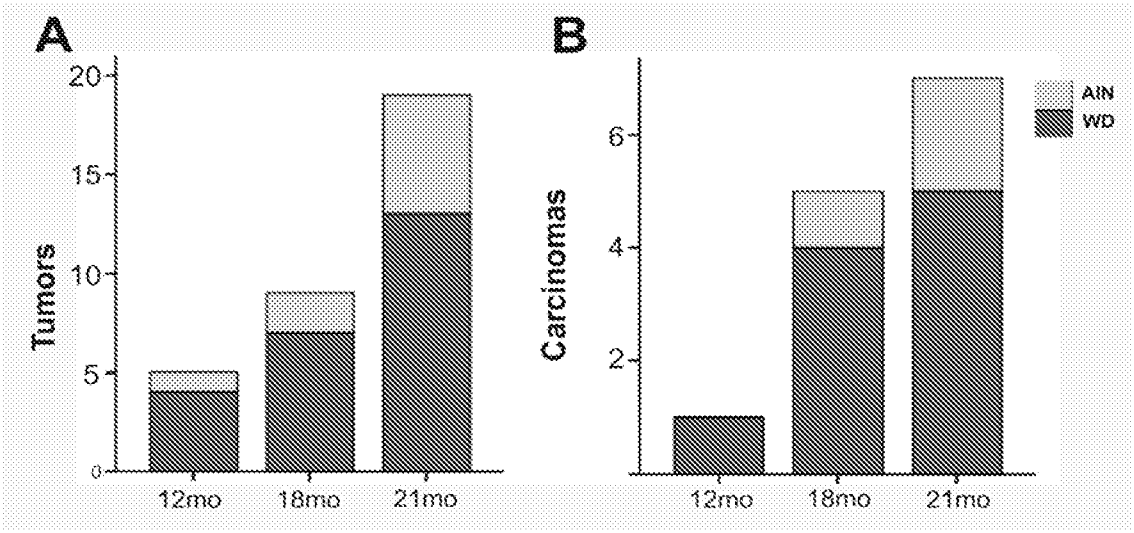
17. The kit for use according to any one of claims 13-16 comprising tools to determine the expression levels of at least ten genes associated with chromosomal segregation *Bub1*, *Mis18a*, *Tpx2*, *Rad9a*, *Pms2*, *Mlh1*, *Cenpe*, *Ncapd3*, *Odf2* and *Dclre1b*.

18. The kit for use according to any one of claims 13-17 comprising tools to determine the expression levels of at least three, four, five, six, seven, eight, nine or ten genes associated with chromosomal segregation selected from the group consisting of *Bub1*, *Mis18a*, *Tpx2*, *Rad9a*, *Pms2*, *Mlh1*, *Cenpe*, *Ncapd3*, *Odf2* and *Dclre1b*.

19. Use of the kit according to any one of claims 13-18 for determining whether a subject is at risk to develop cancer.



Figures 1A, B and C.



Figures 2A and B.

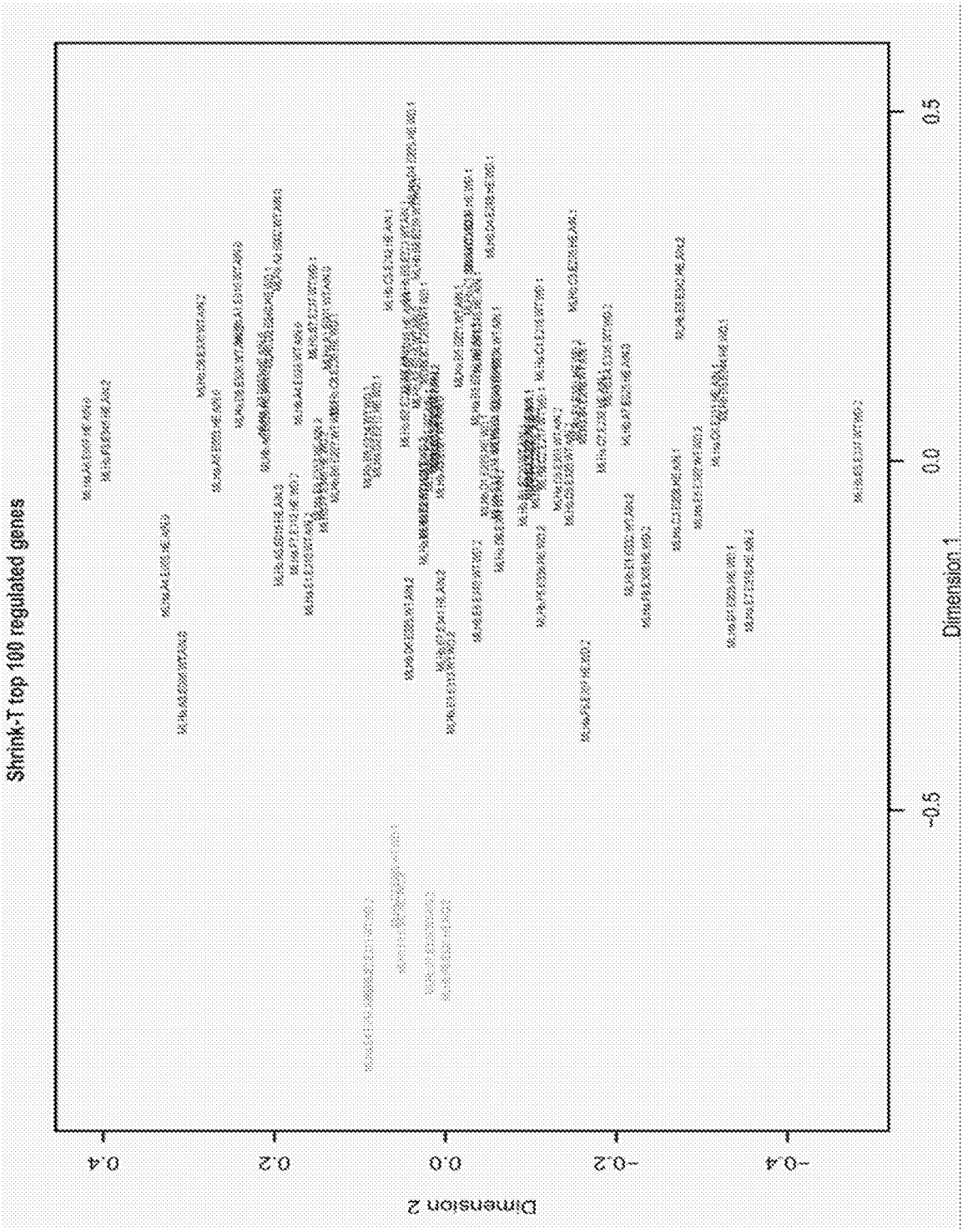

Shrink-T top 100 regulated genes
Dimension 2
0.4
0.2
0.0
-0.2
-0.4
-0.5
0.0
0.5
Dimension 1

Figure 3A.

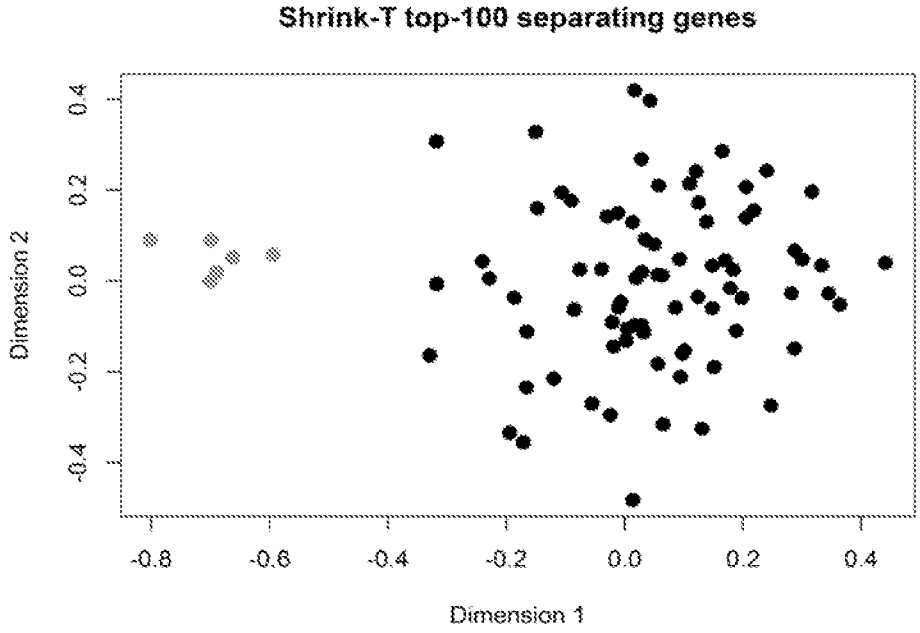
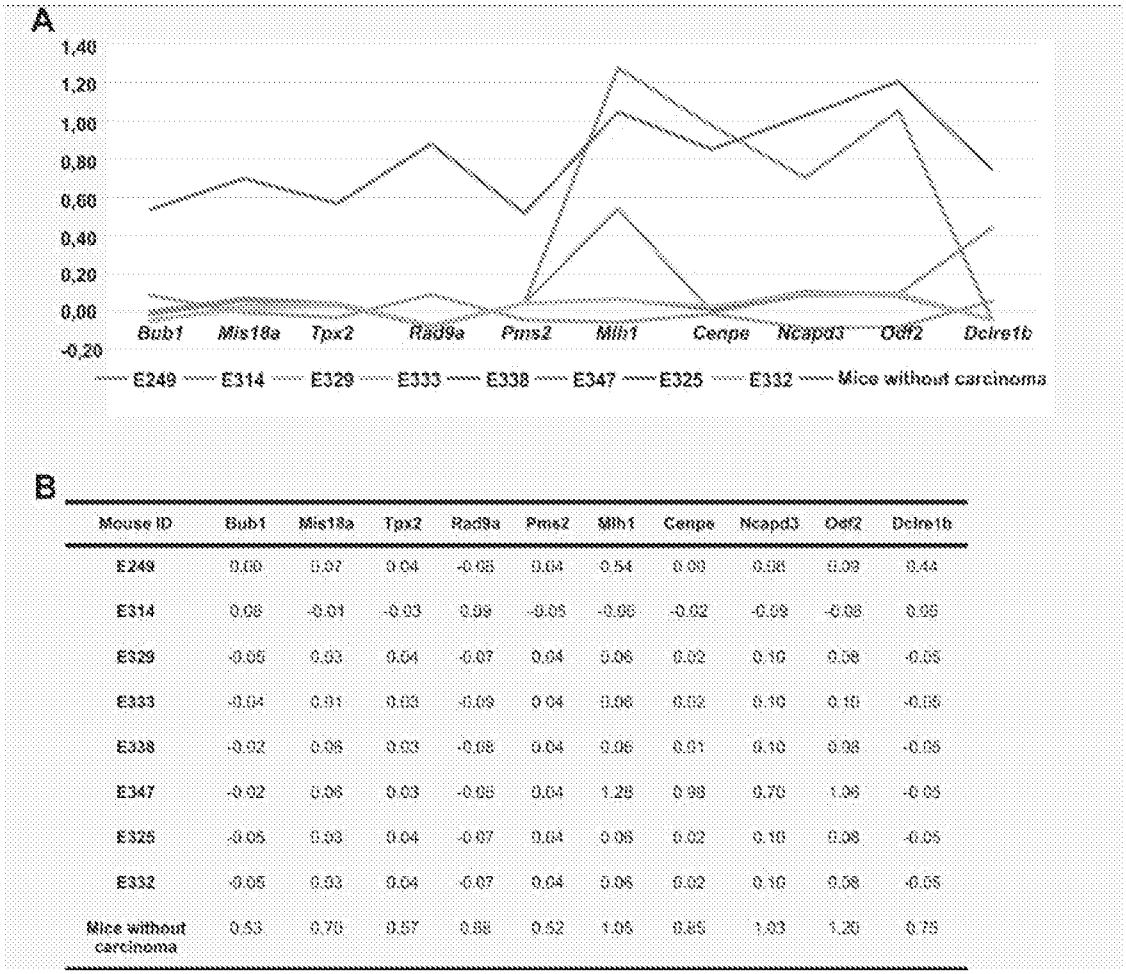


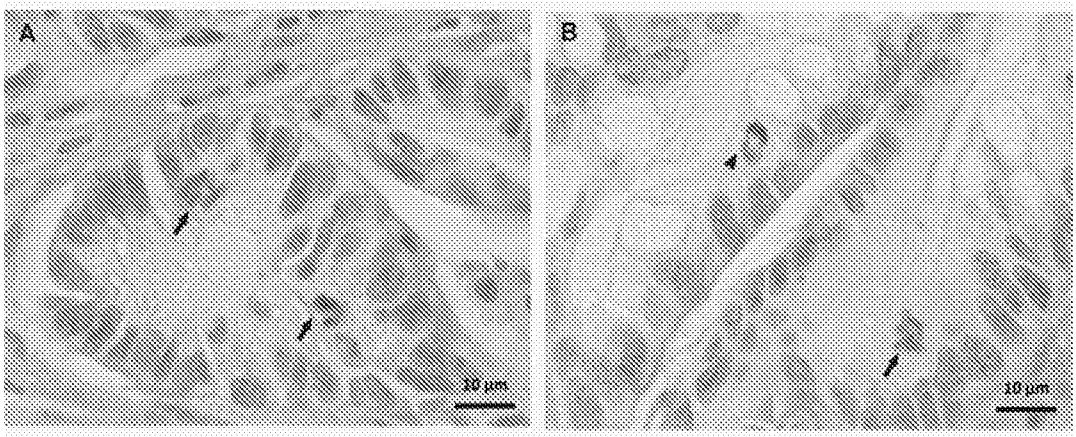
Figure 3B.



B

Mouse ID	Bub1	Mis18a	Tpx2	Rad9a	Pms2	Mlh1	Cenpe	Ncapd3	Odf2	Dclre1b
E249	0.09	0.07	0.04	-0.05	0.04	0.54	0.09	0.08	0.09	0.44
E314	0.08	-0.01	-0.03	0.09	-0.05	-0.06	-0.02	-0.09	-0.08	0.05
E329	-0.05	0.03	0.04	-0.07	0.04	0.06	0.02	0.10	0.08	-0.05
E333	-0.04	0.01	0.03	-0.09	0.04	0.06	0.02	0.10	0.10	-0.05
E338	-0.02	0.06	0.03	-0.08	0.04	0.06	0.01	0.10	0.08	-0.05
E347	-0.02	0.06	0.03	-0.05	0.04	1.28	0.98	0.70	1.06	-0.05
E325	-0.05	0.03	0.04	-0.07	0.04	0.06	0.02	0.10	0.08	-0.05
E332	-0.05	0.03	0.04	-0.07	0.04	0.06	0.02	0.10	0.08	-0.05
Mice without carcinoma	0.53	0.70	0.57	0.88	0.52	1.05	0.85	1.03	1.20	0.75

Figures 4A and B.



Figures 5A and B.

					Tumor characteristics			
ID	Age (months)	Geno type	Diet	Gender	Histopathology	MSI	LOH	Mlh1 expression
E249	12	*Mlh1*$^{+/+}$	WD	M	tubular adenocarcinoma	NA	NA	positive
E314	18	*Mlh1*$^{+/+}$	WD	M	mucinous adenocarcinoma	NA	NA	positive
E329	18	*Mlh1*$^{+/+}$	AIN	F	tubulovillous adenocarcinoma	NA	NA	positive
E333	18	*Mlh1*$^{+/+}$	WD	F	tubular adenocarcinoma	NA	NA	positive
E338	18	*Mlh1*$^{+/-}$	WD	F	tubular adenocarcinoma	MSS	_	positive
E347	18	*Mlh1*$^{+/-}$	WD	F	serrated adenocarcinoma	MSS	_	positive
E402	21	*Mlh1*$^{+/-}$	WD	F	tubular adenocarcinoma	MSS	NA	positive
E409	21	*Mlh1*$^{+/-}$	WD	M	mucinous adenocarcinoma	NA	NA	positive
E410	21	*Mlh1*$^{+/+}$	WD	M	tubular adenocarcinoma	NA	NA	positive
E421	21	*Mlh1*$^{+/+}$	AIN	F	tubular adenocarcinoma	MSS	NA	positive
E437	21	*Mlh1*$^{+/-}$	WD	M	mucinous adenocarcinoma	MSS	_	positive
E444	21	*Mlh1*$^{+/-}$	WD	F	mucinous adenocarcinoma	MSS	_	positive
E453	21	*Mlh1*$^{+/-}$	AIN	F	tubulovillous adenocarcinoma	NA	NA	positive
E214	12	*Mlh1*$^{+/-}$	WD	F	adenoma	NA	NA	NA
E244	12	*Mlh1*$^{+/-}$	WD	F	adenoma	NA	NA	NA
E246	12	*Mlh1*$^{+/-}$	AIN	F	adenoma	NA	NA	NA
E321	18	*Mlh1*$^{+/-}$	WD	F	serrated adenoma	NA	NA	NA
E337	18	*Mlh1*$^{+/+}$	WD	F	adenoma	NA	NA	NA
E402	21	*Mlh1*$^{+/+}$	WD	F	adenoma	NA	NA	NA
E411	21	*Mlh1*$^{+/+}$	WD	F	adenoma	NA	NA	NA
E416	21	*Mlh1*$^{+/+}$	WD	M	adenoma	NA	NA	NA
E446	21	*Mlh1*$^{+/-}$	WD	F	adenoma	NA	NA	NA
E448	21	*Mlh1*$^{+/-}$	AIN	F	adenoma	NA	NA	NA
E217	12	*Mlh1*$^{+/+}$	WD	M	hyperplasia	NA	NA	NA
E307	18	*Mlh1*$^{+/-}$	WD	M	hyperplasia	NA	NA	NA
E328	18	*Mlh1*$^{+/-}$	AIN	F	hyperplasia	NA	NA	NA
E405	21	*Mlh1*$^{+/+}$	WD	F	serrated hyperplasia	NA	NA	NA
E408	21	*Mlh1*$^{+/-}$	AIN	F	hyperplasia	NA	NA	NA
E409	21	*Mlh1*$^{+/+}$	WD	M	hyperplasia	NA	NA	NA
E410	21	*Mlh1*$^{+/+}$	WD	M	hyperplasia	NA	NA	NA
E430	21	*Mlh1*$^{+/-}$	AIN	M	hyperplasia	NA	NA	NA
E433	21	*Mlh1*$^{+/-}$	AIN	F	hyperplasia	NA	NA	NA
E454	21	*Mlh1*$^{+/-}$	WD	F	hyperplasia	NA	NA	NA

WD, Western-style diet; AIN, AIN-93G; M, male; F, female; NA, not available

Figure 6.

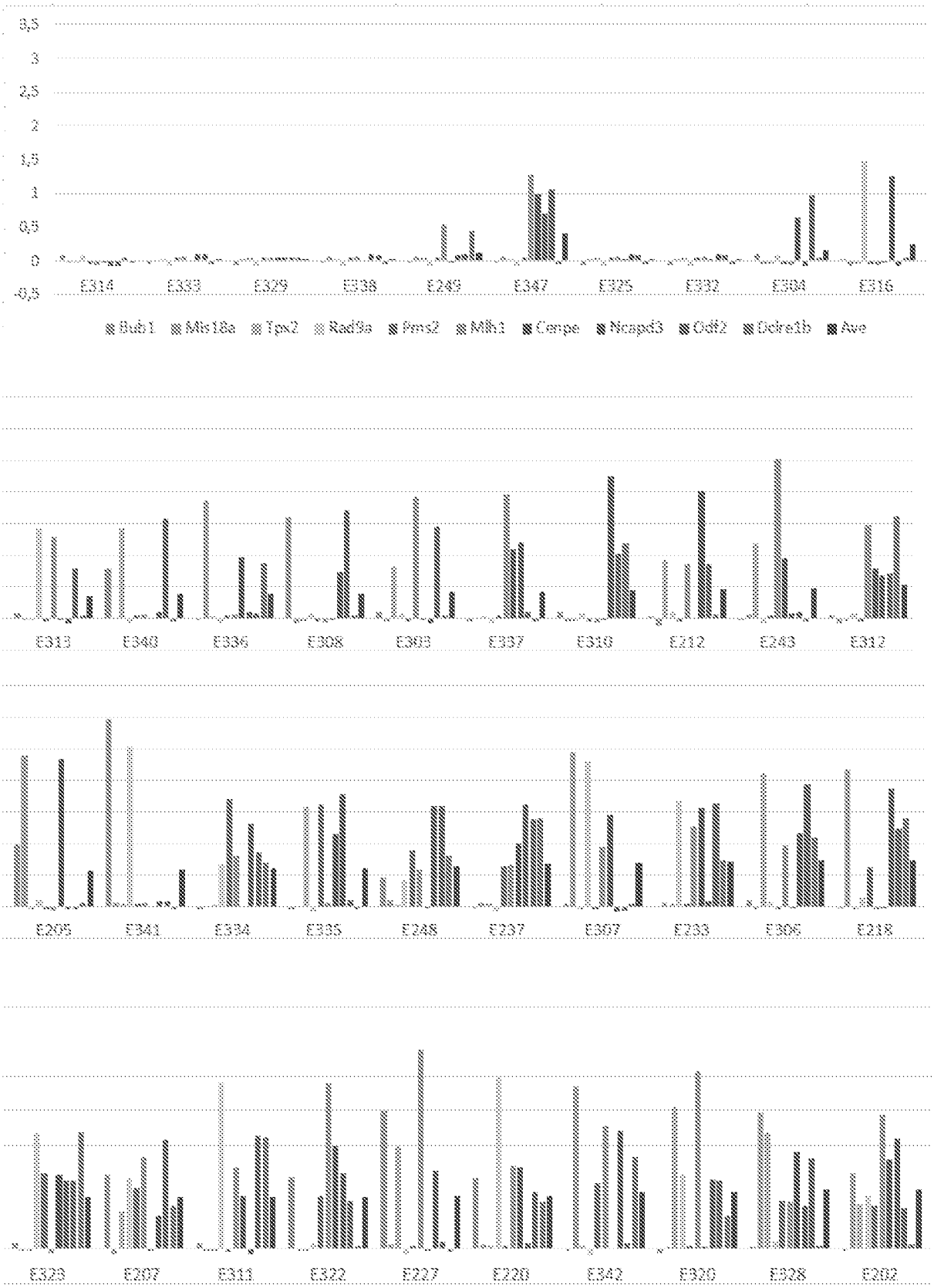


Figure 7. (continues in the next page)

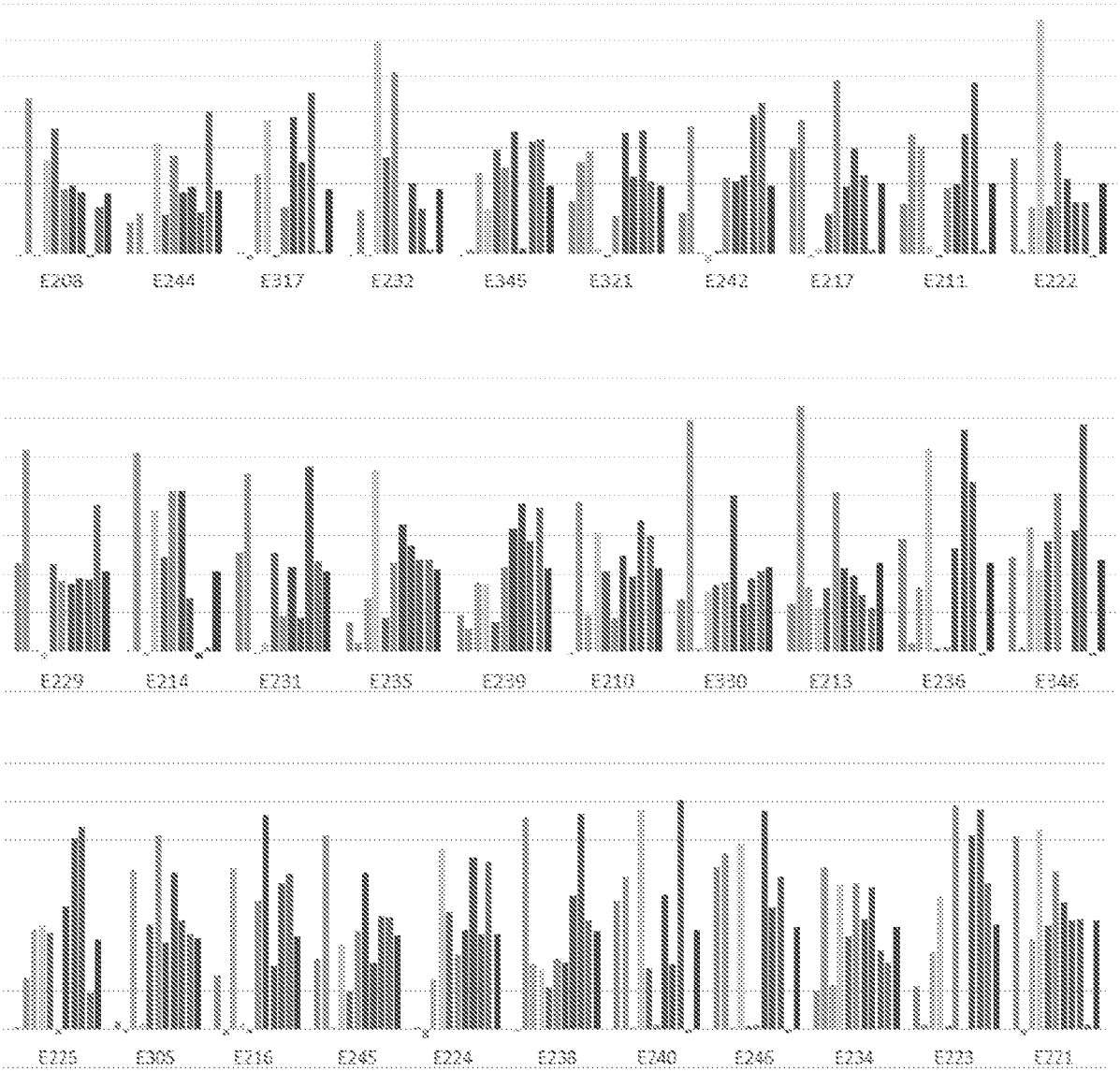


Figure 7. (continues from the earlier page)

Input	MGI.Gene.Marker.ID	Symbol	Name	Feature.Type
Bub1	MGI:1100510	Bub1	budding uninhibited by benzimidazoles 1 homolog (S. cerevisiae)	protein coding gene
Mis18a	MGI:1913828	Mis18a	MIS18 kinetochore protein homolog A (S. pombe)	protein coding gene
Tpx2	MGI:1919369 | | MGI:1342328 | | MGI:1342328	Tpx2 | | Dad | | Dad	TPX2, microtubule-associated protein homolog (Xenopus laevis) | | del	protein coding gene
Rad9a	MGI:1328256 | | MGI:2385231 | | MGI:2385231	Rad9a | | Rad9b | | Rad9b	RAD9 homolog A | | RAD9 homolog B | | RAD9 homolog B	protein coding gene
Pmc2	MGI:104286	Pmc2	postmeiotic segregation increased 2 (S. cerevisiae)	protein coding gene
Mlh1	MGI:1891928	Mlh1	mutL homolog 1 (E. coli)	protein coding gene
Cenpe	MGI:1098230	Cenpe	centromere protein E	protein coding gene
Ncapd3	MGI:2142469	Ncapd3	non-SMC condensin II complex, subunit D3	protein coding gene
Odf2	MGI:1098824 | | MGI:97424 | | MGI:97424	Odf2 | | Odf1 | | Odf1	outer dense fiber of sperm tails 2 | | outer dense fiber of sperm tails 1	protein coding gene
Dclre1b	MGI:2156057	Dclre1b	DNA cross-link repair 1B, PSO2 homolog (S. cerevisiae)	protein coding gene

Figure 8. (continues in the next page)

Ensembl.ID	Entrez.Gene.ID			Cans -No Cans Shrunk T	Cans -No Cans Shrunk P-value	MLHa.A1 E001.W T.AIN.0	MLHa.A2 E002.W T.AIN.0	MLHa.A3 E005.W T.AIN.0
ENSMUSG00000027378			12235	-4,48	0,00	0,06	1,34	0,06
ENSMUSG00000022978			66576	-5,03	0,04	-0,07	0,97	-0,07
ENSMUSG00000027482 || ENSMUSG00000010592 || ENSMUSG00000010592	72119 || 13164 || 13164			-4,47	0,03	1,62	0,55	-0,05
ENSMUSG00000024624 || ENSMUSG00000038569 || ENSMUSG00000038569	19367 || 231724 || 231724			-6,64	0,00	0,09	1,15	0,09
ENSMUSG00000079109			18861	-4,57	0,01	0,48	0,75	-0,05
ENSMUSG00000022498			17350	-3,49	0,03	1,56	1,16	2,06
ENSMUSG00000045318			229641	-4,23	0,02	2,08	1,30	0,90
ENSMUSG00000033024			78658	-5,73	0,00	2,17	1,09	-0,07
ENSMUSG00000026790 || ENSMUSG00000061923 || ENSMUSG00000061923	18236 || 18235 || 18235			-5,01	0,01	0,44	1,67	1,26
ENSMUSG00000027845			140917	-4,66	0,01	0,58	1,00	0,06

MLHa.A4. E003.HE. AIN.0	MLHa.A5. E003.HE. AIN.0	MLHa.A6. E007.HE. AIN.0	MLHa.A7. E013.HE. AIN.0	MLHa.B1. E202.WT. AIN.1	MLHa.B2. E207.WT. AIN.1	MLHa.B3. E210.WT. AIN.1	MLHa.B4. E218.WT. AIN.1	MLHa.B5. E221.WT. AIN.1	MLHa.B7. E213.WT. WD.1	MLHa.B8. E214.WT. WD.1	MLHa.C1. E216.WT. WD.1
0,07	0,94	0,07	0,07	0,00	1,08	0,00	0,00	2,54	0,63	0,02	0,71
-0,05	-0,05	-0,05	-0,05	1,09	-0,08	1,93	2,17	-0,08	3,16	2,55	-0,09
-0,05	0,71	1,82	0,63	0,64	0,54	0,48	-0,04	1,19	0,83	-0,04	2,12
0,69	2,32	0,11	1,13	0,77	1,02	1,53	0,13	2,63	0,58	1,82	0,08
-0,05	-0,05	-0,05	0,62	0,62	0,88	1,04	0,63	1,37	0,82	1,22	-0,05
1,54	1,08	-0,06	1,25	1,95	1,33	0,45	-0,05	2,08	2,05	2,07	1,69
0,00	0,00	1,58	0,00	1,29	0,00	1,24	0,00	1,66	1,08	2,08	2,83
1,12	0,65	1,83	2,32	1,59	0,48	0,97	1,88	1,44	0,98	0,70	0,84
1,45	-0,10	-0,10	0,90	0,59	1,58	1,69	1,24	1,45	0,73	-0,08	1,93
0,05	1,21	0,05	0,05	0,06	0,62	1,50	1,41	0,06	0,57	0,06	2,05

MLHa.C2. E217.WT. WD.1	MLHa.C3. E208.HE. AIN.1	MLHa.C5. E211.HE. AIN.1	MLHa.C6. E231.HE. AIN.1	MLHa.C7. E232.HE. AIN.1	MLHa.D1. E205.HE. WD.1	MLHa.D2. E212.HE. WD.1	MLHa.D3. E224.HE. WD.1	MLHa.D4. E225.HE. WD.1	MLHa.D5. E303.WT. AIN.2	MLHa.D6. E304.WT. AIN.2	MLHa.D7. E305.WT. AIN.2
1,49	-0,03	0,71	1,28	-0,03	0,99	0,02	0,02	0,02	0,10	0,10	0,10
1,88	2,20	1,69	2,28	0,63	2,40	-0,13	-0,13	0,68	-0,04	-0,04	-0,04
-0,04	-0,04	1,50	-0,04	-0,04	-0,04	0,93	0,65	1,30	0,82	-0,04	2,10
0,08	1,33	0,12	0,12	2,98	0,10	0,10	2,38	1,37	0,07	0,07	0,07
0,57	1,76	-0,05	1,28	1,37	-0,05	-0,05	1,54	1,27	-0,04	-0,04	1,38
2,45	0,91	0,94	0,45	2,56	-0,06	0,86	0,99	-0,06	1,93	-0,06	2,56
0,95	0,96	0,99	1,09	0,01	2,33	0,01	1,31	1,61	-0,02	0,84	1,14
1,49	0,88	1,68	0,44	1,01	-0,04	2,02	2,27	2,52	-0,08	-0,08	2,06
1,11	-0,04	2,41	2,37	0,63	-0,04	0,86	1,25	2,66	1,45	0,97	1,44
0,06	0,66	0,06	1,16	0,06	0,06	0,06	2,20	0,49	0,05	0,05	1,26

Figure 8. (continues from the earlier page, continues in the next page)

MLHa.D8. E306.WT. AIN.2	MLHa.E1. E310.WT. AIN.2	MLHa.E2. E311.WT. WD.2	MLHa.E3. E313.WT. WD.2	MLHa.E4. E314.WT. WD.2	MLHa.E5. E322.WT. WD.2	MLHa.E6. E323.WT. WD.2	MLHa.E7. E316.HE. AIN.2	MLHa.E8. E317.HE. AIN.2	MLHa.F3. E328.HE. AIN.2	MLHa.F5. E307.HE. WD.2	MLHa.F6. E308.HE. WD.2
0,10	0,10	0,08	0,08	0,08	1,04	0,08	0,02	0,02	0,02	0,04	1,60
-0,04	-0,04	-0,01	-0,01	-0,01	-0,01	-0,01	-0,08	-0,08	1,98	2,45	-0,07
2,12	-0,04	-0,03	-0,03	-0,03	-0,03	-0,03	-0,04	1,13	1,69	-0,04	-0,04
0,07	0,07	2,42	1,42	0,09	0,09	1,67	1,47	1,89	0,10	2,30	0,08
-0,04	-0,04	-0,05	-0,05	-0,05	0,77	1,09	-0,05	-0,05	0,70	-0,04	-0,04
0,96	-0,06	1,18	1,29	-0,06	2,40	-0,06	-0,06	0,65	0,67	0,85	-0,06
-0,02	-0,02	0,77	-0,02	-0,02	1,48	1,07	0,00	1,92	1,41	1,46	-0,01
1,17	2,24	-0,09	-0,09	-0,09	1,10	0,98	1,26	1,28	0,83	-0,09	0,73
1,34	1,02	1,63	0,78	-0,08	0,69	0,99	-0,07	2,26	1,32	-0,07	1,71
1,09	1,19	1,61	0,05	0,05	0,05	1,68	0,05	0,05	0,05	0,05	0,05

MLHa.F7. E312.HE. WD.2	MLHa.F8. E321.HE. WD.2	MLHb.A1. E010.WT. AIN.0	MLHb.A2. E018.WT. AIN.0	MLHb.A3. E021.WT. AIN.0	MLHb.A4. E024.WT. AIN.0	MLHb.A5. E015.HE. AIN.0	MLHb.A6. E020.HE. AIN.0	MLHb.A7. E025.HE. AIN.0	MLHb.B1. E220.WT. AIN.1	MLHb.B2. E222.WT. AIN.1	MLHb.B3. E223.WT. AIN.1
0,04	0,75	-0,02	-0,02	-0,02	1,32	-0,04	0,72	-0,04	1,02	1,35	0,57
-0,07	1,29	0,06	0,84	2,76	0,06	0,80	0,04	1,99	0,06	0,06	0,06
-0,04	1,46	0,05	0,05	2,15	2,76	0,85	0,05	2,22	0,04	0,65	1,02
0,08	0,08	1,86	-0,09	-0,09	1,95	2,12	3,05	-0,11	2,49	3,29	1,77
-0,04	-0,04	0,99	0,04	0,04	0,04	0,04	0,04	1,66	0,05	0,67	0,05
1,48	0,53	1,90	1,92	1,94	1,08	0,06	0,93	0,06	1,20	1,58	2,96
0,78	1,70	-0,01	1,79	0,76	2,36	1,64	1,69	1,89	1,18	1,06	0,00
0,69	1,09	0,91	1,77	0,08	0,75	0,93	0,97	1,25	0,07	0,74	2,55
0,71	1,75	0,69	0,96	1,59	2,00	0,11	0,11	0,11	0,82	0,72	2,90
1,61	1,02	1,18	-0,05	1,11	1,93	-0,05	1,27	-0,05	0,67	-0,05	1,92

MLHb.B4. E233.WT. AIN.1	MLHb.B5. E234.WT. AIN.1	MLHb.B6. E227.WT. WD.1	MLHb.B7. E237.WT. WD.1	MLHb.B8. E239.WT. WD.1	MLHb.C1. E243.WT. WD.1	MLHb.C2. E249.WT. WD.1	MLHb.C3. E235.HE. AIN.1	MLHb.C4. E236.HE. AIN.1	MLHb.C5. E242.HE. AIN.1	MLHb.C6. E245.HE. AIN.1	MLHb.C7. E246.HE. AIN.1
0,02	0,52	2,00	0,00	0,47	0,00	0,00	0,38	1,45	0,58	0,93	2,15
0,06	2,13	0,07	0,07	0,30	0,07	0,07	0,12	0,12	1,80	2,56	2,33
0,04	0,59	1,49	0,04	0,90	1,19	0,04	0,69	0,85	0,03	0,03	0,03
1,66	1,90	-0,08	-0,08	0,88	-0,08	-0,08	2,33	2,62	-0,11	1,14	2,45
0,05	1,23	0,04	0,64	0,39	0,04	0,04	0,44	0,05	0,05	0,50	0,05
1,27	1,92	2,89	0,65	1,10	2,52	0,54	1,15	0,06	1,07	1,29	0,06
1,56	1,45	0,00	1,00	1,59	0,94	0,00	1,63	1,32	1,03	2,08	2,89
0,07	1,88	1,13	1,61	1,90	0,08	0,08	1,36	2,85	1,12	0,89	1,60
1,63	1,03	0,09	1,39	1,41	0,09	0,09	1,18	2,17	1,96	1,49	2,01
0,72	0,88	-0,05	1,41	1,84	-0,05	0,44	1,19	-0,05	2,12	1,47	-0,05

Figure 8. (continues from the earlier page, continues in the next page)

MLHb.C8. E238.HE. WD.1	MLHb.D1. E229.HE. WD.1	MLHb.D2. E240.HE. WD.1	MLHb.D3. E244.HE. WD.1	MLHb.D4. E248.HE. WD.1	MLHb.D5. E320.WT. AIN.2	MLHb.D6. E325.WT. AIN.2	MLHb.D7. E329.WT. AIN.2	MLHb.D8. E330.WT. AIN.2	MLHb.E1. E332.WT. AIN.2	MLHb.E2. E333.WT. WD.2	MLHb.E3. E334.WT. WD.2
0,00	1,14	1,69	0,45	0,45	-0,05	-0,05	-0,05	0,68	-0,05	-0,04	-0,04
2,80	2,60	2,01	0,57	0,10	0,03	0,03	0,03	2,98	0,03	0,01	0,01
0,87	0,03	0,03	0,03	0,03	2,06	0,04	0,04	0,04	0,04	0,03	0,03
0,78	-0,10	2,89	1,56	0,43	1,08	-0,07	-0,07	0,78	-0,07	-0,09	0,67
0,56	1,13	0,80	0,55	0,89	0,04	0,04	0,04	0,86	0,04	0,04	1,71
0,93	0,92	0,06	1,38	0,56	2,58	0,06	0,06	0,90	0,06	0,06	0,81
0,88	0,87	1,78	0,88	0,00	0,02	0,02	0,02	2,01	0,02	0,02	0,02
1,76	0,94	0,86	0,95	1,80	1,01	0,10	0,10	0,62	0,10	0,10	1,32
2,85	0,94	3,03	0,59	1,59	0,99	0,08	0,08	0,95	0,08	0,10	0,87
1,43	1,89	-0,05	1,99	0,80	0,48	-0,05	-0,05	1,03	-0,05	-0,05	0,69

MLHb.E4. E335.WT. WD.2	MLHb.E5. E337.WT. WD.2	MLHb.E6. E340.WT. WD.2	MLHb.E7. E341.HE. AIN.2	MLHb.E8. E342.HE. AIN.2	MLHb.F3. E345.HE. AIN.2	MLHb.F5. E336.HE. WD.2	MLHb.F6. E338.HE. WD.2	MLHb.F7. E346.HE. WD.2	MLHb.F8. E347.HE. WD.2
-0,04	-0,04	0,79	2,97	0,00	0,00	-0,02	-0,02	1,21	-0,02
0,01	0,01	0,01	0,06	2,36	0,06	1,87	0,06	0,06	0,06
1,58	0,03	1,42	0,04	0,04	1,16	0,03	0,03	1,60	0,03
-0,09	-0,09	-0,09	2,53	-0,10	0,65	-0,08	-0,08	1,05	-0,08
1,62	0,04	0,04	0,04	0,94	1,47	0,04	0,04	1,41	0,04
0,06	1,96	0,06	0,06	1,78	1,22	0,06	0,06	2,03	1,28
1,14	1,10	0,02	0,01	0,01	1,73	0,97	0,01	0,01	0,98
1,77	1,20	0,10	0,09	1,71	0,09	0,10	0,10	1,56	0,70
0,10	0,10	1,58	0,09	0,09	1,59	0,08	0,08	2,91	1,06
-0,05	-0,05	-0,05	-0,05	1,33	1,62	0,88	-0,05	-0,05	-0,05

Figure 8. (continues from the earlier page)

INTERNATIONAL SEARCH REPORT

International application No
PCT/FI2018/050419

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

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

B. FIELDS SEARCHED

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

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

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	PINO M S ET AL: "The Chromosomal Instability Pathway in Colon Cancer", GASTROENTEROLOGY, W.B. SAUNDERS CO, US, vol. 138, no. 6, 1 May 2010 (2010-05-01), pages 2059-2072, XP027032386, ISSN: 0016-5085 [retrieved on 2010-04-23] the whole document	1-3,13, 19
X	WO 2007/082099 A2 (GENOMIC HEALTH INC [US]; NASBP FOUNDATION INC [US]; COWENS WAYNE [US];) 19 July 2007 (2007-07-19) the whole document	1-3, 11-13,19
A	WO 2012/129008 A1 (BAYLOR RES INST [US]; GOEL AJAY [US]; BOLAND C RICHARD [US]; BALAGUER) 27 September 2012 (2012-09-27) the whole document	1-19

-/--

[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
19 July 2018	30/07/2018

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

1

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

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2015/121465 A1 (VENTANA MED SYST INC [US]; UNIV CALIFORNIA [US]; US DEPT VETERANS AFFA) 20 August 2015 (2015-08-20) the whole document	1-19
X	WO 2004/061423 A2 (WYETH CORP [US]; MARTINEZ ROBERT VINCENT [US]; BROWN EUGENE [US]; LIU) 22 July 2004 (2004-07-22) the whole document	1-3, 11-13,19
Y	WO 99/47638 A2 (UNIV JOHNS HOPKINS MED [US]; VOGELSTEIN BERT [US]; KINZLER KENNETH W []) 23 September 1999 (1999-09-23) the whole document	1-19
Y	PING WEI ET AL: "TPX2 is a novel prognostic marker for the growth and metastasis of colon cancer", JOURNAL OF TRANSLATIONAL MEDICINE, BIOMED CENTRAL, vol. 11, no. 1, 17 December 2013 (2013-12-17), page 313, XP021171385, ISSN: 1479-5876, DOI: 10.1186/1479-5876-11-313 the whole document	1-19
Y	SERGIO ALONSO ET AL: "Methylation of MGMT and ADAMTS14 in normal colon MUCOSA: biomarkers of a field defect for cancerization preferentially targeting elder African-Americans", ONCOTARGET, vol. 6, no. 5, 20 February 2015 (2015-02-20), XP055493727, DOI: 10.18632/oncotarget.2852 the whole document	1-19

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/FI2018/050419

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2007082099 A2	19-07-2007	AU 2007204826 A1	19-07-2007
		BR PI0706511 A2	29-03-2011
		CA 2636984 A1	19-07-2007
		DK 1974058 T3	01-09-2014
		EP 1974058 A2	01-10-2008
		EP 2377950 A1	19-10-2011
		EP 2400036 A1	28-12-2011
		EP 2407553 A1	18-01-2012
		EP 2412820 A1	01-02-2012
		EP 2412821 A1	01-02-2012
		EP 2412822 A1	01-02-2012
		EP 2412823 A1	01-02-2012
		EP 2412824 A1	01-02-2012
		ES 2491222 T3	05-09-2014
		HK 1120089 A1	06-03-2015
		IL 192794 A	31-07-2013
		JP 5297202 B2	25-09-2013
		JP 5486664 B2	07-05-2014
		JP 5486717 B2	07-05-2014
		JP 5486718 B2	07-05-2014
		JP 5486719 B2	07-05-2014
		JP 5486720 B2	07-05-2014
		JP 2009523028 A	18-06-2009
		JP 2013074894 A	25-04-2013
		JP 2013176398 A	09-09-2013
		JP 2013215201 A	24-10-2013
		JP 2013223503 A	31-10-2013
		JP 2013226150 A	07-11-2013
		JP 2013226151 A	07-11-2013
		KR 20090003178 A	09-01-2009
		NZ 569788 A	26-08-2011
		NZ 593224 A	26-10-2012
		NZ 593225 A	26-10-2012
		NZ 593226 A	26-10-2012
		NZ 593227 A	26-10-2012
		NZ 593228 A	26-10-2012
		NZ 593229 A	26-10-2012
		US 2009291434 A1	26-11-2009
		US 2010190173 A1	29-07-2010
		US 2011039269 A1	17-02-2011
		US 2011039270 A1	17-02-2011
		US 2011039271 A1	17-02-2011
		US 2011039272 A1	17-02-2011
		US 2011097759 A1	28-04-2011
		US 2011111421 A1	12-05-2011
		US 2012171688 A1	05-07-2012
		US 2013102492 A1	25-04-2013
		WO 2007082099 A2	19-07-2007
WO 2012129008 A1	27-09-2012	AR 085818 A1	30-10-2013
		CA 2830329 A1	27-09-2012
		EP 2686451 A1	22-01-2014
		TW 201243331 A	01-11-2012
		US 2012238463 A1	20-09-2012
		WO 2012129008 A1	27-09-2012
WO 2015121465 A1	20-08-2015	AU 2015216886 A1	25-08-2016
		CA 2939697 A1	20-08-2015

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/FI2018/050419

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
		EP 3108254 A1	28-12-2016
		JP 2017508147 A	23-03-2017
		US 2017038385 A1	09-02-2017
		WO 2015121465 A1	20-08-2015
WO 2004061423 A2	22-07-2004	AT 401575 T	15-08-2008
		AU 2004203749 A1	22-07-2004
		CA 2511907 A1	22-07-2004
		DK 1581812 T3	20-10-2008
		EP 1581812 A2	05-10-2005
		EP 1998177 A2	03-12-2008
		ES 2309485 T3	16-12-2008
		PT 1581812 E	22-09-2008
		SI 1581812 T1	31-10-2008
		US 2004265230 A1	30-12-2004
		WO 2004061423 A2	22-07-2004
WO 9947638 A2	23-09-1999	AU 3187399 A	11-10-1999
		WO 9947638 A2	23-09-1999